UNIVERSITI TEKNOLOGI MARA

DEVELOPMENT OF A WARRANTY MANAGEMENT SYSTEM MODEL USING DATA-DRIVEN STRATEGIES FOR THE MALAYSIAN AUTOMOTIVE INDUSTRY

NOR HISSHAM BIN ABDUL HAMID

PhD

June 2026

UNIVERSITI TEKNOLOGI MARA

DEVELOPMENT OF A WARRANTY MANAGEMENT SYSTEM MODEL USING DATA-DRIVEN STRATEGIES FOR THE MALAYSIAN AUTOMOTIVE INDUSTRY

NOR HISSHAM BIN ABDUL HAMID

Thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)

Faculty of Mechanical Engineering

June 2026

CONFIRMATION BY PANEL OF EXAMINERS

I certify that a Panel of Examiners has met on 22 April 2026 to conduct the final examination of Nor Hissham Bin Abdul Hamid on his Doctor of Philosophy thesis entitled "Development of a Warranty Management System Model Using Data-Driven Strategies for the Malaysian Automotive Industry" in accordance with Universiti Teknologi MARA Act 1976 (Akta 173). The Panel of Examiners recommends that the student be awarded the relevant degree. The Panel of Examiners was as follows:

Johnson Lim Soon Chong, PhD
Associate Professor
Faculty of Mechanical Engineering
Universiti Teknologi MARA
(Chairman)

Amirul Abdul Rashid, PhD
Associate Professor
Faculty of Mechanical Engineering
Universiti Teknologi MARA
(Internal Examiner)

Mohd Shahir Kasim, PhD
Professor
Fakulti Reka Bentuk Inovatif dan Teknologi
Universiti Sultan Zainal Abidin
(External Examiner)

PROFESSOR DR HJH ZURAEDA IBRAHIM
Dean
Institute of Postgraduates Studies
Universiti Teknologi MARA
Date: 30 June 2026

AUTHOR'S DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of Universiti Teknologi MARA. It is original and is the results of my own work, unless otherwise indicated or acknowledged as referenced work. This thesis has not been submitted to any other academic institution or non-academic institution for any degree or qualification.

I, hereby, acknowledge that I have been supplied with the Academic Rules and Regulations for Post Graduate, Universiti Teknologi MARA, regulating the conduct of my study and research.

Name of Student	:	Nor Hissham Bin Abdul Hamid
Student I.D. No.	:	2023686556
Programme	:	Doctor of Philosophy (Mechanical Engineering) – CEEM950
Faculty	:	Mechanical Engineering
Thesis Title	:	Development of a Warranty Management System Model Using Data-Driven Strategies for the Malaysian Automotive Industry
Signature of Student	:	……
Date	:	June 2026

ABSTRACT

The increasing complexity of automotive systems, combined with large volumes of heterogeneous warranty data, presents significant challenges in achieving timely fault detection and effective warranty cost control. Conventional warranty management approaches remain predominantly reactive and are often constrained by data inconsistency, delayed reporting, and limited predictive capability. This study addresses these limitations through the development of a data-driven warranty management system model tailored to the Malaysian automotive industry, with a focused case study on Alternator Assembly components supplied by local company (Company MNC) to local car manufacturing company (Company A). A total of eight years of real-world warranty data were systematically managed through structured extraction, cleansing, validation, and classification processes to ensure data reliability and analytical consistency. Particular attention was given to resolving tangible data issues, including missing entries, duplicate claims, inconsistent failure coding, and non-trouble-found (NTF) cases arising from misdiagnosis. These corrections were essential to minimise bias and improve the robustness of subsequent analyses. Advanced analytical techniques, including statistical analysis, machine learning models, and data mining methods, were applied to identify failure patterns, temporal trends, and anomaly behaviours. Predictive modelling was validated using historical datasets, achieving acceptable performance with coefficient of determination ($R^2 \geq 0.60$), demonstrating the capability of the proposed approach to handle real-world, imbalanced warranty data. However, the study also recognises inherent uncertainties influenced by external factors such as variations in service centre diagnostic practices, environmental operating conditions, and customer usage behaviour, which may affect model generalisation. In parallel, root cause analysis and component repair feasibility studies were conducted to evaluate cost-effective alternatives to full component replacement, supporting more targeted and sustainable warranty strategies. These analytical outputs were integrated into the Automotive Warranty Analysis System (AWAS), an interactive decision-support platform developed using Microsoft Power BI, enabling enhanced data transparency, early failure detection, and real-time monitoring. The results demonstrate measurable improvements in warranty management performance, including enhanced identification of high-risk components, reduction in unnecessary replacements through targeted repair strategies, and improved decision-making efficiency. The proposed framework also exhibits scalability and adaptability for broader industrial applications. This study contributes to the advancement of automotive warranty management by integrating predictive analytics, data governance, and engineering decision-making into a unified system. The findings support a transition from reactive to predictive warranty strategies, ultimately improving product reliability, reducing operational costs, and enhancing customer satisfaction.

ACKNOWLEDGEMENT

First and foremost, I extend my heartfelt gratitude to God for granting me the opportunity to pursue my doctoral studies and for providing the strength and perseverance to successfully complete this rigorous journey. Without His blessings, this achievement would not have been possible. I am profoundly indebted to Associate Professor Ir Ts Dr. Nor Fazli Bin Adull Manan, my supervisor, for his unwavering support, insightful guidance, and constructive feedback throughout the course of my research. His expertise and encouragement have been instrumental in shaping this thesis. My gratitude extends to my colleagues in the Sales Quality Engineering Department for their dedicated assistance and support in various aspects of this research. Furthermore, I am deeply grateful to my fellow PhD candidates, Ts. Mohd Fauzi Ismail, Ms. Femmay Jika, and Ts. Tajul Farhan Tamimuddin, for their friendship, collaboration, and moral support. Their camaraderie has been a source of great strength and motivation throughout this journey. I would also like to acknowledge the invaluable assistance and resources provided by Universiti Teknologi MARA (UiTM), which have greatly facilitated the successful completion of this thesis. The academic and administrative staff at UiTM have been exceptionally supportive and accommodating, for which I am truly thankful. Lastly, this thesis is dedicated to my late father, Allahyarham Haji Abdul Hamid Bin Taib, my late mother, Allahyarhamah Hajah [redacted], my supportive wife, Noranizah Binti Mohd Sa'at, my sons, Muhammad Danish Irfan and Muhammad Danial Iman, my precious daughter, Daryn Aisya, and my extended family. Their unwavering support, patience, and encouragement have been the cornerstone of my academic journey. This achievement is a testament to their love and dedication. Alhamdulillah.

TABLE OF CONTENTS

		Page
CONFIRMATION BY PANEL OF EXAMINERS		**ii**
AUTHOR'S DECLARATION		**iii**
ABSTRACT		**iv**
ACKNOWLEDGEMENT		**v**
TABLE OF CONTENTS		**vi**
LIST OF TABLES		**xi**
LIST OF FIGURES		**xiii**
LIST OF ABBREVIATIONS		**xvi**
CHAPTER 1 INTRODUCTION		**1**
1.1	Research Background	1
1.2	Problem Statement	2
1.3	Objectives	7
1.4	Research Scope and Limitation	7
1.5	Significant of Research	9
CHAPTER 2 LITERATURE REVIEW		**12**
2.1	Introduction	12
2.2	Company Overview	13
2.3	Plant Locations and Products	16
	2.3.1 Electronics Plant	17
	2.3.2 Vehicle System Plant	18
2.4	Malaysian Market Overview and Warranty Data Visualization	18
2.5	Predictive Warranty System Model	22
2.6	Regression Model	24
2.7	Validate The Model for Component Repair	25
2.8	Benchmarking Analysis	28
2.9	Automotive Warranty Cost Forecasting Overview	32

2.9.1 Industry Context and Economic Impact 32
2.9.2 Traditional Forecasting Approaches 34
2.9.3 Evolution Towards Advanced Analytics 35
2.10 Statistical Methods for Warranty Prediction 36
2.10.1 Exploratory Data Analysis (EDA) 36
2.10.2 Process of Exploratory Data Analysis 37
2.10.3 Linear Regression Applications 39
2.10.3.1 Linear Regression 41
2.10.3.2 Random Forest 43
2.10.3.3 XGBoost 44
2.10.4 Phase 4: Model Walk-Forward Validation Framework 44
2.10.4.1 Iterative Validation Design 45
2.10.4.2 Conservative Model Configurations 45
2.10.4.3 Comprehensive Performance Assessment 45
2.10.4.4 Cross-Iteration Analysis 45
2.10.5 Phase 5: Production Forecasting 45
2.10.5.1 Unified Training Dataset 46
2.10.5.2 Production Model Training Implementation 46
2.10.5.3 FY2024 Forecast Generation Methodology 46
2.10.5.4 Comprehensive Comparison Framework 46
2.10.5.5 Validation Actual Data and Identify the Best Performing Model 46
2.10.6 ARIMA Time Series Models 47
2.10.6.1 Stationarity and Autocorrelation Testing 49
2.11 Machine Learning Applications in Warranty Forecasting 52
2.11.1 Random Forest Methods 52
2.11.2 Gradient Boosting Techniques 53
2.11.3 Ensemble Methods 54
2.12 Comparative Studies: AI vs Statistical Methods 56
2.12.1 Performance Comparison Literature 56
2.12.2 Performance Metrics in Comparative Studies 56
2.12.3 Expected Out-of-Sample Performance 58
2.12.4 Cross-Validation for Time Series 58
2.13 Big Data Analysis for Warranty and Quality Improvement 59

2.14 Product Design Specification 60
2.14.1 Data Exploration, Manipulation, And Cleaning Techniques 61
2.14.2 Advanced Analytical Techniques For Warranty Data Analysis (WDA) 62
2.14.3 Enhance Decision-Making with Business Intelligence (BI) 63
2.14.4 Data Preparation and Extraction, Transformation and Loading Process 63
2.14.5 Dashboard Development and Visualization with Power BI 64
2.14.6 Application of Plan Do Check Act Cycle in Data Analytics 64
2.14.7 Time-Based Analysis and Predictive Failure Modelling 65
2.14.8 Warranty Cost Analysis and Financial Trends Evaluation 65
2.14.9 Integration of Quality 4.0 and Business Intelligence 66
2.15 Literature Review Summary 66
2.16 Research Gap 68

CHAPTER 3 RESEARCH METHODOLOGY 70
3.1 Introduction 70
3.2 Product Design Specification 71
3.2.1 Phase 1: Information Collection 71
3.2.2 Phase 2: Concept Generation and Evaluation 75
3.2.3 Phase 3: Product Design 80
3.2.4 Phase 4: Parametric Analysis and Dashboard Design 84
3.2.4.1 Parametric Warranty Data Structuring 85
3.2.4.2 Dashboard Design and Analytical Implementation 86
3.2.4.3 Power BI Integration and Predictive Capability 88
3.2.4.4 Strategic Impact 89
3.3 Model Validation 89
3.3.1 Software and Tools Implementation 90
3.3.2 Methodological Framework Overview 90
3.3.2.1 Phase 1: Data Preparation and Structuring 91
3.3.2.2 Phase 2: Exploratory Data Analysis (EDA) 95
3.3.2.3 Phase 3: Model Development 106
3.4 Investigation and Determine Root Cause 110
3.4.1 Customer Feedback Activity 111

3.4.2 Conduct Component Repair 115
3.5 Warranty claim analysis 116
3.6 Develop a Predictive Warranty System Model 118
3.7 Validate The Model for Component Repair 119
3.7.1 Feasibility Study To Conduct Component Repair 128
3.7.1.1 SWOT analysis 128
3.7.1.2 Business Model Canvas (BMC) analysis 129
3.8 Process of Component Repair: Comparison of Repair Concepts 132

CHAPTER 4 RESULT AND DISCUSSION 134
4.1 Introduction 134
4.2 Development of Novel Big Data Analysis 134
4.2.1 Evolution of Warranty Management 135
4.2.2 Role of AI and Business Intelligence in Warranty Analysis 136
4.2.3 Overview of Dashboard Development 137
4.2.4 Data Selection and Preparation 138
4.2.4.1 Time-Based Warranty Trend Analysis 140
4.2.4.2 Model Distribution of Warranty Claims 141
4.2.4.3 Part Name Failure Analysis 142
4.2.4.4 Cause Description Categorization 143
4.2.4.5 Country-Based Claim Analysis 144
4.2.4.6 Warranty Claims by Failed Part 144
4.2.4.7 Data Preparation and Visualization for Warranty Optimization 145
4.3 AWAS Dashboard Design 149
4.4 Warranty Trends and KPI Insights 150
4.5 System Impact 150
4.6 Interface Dashboard 150
4.7 Model Development Results 153
4.7.1 Training Performance Analysis 153
4.7.2 Bias-Variance Trade-off in Warranty Datasets 154
4.7.3 Overfitting Detection and Practical Implication 154
4.7.4 Multi-Model Forecast Generation 154
4.8 Validation Against Actual Data 157

4.8.1 FY24 Actual Warranty Performance 157
4.8.2 Actual vs. Predicted Performance and Ranking 158
4.8.3 Model Selection Recommendations 161
4.9 Warranty Data Visualization 162
4.9.1 Warranty Claims Trend: Quantity Versus Paid Amount 162
4.9.2 Sales Growth Context 164

CHAPTER 5 CONCLUSION AND RECOMMENDATION 166
5.1 Conclusion 166
5.2 Recommendations for Future Work 167
REFERENCES 169
APPENDICES 189
AUTHOR'S PROFILE 199

LIST OF TABLES

Tables	Title	Page
Table 1.1	Market Estimation	6
Table 2.1	Benchmarking Based on Product Supply Chain [53]	29
Table 2.2	Domain Parameter	67
Table 2.3	Correspondence Index	68
Table 2.4	Measurement Model	68
Table 3.1	Decision Software	75
Table 3.2	Morphological Chart Comparison of Software	77
Table 3.3	Parameter of Warranty Data	86
Table 3.4	Vehicle Model Distribution and Claim Characteristics	100
Table 3.5	Seasonal Pattern Strength Analysis by Vehicle Model	101
Table 3.6	Part Category Analysis Summary	103
Table 3.7	Feature Engineering Priority Matrix	103
Table 3.8	Feature Engineering Summary by Model Type	109
Table 3.9	Feature Group Summary and Missing Value Treatment	109
Table 3.10	Rule of Thumb for Results [143]	112
Table 3.11	Cronbach's Alpha for 3 Years	112
Table 3.12	Analysis Links the Feedback Data by Customers	113
Table 3.13	Summary of joint investigation for Alternator abnormal noise cases	123
Table 3.14	Comparison of Belt Tension, Electrical Load, and Vibration Characteristics	125
Table 3.15	Belt Tension Confirmation from Good Vehicle and New Vehicles	126
Table 3.16	SWOT analysis	129
Table 3.17	Business Model Canvas Analysis	132
Table 4.1	Training Performance Comparison	153
Table 4.2	FY24 Annual Total Comparison by Model	156
Table 4.3	FY24 Forecasted Monthly Warranty Costs (Apr 2024 – Mar 2025)	156

Table 4.4	FY24 Actual Warranty Data Summary	157
Table 4.5	Model Ranking (by MAPE – Mean Absolute Percentage Error)	159
Table 4.6	Monthly Warranty Cost Actual vs Forecast Comparison by Model	161

LIST OF FIGURES

Figures	Title	Page
Figure 1.1	Company MNC Warranty Claim Trend [1]	2
Figure 1.2	Company MNC Domestic Sales by OE Customer	5
Figure 1.3	Malaysia Market Economic Outlook and Automotive Industry Trends	6
Figure 1.4	Breakdown Warranty Claim Trend by Customers	7
Figure 2.1	Company MNC Sales by Customer	15
Figure 2.2	Electronic Plant Assembly Products	18
Figure 2.3	Vehicle System Plant Products	18
Figure 2.4	Malaysia Total Industry Volume (TIV) January to December 2023 vs 2022; Trend by Month [21]	19
Figure 2.5	Company MNC Sales by Product Manufactures	22
Figure 2.6	Company MNC Top 2 Supplier Contribute Design Issue in 2020 [49]	26
Figure 2.7	Company A Promotes Warranty Repair Cost Reduction with Model Concepts [49]	27
Figure 2.8	Total Warranty Claim Alternator for Company E	30
Figure 2.9	Company MNC Warranty Claim vs Vehicle Sales by Company A	33
Figure 2.10	Framework Product Design Specification Approach	61
Figure 2.11	PDCA Cycle for Warranty Data Analysis	61
Figure 3.1	Flowchart of Warranty Management System Model System	70
Figure 3.2	Warranty Claim Raw Data.	72
Figure 3.3	Data Collection Process	74
Figure 3.4	Conceptual Interface Design	78
Figure 3.5	Power Query Editor	80
Figure 3.6	Visualizations in Microsoft Power BI to Improve Interactive of Dashboards	82
Figure 3.7	MVC Framework for System Architecture	83
Figure 3.8	Layout ff Automotive Warranty Analysis System Dashboard	87

Figure 3.9	Detailed GUI Design for Warranty Management System	88
Figure 3.10	Line Graphs Showing Monthly Warranty Claim Trends with COVID-19 Impact.	96
Figure 3.11	Line Graphs Showing Rolling Mean and Rolling Standard Deviation of Monthly Warranty Claim Amounts with COVID-19 Impact	98
Figure 3.12	Average Monthly Warranty Claim Amounts by Calendar Month	99
Figure 3.13	Customer Satisfaction Feedback for Market Claim	114
Figure 3.14	Warranty Claim Amount vs Quantity From FY18 ~ FY23	117
Figure 3.15	Identify Focus Product Based on Amount	117
Figure 3.16	Identify Focus Product Based on Quantity	118
Figure 3.17	Investigation Results Bearing Failure	119
Figure 3.18	Process Failure Mode and Effects Analysis (FMEA) - Bearing	120
Figure 3.19	Fault Tree Analysis [FTA] – Bearing Failure [75]	121
Figure 3.20	Estimate Failure Analysis by Fault Tree Analysis [75]	126
Figure 3.21	Complained & Good Vehicles (Used) [75]	127
Figure 3.22	New vehicles after Quality Control check by Company A Manufacturing [75]	127
Figure 4.1	Evolution of Warranty and Management	135
Figure 4.2	Warranty Data Selection For Dealer Code	139
Figure 4.3	Warranty Data Selection for Cause Description	139
Figure 4.4	Warranty Data Selection for Claimed (Year/Month)	140
Figure 4.5	Count of L/O Date and Repair Date	141
Figure 4.6	Count of Quantity by Model Name	142
Figure 4.7	Count of Quantity by Part Name	143
Figure 4.8	Count of Quantity by Cause Description	143
Figure 4.9	Sum of Quantity by Claim Country	144
Figure 4.10	Count of Quantity by Part Name	145
Figure 4.11	Raw Warranty Data in Excel	146
Figure 4.12	Five-Year Warranty Claims Trend	147
Figure 4.13	Warranty Cost Distribution by Year	147
Figure 4.14	Warranty Cost Distribution by Year	148
Figure 4.15	Automotive Warranty Analysis System (AWAS) dashboard.	149

Figure 4.16	Graphical User Interface	152
Figure 4.17	Historical Warranty Costs Versus FY2024 Forecasts (Multiple Models)	155
Figure 4.18	FY2024 Monthly Forecast Comparison by Model	155
Figure 4.19	Actual versus Predicted Warranty Claim Cost Performance for FY2024	159
Figure 4.20	Directional Accuracy of Forecasting Models in Predicting Warranty Cost Trends	160
Figure 4.21	Total Warranty Claim Based on Quantity vs Paid Amount	163
Figure 4.22	Total Company A Alternator Assy Warranty Claim Trend	163
Figure 4.23	Total Company MNC Total Sales Trend	165

LIST OF ABBREVIATIONS

Abbreviations	
AI	Artificial Intelligence
ARIMA	Autoregressive Integrated Moving Average
AWAS	Automotive Warranty Analysis System
EDA	Exploratory Data Analysis
EDER	Early Detection Early Resolution
FMEA	Failure Mode and Effects Analysis
FTA	Fault Tree Analysis
FY	Fiscal Year
GUI	Graphical User Interface
GWAMS	Global Warranty Analysis Management System
KPI	Key Performance Indicator
MAPE	Mean Absolute Percentage Error
MLRA	Multiple Linear Regression Analysis
OEM	Original Equipment Manufacturer
OFP	Original Failure Parts
PDCA	Plan–Do–Check–Act
PDS	Product Design Specification
QA	Quality Assurance
RMSE	Root Mean Square Error
TIV	Total Industry Volume
SARIMA	Seasonal Autoregressive Integrated Moving Average
WDA	Warranty Data Analysis

CHAPTER 1
INTRODUCTION

1.1 Research Background

The automotive industry, a hub of innovation and technological advancement, faces the escalating challenge of managing warranty claims. Warranty claims, which arise from product defects or component failures, impose significant financial burdens on manufacturers and can negatively affect customer satisfaction and brand reputation. As vehicles become more sophisticated, the complexity of managing warranty claims has increased, rendering traditional reactive approaches inadequate. The complexity of modern vehicles, with their advanced technologies, has intensified the challenge of identifying, resolving, and preventing warranty-related issues. The traditional reactive model, which addresses issues only after they arise, is no longer sufficient. Automotive manufacturers must now adopt proactive strategies, identifying potential issues before they lead to warranty claims, to ensure a smoother and more efficient operational cycle. In response to this shift, data-driven strategies have emerged as a solution for the automotive industry. By leveraging big data analytics, predictive maintenance models, and sustainable practices, manufacturers aim to mitigate the financial impact of warranty claims and enhance product quality. Data-driven decision-making, informed by insights from warranty data analytics, promises to transform quality control, production processes, and customer satisfaction. This research explores data-driven strategies for mitigating warranty claims in the automotive sector. Through comprehensive analysis of big data analytics, predictive maintenance, and sustainable practices, the study aims to provide actionable insights for automotive manufacturers. The objective is to equip manufacturers with the knowledge and tools needed to prevent warranty claims through strategic, data-informed interventions. The convergence of automotive engineering and data science offers unprecedented opportunities for redefining industry norms. This study contributes to the advancement of warranty claim management in the automotive sector by developing a framework to reduce warranty claims, streamline big data analysis, and secure sales profitability, ultimately fostering sustained customer satisfaction.

1.2 Problem Statement

The increasing complexity of modern automotive systems, combined with the rapid expansion of heterogeneous warranty data, has intensified the challenges associated with effective warranty management. Despite ongoing quality improvement initiatives within the industry, existing practices remain predominantly reactive, limiting the ability to perform early fault detection and proactive decision-making. This limitation is particularly critical in high-volume automotive environments, where even marginal inefficiencies can lead to substantial financial impact and reduced customer satisfaction. Empirical evidence from Company MNC, Figure 1.1 illustrates the annual warranty claim trend for Company MNC from FY2019 to FY2023. The warranty claim amount remained high in FY2019 and FY2020 at RM648K and RM643K, respectively, before showing a gradual decline to RM504K in FY2021, RM445K in FY2022, and RM388K in FY2023. Overall, the warranty claim amount reduced by RM260K, representing an approximate reduction of 40.1% over the five year period. Although the trend indicates improvement in warranty cost performance, the remaining claim amount in FY2023 is still significant, suggesting that recurring warranty issues and operational inefficiencies continue to exist. The results highlight the need for a more structured and data driven warranty management approach to improve failure identification, forecasting capability, and proactive decision making for long term warranty cost reduction and product reliability improvement. This highlights the need for more structured and meaningful visualisation techniques that reflect both performance trends and process dynamics.

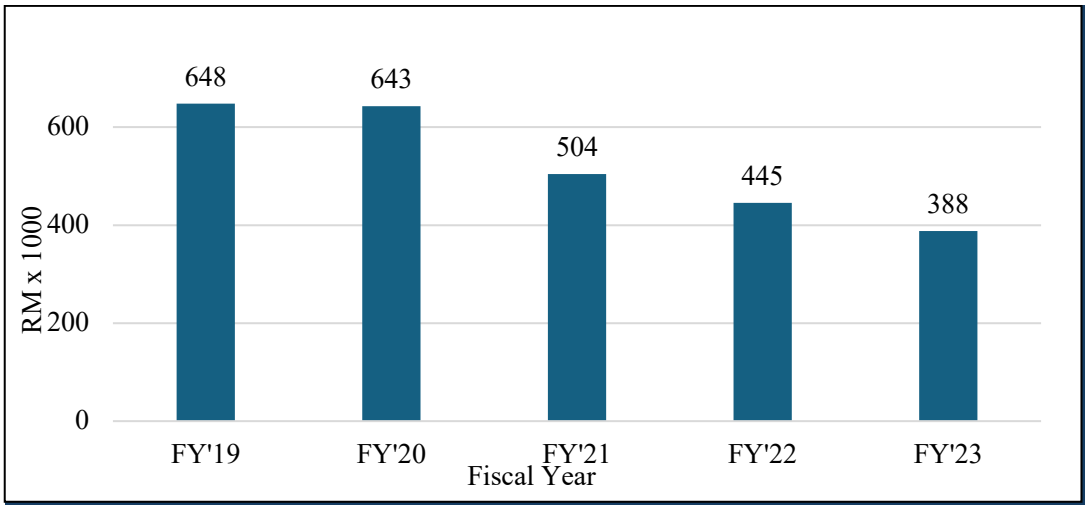


Figure 1.1 Company MNC Warranty Claim Trend [1]

Frequent warranty claims not only erode customer trust but also jeopardize brand loyalty and market share, particularly for companies like local car manufacturing company (Company A), which currently holds the largest market share in Malaysia. Company A dominates the Malaysian automotive market with a 36% market share, equivalent to approximately 204,000 units sold in FY2023. This is followed by Company D at 16% (92,000 units), Company C at 13% (72,000 units), and Company B at 11% (64,000 units) [2]. This distribution clearly indicates that Company A operates at a significantly larger scale compared to its other local car manufacturing company competitors. The high vehicle population directly increases warranty exposure, as the probability of failure occurrences is proportional to the number of units in operation. Consequently, even minor inefficiencies in warranty management may result in disproportionately large financial and operational impacts, further reinforcing the need for a more robust and predictive warranty management approach.

A fundamental limitation identified in the current warranty management approach is the absence of an integrated analytical system. At present, there are no prominent systems currently used in the company that systematically integrate predictive analytics, data visualisation, and engineering decision-making into a unified warranty management framework. Existing practices rely heavily on descriptive reporting and retrospective analysis, which restrict the organisation's ability to detect early failure patterns, predict future risks, and implement timely corrective actions. Furthermore, deficiencies in data representation reduce the effectiveness of current analytical practices.

Another significant gap lies in the lack of benchmarking and performance thresholds. The current system does not define benchmark values, such as acceptable warranty cost relative to sales volume or claim rate per vehicle, which are essential for evaluating organisational performance against industry standards. Additionally, the absence of clearly defined targets or thresholds prevents the identification of performance gaps, making it difficult to validate whether the observed improvements meet strategic or operational requirements. As a result, performance evaluation remains largely qualitative and lacks a robust basis for decision-making.

From a data integrity perspective, warranty datasets are often affected by tangible errors, including duplicate records, missing entries, inconsistent failure coding, and non-trouble-found (NTF) cases due to diagnostic inaccuracies. It is estimated that NTF cases may contribute approximately 10–15% of total warranty claims, leading to

inflated cost reporting and distorted failure patterns if not properly managed. These data quality issues significantly impact the reliability of statistical analysis and reduce confidence in the derived insights. Moreover, although statistical and analytical methods are applied in warranty analysis, their practical impact remains limited due to insufficient integration with real operational factors and engineering decision-making processes. The lack of translation from statistical outputs into actionable insights reduces the effectiveness of these methods in improving warranty performance and cost efficiency. This indicates the need for a more holistic approach that combines data analytics with engineering judgement and operational context.

In addition, the current system lacks consistent quantitative validation using percentage-based performance indicators. Key metrics, such as percentage reduction in warranty cost, improvement in claim rate, and cost per unit, are not systematically defined or monitored, weakening the justification for improvement initiatives and limiting the ability to demonstrate measurable impact. Without such quantification, it is difficult to establish a strong evidence-based argument for system enhancement. Given these limitations, there is a clear need for the development of an integrated, data-driven warranty management system that addresses data quality issues, incorporates predictive analytics, establishes benchmarking and performance thresholds, and enhances data visualisation. Such a system is essential to transition from reactive to proactive warranty management, enabling more accurate decision-making, improved cost control, and enhanced product reliability. Ultimately, this will support the organisation in sustaining its competitive advantage and achieving long-term operational excellence in the automotive industry.

The complexity of modern vehicles generates vast amounts of data, presenting a challenge for manufacturers who struggle to transition from reactive to proactive warranty management. The core issue involves not just reducing warranty claims but revolutionizing the approach through data-driven strategies. Traditional methods lack the precision necessary for contemporary automotive technologies, leaving manufacturers vulnerable to recurring issues, increased costs, and diminished customer satisfaction. To address these challenges, this research develops a framework that leverages data-driven strategies using suitable system to expedite analysis and improve accuracy. Additionally, it promotes the repair of components to the same manufacturing standards as new parts, thereby reducing warranty costs. This approach transforms warranty management in the automotive industry by enhancing product reliability,

customer satisfaction, and operational efficiency through proactive data-driven solutions.

Company MNC domestic sales performance increased steadily from RM594 million in FY2019 to RM1002 million in FY2023, representing an overall growth of RM408 million or approximately 68.7% shown in Figure 1.2. The strongest increase occurred between FY2021 and FY2022, reflecting recovery in manufacturing activities, stronger customer demand, and higher OE market contribution. The expanding sales volume implies broader product deployment in the field, which increases operational exposure to warranty risks and reinforces the importance of strengthening warranty monitoring, reliability assessment, and analytical decision support capability.

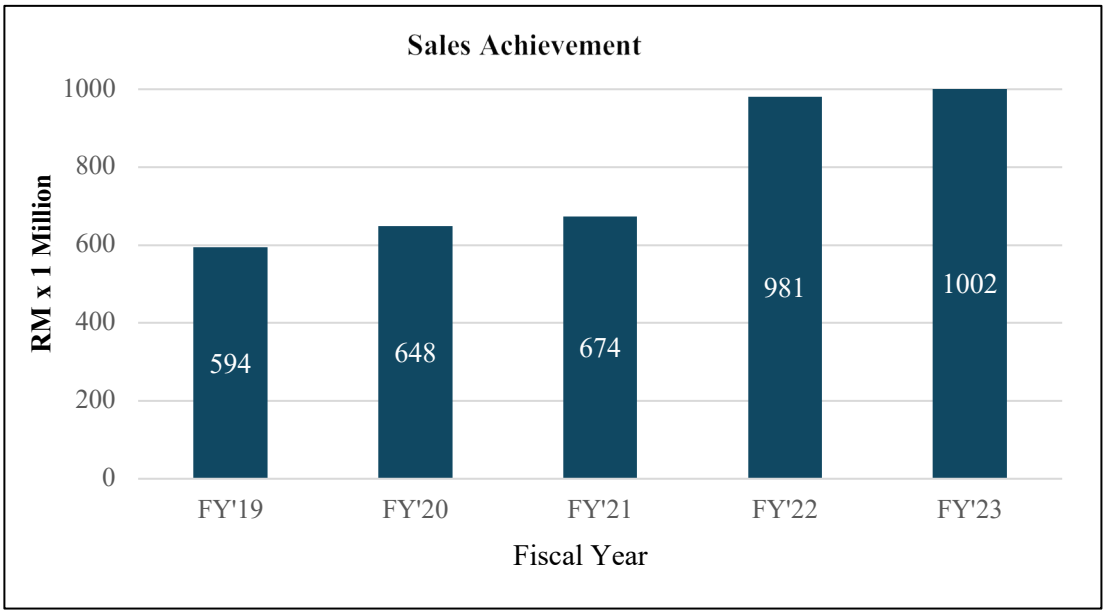


Figure 1.2 Company MNC Domestic Sales by OE Customer

Malaysia automotive market performance was closely influenced by national economic conditions, where Total Industry Volume (TIV) declined from 567 thousand units in FY2019 to 556 thousand units in FY2020 following negative GDP growth during the COVID 19 disruption period. Recovery accelerated from FY2021 onward, reaching 789 thousand units in FY2023, while future projections indicate approximately 2% annual growth supported by positive GDP outlook, new vehicle demand, and backlog recovery activities as shown in Figure 1.3. The Table 1.1 anticipated market expansion suggests increasing vehicle population and business activity, creating greater demand for predictive warranty analysis and stronger cost management capability.

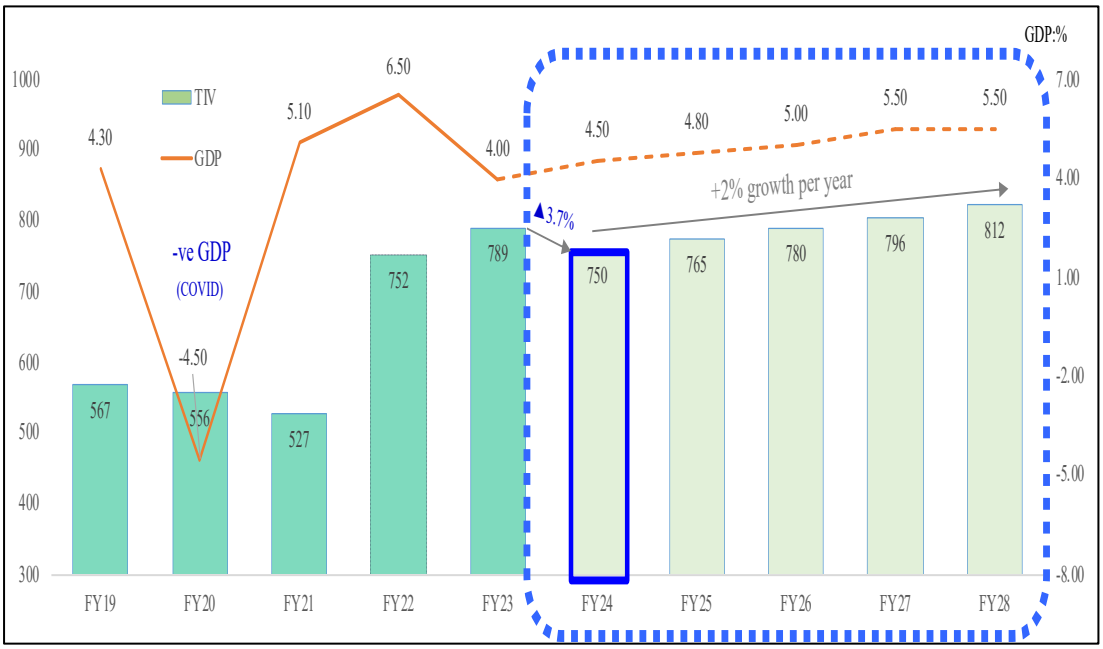


Figure 1.3 Malaysia Market Economic Outlook and Automotive Industry Trends

Table 1.1
Market Estimation

Impact	FY24 Estimation	FY25 ~ FY28
Positive	• Positive economic growth with GDP 4% ~ 5%. • New models (FY23 & FY24) continue to have good demand. • Company A continue recover FY23 back orders (>50K)	Expected growth +2% year on year in line with positive GDP growth.
Negative	Drop sales due to other makers (Company B, Company D, Company C) already cleared FY23 back orders.	

Figure 1.4 illustrates warranty claims between FY2018 and FY2023 were largely concentrated within Company A, contributing 72% of the total claim amount, followed by Company C at 14%, Company B at 13%, and other customers at 1%. The distribution indicates that customer application conditions and field operating environments had a significant influence on warranty behaviour. Although claim levels improved in the later fiscal years, the dominant contribution from Company A indicates the continued presence of recurring reliability concerns, justifying focused investigation, customer specific monitoring, and targeted corrective strategies to improve long term warranty performance.

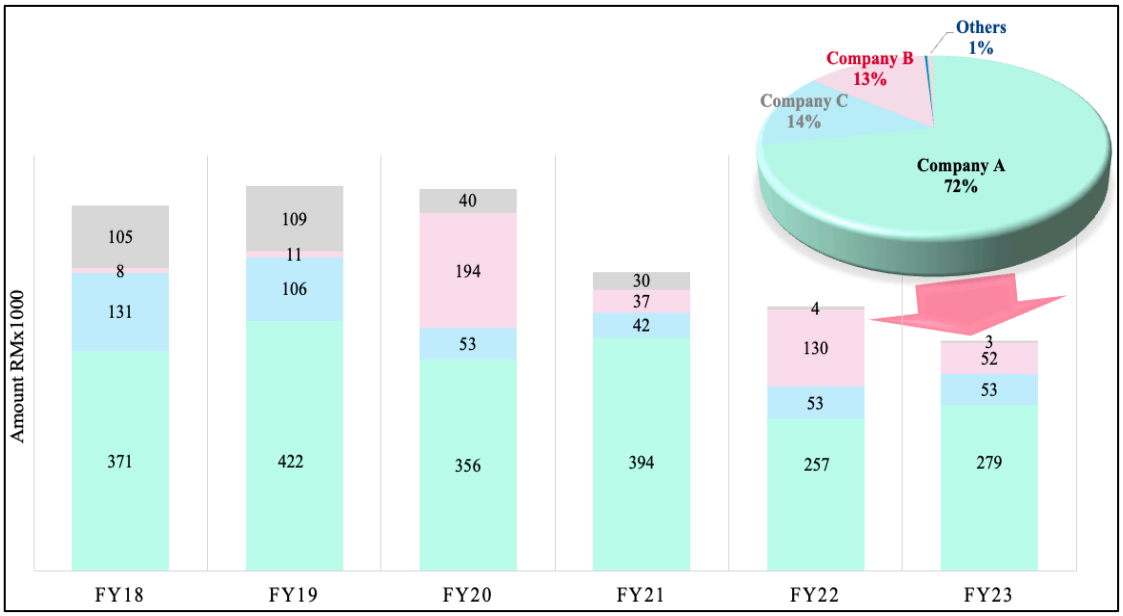


Figure 1.4 Breakdown Warranty Claim Trend by Customers

1.3 Objectives

To address the escalation of warranty claims, a transformation in management practices is essential. This transformation relies on data-driven strategies that revolutionize warranty claim management within the automotive sector. In this context, the research seeks to achieve the following objectives:

1. To analyse the Malaysian market overview and visualize warranty data over a eight-year period, identifying patterns, trends, and anomalies in the data.
2. To develop a predictive warranty system model for identifying factors influencing warranty claims and predicting future claim occurrences and costs.
3. To validate the model for component repair in mitigating warranty costs through the investigation of part recovery.

1.4 Research Scope and Limitation

This This research aims to comprehensively explore data-driven strategies for mitigating warranty claims in the automotive industry, with a primary focus on the application and impact of big data analytics and predictive maintenance models. The

scope is designed to ensure both academic rigour and industrial relevance, particularly within the Malaysian automotive context.

Firstly, the geographical scope is defined within the domestic automotive market in Malaysia, focusing on local automotive manufacturers, suppliers, and operational practices. This domestic focus ensures that the findings are directly applicable to the Malaysian automotive ecosystem, taking into account local operating conditions, service environments, customer usage patterns, and regulatory context. The primary focus is on the automotive manufacturing sector, including original equipment manufacturers (OEMs) and their supply chains. The study covers a range of vehicle categories, including passenger vehicles and light commercial vehicles, to ensure broader applicability of the findings. In terms of product scope, this study places particular emphasis on the Alternator Assembly component, as it represents one of the most data-intensive and warranty-critical components within the dataset. The alternator is selected due to its high frequency of warranty claims, availability of structured historical data, and significant contribution to warranty cost, making it a suitable case for detailed analysis and predictive modelling. While other components supplied by Company MNC are considered at a general level, the alternator serves as the primary case study to enable in-depth technical, statistical, and engineering evaluation. Key aspects of the research include the investigation of data-driven strategies through the application of advanced analytical techniques. This involves examining the use of big data analytics to identify patterns, anomalies, and trends related to warranty claims. The study further explores predictive maintenance models aimed at anticipating failures before occurrence, thereby supporting proactive decision-making. Detailed analyses are conducted on the implementation and effectiveness of these strategies, including their integration into warranty management processes.

Additionally, the research examines sustainable and operational practices within automotive manufacturing, including process improvement initiatives, quality control mechanisms, and supply chain coordination. These factors are analysed to understand their contribution to reducing warranty claims and enhancing overall product reliability. The study incorporates real-world case data and expert insights to strengthen its practical relevance. Warranty data is sourced from Company MNC, covering operations from its manufacturing plant in Bangi and sales office in Shah Alam. The dataset includes warranty claims collected from customers across Malaysia, including both Peninsular and East Malaysia, ensuring comprehensive domestic coverage. Functional

activities considered in this research include warranty claim investigation, Original Failure Part (OFP) analysis, quality data collection via service dealer networks, and reporting of quality issue resolution progress.

A financial perspective is also incorporated through cost-related evaluation, examining the economic impact of warranty claims and the potential benefits of implementing data-driven strategies. This includes assessing cost trends, identifying high-cost components, and evaluating opportunities for cost reduction through improved decision-making. Despite the comprehensive scope, several limitations are acknowledged. The study is constrained by the availability and accessibility of proprietary industrial data, which may limit the extent of generalisation across different organisations. Additionally, the dynamic nature of the automotive industry, including rapid technological advancements and changing customer usage behaviour, may influence warranty patterns over time. Furthermore, the focus on the alternator component, while enabling in-depth analysis, may limit the direct applicability of findings to other automotive components with different failure characteristics. These limitations are recognised and addressed to ensure the reliability, transparency, and practical relevance of the research outcomes.

By defining a focused yet comprehensive scope, this research aims to contribute significantly to the understanding and application of data-driven strategies for warranty claim mitigation. The findings are expected to provide actionable insights for industry practitioners, supporting the transition towards proactive, predictive warranty management within the Malaysian automotive industry.

1.5 Significant of Research

The automotive industry, at the forefront of technological innovation, faces the significant challenge of rising warranty claims. As vehicles evolve into sophisticated hubs of interconnected technologies, manufacturers contend with the financial strain, reputational risks, and operational inefficiencies these claims cause. This research advocates a shift from reactive to proactive warranty claim management through data-driven strategies. It holds profound significance for industry stakeholders and academia, offering transformative insights into proactive warranty management. By understanding financial implications, optimizing operational efficiency, and enhancing product quality, manufacturers can move towards a more resilient and efficient future. The

financial benefits of this research are substantial. In an industry where warranty claim costs are high, the cost-benefit analysis provided here acts as a strategic guide. It helps decision-makers understand how data-driven strategies can reduce warranty claims and yield significant returns on investment, empowering manufacturers to make informed financial decisions.

Operational efficiency is essential in automotive manufacturing. This research explores sustainable practices and resilient supply chains to optimize workflows, reduce defective components, and minimize warranty claims, positioning manufacturers at the forefront of technological innovation. Integrating advanced technologies like big data analytics and predictive maintenance models represents a significant technological leap. Quality enhancement, closely linked to customer satisfaction, is a core focus of this research. By examining sustainable practices and employee training, the study highlights the connection between environmentally conscious manufacturing and superior product quality. Proactive issue resolution becomes a reality, significantly enhancing customer satisfaction.

Academically, this research expands the boundaries of knowledge, serving as a comprehensive resource for scholars exploring data science and automotive manufacturing. As the industry evolves, this research will be foundational in understanding and driving innovation. Sustainable practices explored in this research align with global environmental responsibility. By reducing the ecological footprint through green manufacturing, manufacturers fulfil corporate social responsibilities and contribute to a more sustainable industry. The recommendations from this research are strategic guidelines rooted in industry best practices, helping decision-makers in the automotive industry steer towards a future-ready, resilient manufacturing ecosystem.

The significance of this research transcends theoretical analysis, offering tangible improvements across the automotive landscape. By integrating traditional practices with data-driven strategies, it guides the industry towards effectively addressing and mitigating warranty claims through proactive interventions. For Company MNC, this research aims to directly reduce customer claims and costs. By gaining customer trust, there is potential for future business with new model changes, sustaining profitability and growth. Good relationships and collaboration with customers can help mitigate unnecessary claims from service centers and prevent erroneous judgments. The concepts and activities explored here have potential for widespread application in the automotive industry, promoting effective component

repair. End-users will benefit from quick problem resolution and repairs that meet standard manufacturing specifications.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

The automotive industry in Malaysia is a vital sector, significantly contributing to the nation's economic development, industrial growth, and employment rates. It is essential not only for its economic metrics but also for fostering technological advancements and infrastructure development within the country. Customer satisfaction and loyalty are paramount in this sector, as they are crucial for maintaining a competitive edge and securing a robust market share. To achieve these objectives, the industry must excel in areas such as market uniqueness, relationship quality, and service excellence, ensuring that customer expectations are consistently met [3].

Over the past few years, there has been a notable increase in the adoption of quality improvement tools and lean management practices aimed at enhancing service delivery and customer satisfaction in the Malaysian automotive industry. Techniques like Design Failure Mode Effect Analysis (DFMEA) and the Customer Approach Matrix (CAM) have been widely implemented to identify and mitigate potential defects in automotive products, thereby reducing warranty claims and improving product reliability. Despite these efforts, local automotive companies continue to face significant challenges, particularly when competing with Japanese firms that are renowned for achieving high levels of customer satisfaction [4].

The continuous pursuit of customer satisfaction has prompted Malaysian automotive companies to refine their strategies and adopt innovative practices. This includes a deeper focus on understanding and addressing the root causes of warranty claims, improving product design, and enhancing overall service quality. These efforts are crucial for maintaining customer loyalty and ensuring long-term success in a highly competitive market. The industry's ability to meet these challenges head-on will determine its future trajectory and its capacity to contribute sustainably to Malaysia's economic growth [5].

Warranty-claim cost reduction is a critical activity for local company (Company MNC) to protect and sustain the company's profits. Company MNC focuses on

advanced mobility, aiming to positively impact global transportation and contribute to greater prosperity. Established in 1949, Company MNC Japan is a Tier 1 automotive supplier providing thermal systems, powertrain, mobility, electrical, and electronics to automakers and other manufacturers worldwide. As a Fortune 500 company, Company MNC has an extensive product portfolio and a significant global impact. The company's commitment to innovation and quality has positioned it as a leader in the automotive industry. Company MNC advanced technologies are designed to enhance vehicle performance, safety, and environmental sustainability [6].

Company MNC plays a vital role in this global network, focusing on reducing warranty claims to maintain high customer satisfaction. The company's efforts in warranty claim cost reduction involve rigorous quality control measures and continuous process improvements. These initiatives are crucial for sustaining profitability and ensuring the delivery of reliable products. The reduction of warranty claims not only protects Company MNC financial interests but also enhances its reputation as a reliable supplier. By minimizing defects and improving product durability, Company MNC ensures that customers receive high-quality automotive components. This strategy helps build long-term relationships with clients and fosters customer loyalty [7].

Moreover, Company MNC commitment to quality and innovation drives its efforts to meet the evolving needs of the automotive industry. The company invests in research and development to create advanced technologies that improve vehicle efficiency and safety. These innovations are critical for addressing the challenges of modern transportation and supporting sustainable mobility. In summary, Company MNC focus on warranty-claim cost reduction is a key element of its strategy to maintain profitability and enhance customer satisfaction. The company's global presence, extensive product portfolio, and commitment to quality make it a leader in the automotive industry [8].

2.2 Company Overview

Company MNC Corporation operates in 35 countries and regions, enabling it to deliver global and regional solutions that meet societal needs. This extensive network allows Company MNC to share different perspectives and leverage diverse expertise to develop innovative solutions. Company MNC benefits from this global reach, enhancing its capabilities and market presence. The company's mission is to explore the

deeper meaning of happiness and build a future with greater prosperity for all. Company MNC achieves this through world-class problem-solving capabilities and by providing technology that supports daily well-being. This holistic approach ensures that Company MNC innovations positively impact society and the environment [9].

Company MNC strategic focus on environmental sustainability and the advancement of a resilient, self-reliant society is evidenced through its continuous technological innovation and engineering excellence. The company leverages its advanced technology and productivity to extend a variety of automotive systems, providing growth to vehicle manufacturers worldwide. This focus on sustainability and innovation positions Company MNC as a forward-thinking industry leader. Company MNC product offerings include air conditioning systems, powertrain systems, safety and cockpit systems, automotive service parts and accessories, and service stations. These products cater to the diverse needs of automotive customers worldwide, ensuring that Company MNC remains a trusted partner for vehicle manufacturers. The company's comprehensive product portfolio supports its goal of delivering high-quality, reliable solutions [10].

The powertrain systems offered by Company MNC are particularly noteworthy, encompassing gasoline, diesel, hybrid, electric vehicles, and fuel cell vehicles. These systems are essential for enhancing vehicle performance and reducing environmental impact. Company MNC expertise in powertrain technology ensures that its products meet the highest standards of quality and efficiency. Overall, Company MNC exemplifies the company's commitment to excellence and innovation. By leveraging its global network and advanced technologies, Company MNC continues to drive progress in the automotive industry and contribute to a sustainable future [11].

Company MNC advanced technologies are designed to enhance vehicle performance, safety, and environmental sustainability. The company invests in research and development to create innovative solutions that address the challenges of modern transportation. These innovations are critical for improving vehicle efficiency and safety, supporting sustainable mobility, and reducing the environmental impact of transportation. Company MNC commitment to quality and innovation drives its efforts to meet the evolving needs of the automotive industry and contribute to a more sustainable future [12].

The company's global network enables it to leverage diverse expertise and perspectives to develop innovative solutions that meet the needs of its customers.

Company MNC extensive product portfolio includes air conditioning systems, powertrain systems, safety and cockpit systems, automotive service parts and accessories, and service stations. These products cater to the diverse needs of automotive customers worldwide, ensuring that Company MNC remains a trusted partner for vehicle manufacturers. The company's comprehensive product portfolio supports its goal of delivering high-quality, reliable solutions [13].

Company MNC was established in Malaysia in 1980 as a joint venture with its parent company and has grown into a major automotive component manufacturer. The company employs approximately 1,700 associates and supplies automotive components to several leading vehicle assemblers in Malaysia. Its customer base expanded progressively, beginning with Company B in 1983, followed by Company C in 1987, Company A in 1994, and Company D in 2000.

The customer sales profile reflects a diversified business presence across both domestic and export markets. Export activities contribute the largest share at 32%, followed by Company B at 30% and Company A at 28%, while the aftermarket segment accounts for 6%. Smaller contributions are recorded from Company D (2%), Company C (1%), and other customers (1%). This distribution indicates that Company MNC operates within a broad market environment influenced by multiple customer segments, production demands, and application requirements. Such market dependency may affect operational performance, warranty exposure, and long term business sustainability. The sales distribution by customer is presented in Figure 2.1.

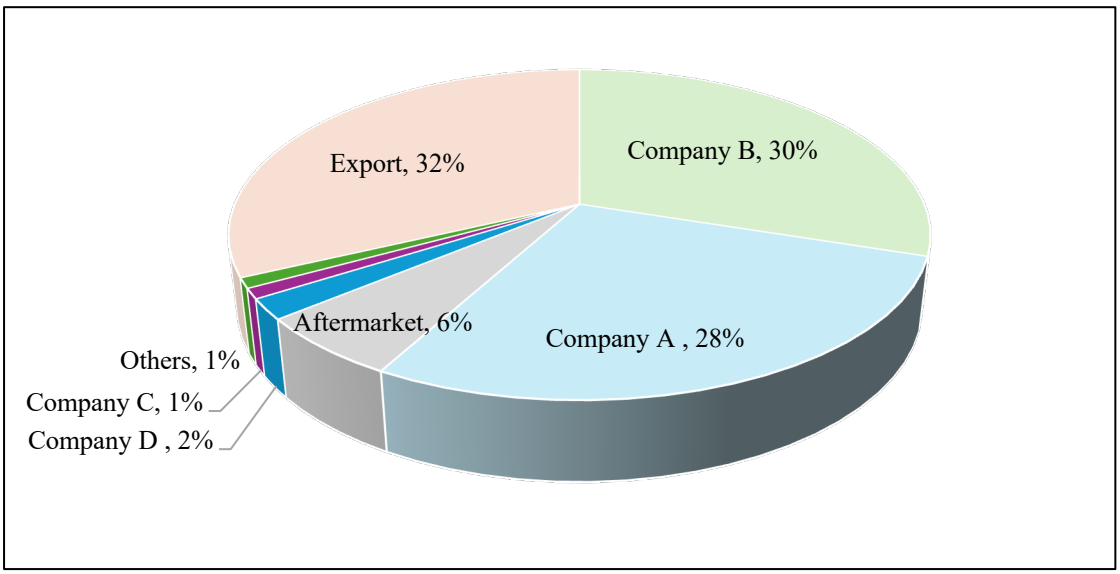


Figure 2.1 Company MNC Sales by Customer

Company MNC plays a vital role in the company's global network, focusing on reducing warranty claims to maintain high customer satisfaction. The company's efforts in warranty claim cost reduction involve rigorous quality control measures and continuous process improvements. These initiatives are crucial for sustaining profitability and ensuring the delivery of reliable products. By minimizing defects and improving product durability, Company MNC ensures that customers receive high-quality automotive components. This strategy helps build long-term relationships with clients and fosters customer loyalty [14].

2.3 Plant Locations and Products

Company MNC operates multiple facilities to support automotive component manufacturing and supply chain operations. The company's operational structure is divided into Thermal and Electronics manufacturing plants, supported by logistics and customer interface units, which collectively contribute to warranty performance and product quality outcomes.

The Vehicle System Plant, located in Bangi, manufactures key automotive components including heating, ventilation, and air conditioning (HVAC) systems, condensers, radiators, and alternator assemblies. These components are directly associated with vehicle reliability and warranty performance, as failures in thermal and electrical subsystems often contribute significantly to warranty claims [15]. Among these, the alternator assembly is particularly critical due to its functional role in electrical power generation and its higher exposure to failure under varying operational conditions, making it a key focus component in this study.

The Electronics Plant, also located in Bangi, produces electronic control units (ECUs), application-specific integrated circuits (ASICs), magnetic resistance elements (MREs), and programmable logic controllers (PLCs). These components support advanced vehicle functionalities, including control systems and safety features. While electronic components are essential for modern vehicle performance, this study places greater analytical emphasis on mechanical-electrical components such as the alternator, where warranty claim data is more extensive and suitable for predictive modelling [16]. From a market perspective, the Vehicle System Plant primarily serves the domestic automotive market, supplying components to local car manufacturing, whereas the Electronics Plant is largely oriented toward export markets. This domestic operational

focus is particularly relevant to the present study, as warranty claim data is collected from Malaysian customers across Peninsular and East Malaysia, ensuring alignment between production, field performance, and warranty analysis within a consistent operating environment [17].

Supporting these manufacturing operations, the Selaman Warehouse in Bangi facilitates inventory management and distribution of components from both Thermal and Electronics plants, as well as imported parts from overseas group companies. Efficient logistics operations are essential for maintaining product availability and supporting timely warranty replacement processes. Additionally, the Sales Office in Glenmarie serves as a key interface between Company MNC and automotive assemblers, enabling the collection of field quality data and customer feedback, which are critical inputs for warranty analysis and continuous improvement [18]. The integration of manufacturing, logistics, and customer feedback systems provides a structured data environment for warranty analysis; however, the absence of an integrated analytical platform limits the effective utilisation of this data for predictive decision-making. This limitation has been recognised in prior studies, where fragmented data systems reduce the effectiveness of warranty analytics and delay corrective actions [19]. This reinforces the need for a data-driven warranty management system capable of transforming raw operational data into actionable insights.

Figures 2.2 and 2.3 illustrate the product categories and plant distribution within Company MNC, highlighting the distinction between electronics and thermal product lines. The Vehicle System Plant's focus on domestic supply and high-impact components such as alternators further justifies its selection as the primary case study environment for this research.

2.3.1 Electronics Plant

Type of Product: Electronics & Semiconductor Product
Customer: Mainly for Export market

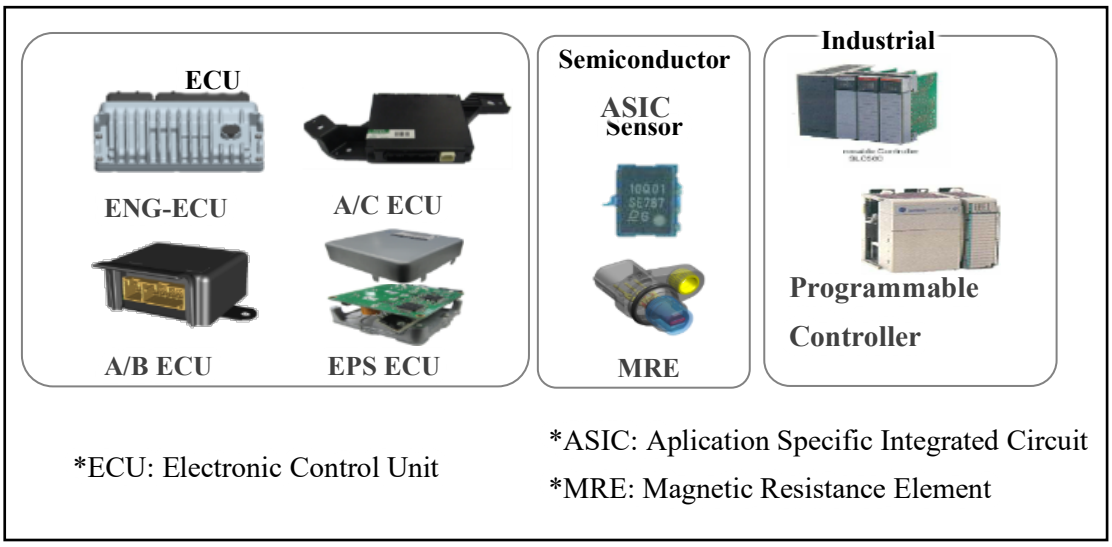


Figure 2.2 Electronic Plant Assembly Products

2.3.2 Vehicle System Plant

Type of Product: Air-Conditioning, Powertrain and Body Electrical components

Customer : Mainly for the Domestic market

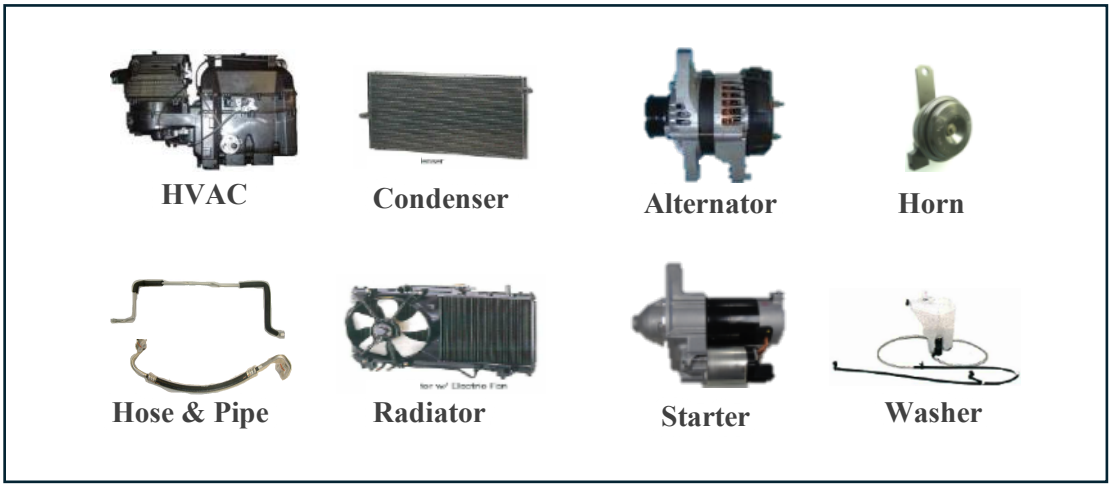
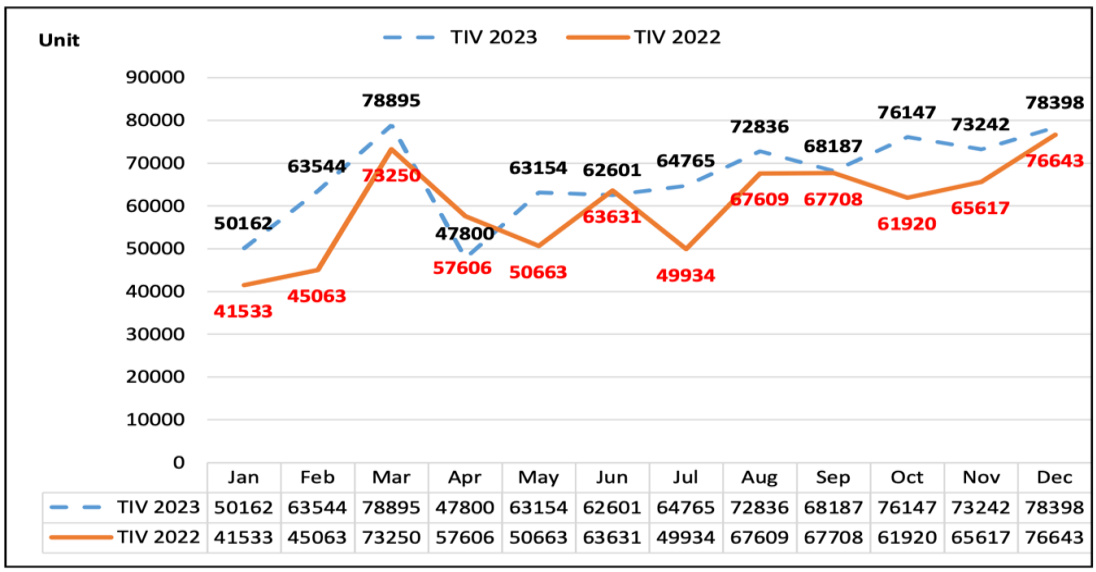


Figure 2.3 Vehicle System Plant Products

2.4 Malaysian Market Overview and Warranty Data Visualization

In 2023, the Malaysian automotive market experienced significant growth, with new motor vehicle sales increasing by 11%, primarily driven by the passenger car segment. The Malaysia Total Industry Volume (TIV) reached 78,398 units, reflecting strong post-pandemic recovery supported by favourable economic conditions and

government initiatives [20]. The monthly trend of vehicle sales is illustrated in Figure 2.4, showing consistent demand throughout the year [21].



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
TIV 2023	50162	63544	78895	47800	63154	62601	64765	72836	68187	76147	73242	78398
TIV 2022	41533	45063	73250	57606	50663	63631	49934	67609	67708	61920	65617	76643

Figure 2.4 Malaysia Total Industry Volume (TIV) January to December 2023 vs 2022; Trend by Month [21]

While this growth indicates a healthy market environment, it also implies a proportional increase in vehicle population and, consequently, higher exposure to warranty-related risks. As the number of vehicles in operation increases, the likelihood of component failures and warranty claims correspondingly rises, placing greater pressure on manufacturers to ensure product reliability and effective warranty management. Despite strong market performance, the Malaysian automotive industry remains highly competitive, particularly with the presence of established Japanese manufacturers known for superior reliability and customer satisfaction [22]. To remain competitive, domestic companies have adopted structured quality approaches such as Design Failure Mode and Effects Analysis (DFMEA) and Corrective Action Management (CAM). However, warranty claims continue to persist due to unresolved product reliability issues and limitations in current analytical practices.

From a customer perspective, common concerns associated with failed components are primarily related to vehicle performance degradation, unexpected breakdowns, warning indicators (such as battery or charging system alerts), and inconvenience due to repair downtime. In particular, failures in critical components such as the alternator directly affect vehicle operability, as they disrupt the electrical charging system, leading to battery discharge and potential engine stalling. These issues significantly impact customer satisfaction and trust in product reliability. Analysis of

warranty data further indicates that failed components are often associated with recurring technical issues. For alternator assemblies, typical failure modes include bearing wear, voltage regulator malfunction, diode failure, and thermal degradation due to prolonged exposure to high operating temperatures. These failures are commonly attributed to root causes such as material fatigue, insufficient cooling performance, electrical overloading, and variation in manufacturing quality. Additionally, customer usage patterns, environmental conditions (e.g., high humidity and temperature in the Malaysian climate), and inconsistent maintenance practices contribute to accelerated component degradation. These external factors introduce variability in failure behaviour, making it challenging to predict warranty trends using conventional methods.

Visualisation of warranty data has been identified as a key enabler for understanding such complex relationships between customer concerns, failure modes, and root causes. By analysing patterns in warranty claims, organisations can identify high-risk components and recurring failure trends. However, conventional visualisation approaches are often limited to descriptive reporting, lacking integration with predictive analytics and time-based analysis. Current visualisation practices also do not adequately capture time-related parameters, such as failure occurrence intervals, claim processing duration, and repair lead time, which are critical for evaluating the efficiency of warranty management processes. The absence of these time-consuming factors limits the ability to assess operational performance and identify bottlenecks in the warranty lifecycle.

Another challenge lies in customer behaviour related to after-sales service. A significant number of customers shift to unauthorised service centres after the free service period, resulting in inconsistent maintenance practices and reduced traceability of failure data. This behaviour contributes to increased warranty claims and complicates root cause analysis due to incomplete or inaccurate service records [23]. To address these challenges, customer education initiatives have been introduced to encourage proper vehicle usage and adherence to authorised service maintenance. Effective education programs can reduce misuse-related failures and improve long-term product reliability. In addition, targeted communication strategies based on warranty data insights can help address specific customer concerns and improve service retention rates [24]. The integration of big data analytics with traditional quality tools offers significant potential to enhance warranty management effectiveness. Combining DFMEA with

machine learning techniques enables more accurate prediction of potential failures and identification of critical risk factors. Furthermore, real-time data analytics platforms allow continuous monitoring of product performance, enabling early detection of anomalies and proactive intervention before failures escalate [25].

However, despite the availability of large volumes of warranty and operational data, the lack of an integrated analytical system limits the ability of organisations to transform this data into actionable insights. This gap highlights the need for a comprehensive data-driven warranty management framework, such as the proposed Automotive Warranty Analysis System (AWAS), which integrates predictive analytics, data visualisation, and engineering decision-making [26]. Company MNC's sales distribution by product category is presented in Figure 2.5, which further supports the identification of high-impact components for analysis. The dominance of thermal and electrical components, particularly alternators, reinforces their significance in warranty cost contribution and justifies their selection as the primary focus in this study.

A warranty represents a contractual agreement specifying the conditions under which a supplier will repair, replace, or compensate for a defective item for the buyer or end-user. Warranty expenses are those associated with the repair, replacement, or compensation provided to the consumer for any product damage. Essentially, the supplier or manufacturer commits to repairing or replacing a sold product within a stipulated period if it becomes damaged or fails to function correctly as per the warranty terms. Excessive warranty expenditure can adversely impact the annual budget allocation, potentially resulting in profit loss and hindering future business growth.

Company MNC warranty cost encompasses the total claims received from local car manufacturing, including part costs, labor, and taxes. Monthly, Company MNC receives and disseminates warranty claims data accurately and promptly through the Global Warranty Analysis Management System (GWAMS). This system facilitates the Early Detection Early Resolution (EDER) initiative, enabling Company MNC to collaborate effectively with parent company, thereby reducing warranty costs. Warranty data is an essential component of quality data, crucial for confirming product quality across various Company MNC plants.

Analysing five-year claims data through GWAMS allows for identifying trends and patterns, with Pareto charts used to highlight high-frequency or abnormal claims, thus facilitating rapid problem analysis and resolution. Company MNC identified an

abnormal number of claims concerning the Alternator Assembly, with many claims related to abnormal noise during idling conditions, as shown in Figures 2.5.

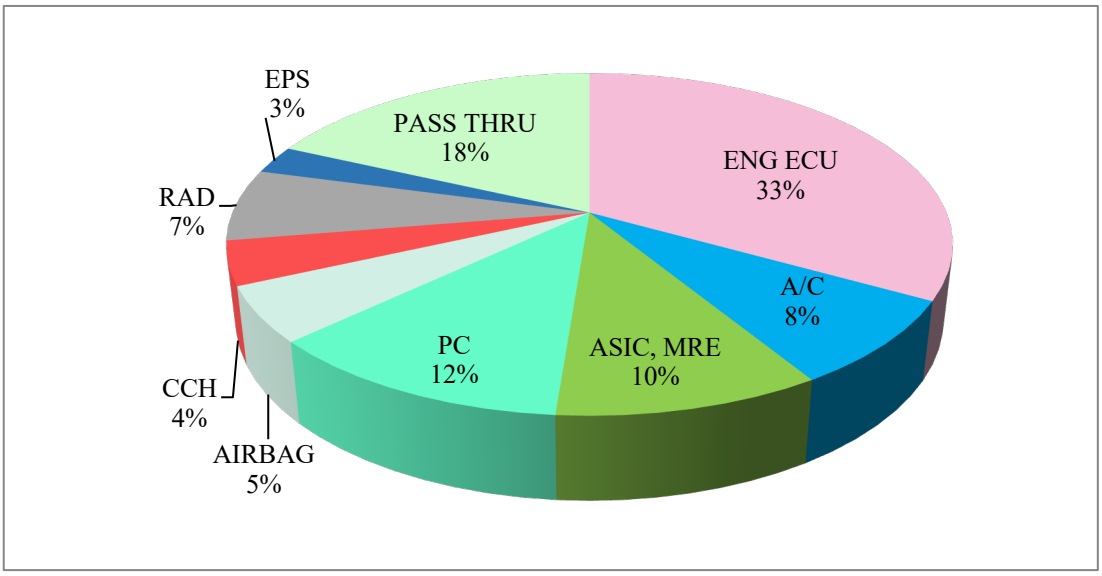


Figure 2.5 Company MNC Sales by Product Manufactures

2.5 Predictive Warranty System Model

The integration of big data analytics into the automotive industry represents a significant advancement in the quest to improve customer satisfaction and reduce warranty claims. Developing robust statistical models, machine learning algorithms, and data mining techniques are essential components of this process. These tools enable the analysis of vast amounts of data, providing valuable insights into customer behaviour, product performance, and service quality. By identifying patterns and predicting future failures, big data analytics can help automotive companies formulate effective strategies to address warranty issues proactively [27].

One of the key analytical methods employed in big data analytics is the Analytical Hierarchy Process (AHP), which helps prioritize factors influencing warranty claims. Another method, the Decision-Making Trial and Evaluation Laboratory (DEMATEL), examines the inter-relationships among these factors, allowing companies to make data-driven decisions that enhance product reliability and customer satisfaction. By leveraging these advanced analytical techniques, automotive firms can gain a competitive edge in the market and improve their overall performance [28] [29].

Furthermore, the integration of big data analytics with traditional quality management tools can significantly enhance the effectiveness of warranty management practices. For example, combining DFMEA with machine learning algorithms allows companies to predict potential failures more accurately and develop targeted strategies to mitigate these risks. This approach not only reduces costs associated with warranty claims but also improves customer satisfaction by ensuring higher product reliability. In addition to predictive analytics, big data can be utilized to enhance real-time monitoring and decision-making processes within the automotive industry. By implementing advanced data analytics platforms, companies can continuously monitor product performance and detect anomalies before they lead to failures. This proactive approach helps in reducing downtime, improving product reliability, and enhancing customer satisfaction [30].

Moreover, big data analytics facilitates the development of personalized customer experiences by analyzing customer preferences and behaviour patterns. Understanding customer needs and expectations enables automotive companies to tailor their services and products to meet individual requirements, thereby increasing customer satisfaction and loyalty. This level of personalization can significantly enhance the customer experience and foster long-term relationships [31] [32] [33]. Overall, the adoption of big data analytics in the automotive industry offers numerous benefits, including improved product reliability, reduced warranty claims, and enhanced customer satisfaction. By leveraging advanced analytics, companies can transform their warranty management practices, achieve cost savings, and gain a competitive advantage in the market. The strategic use of big data analytics is thus essential for the long-term success and sustainability of the automotive industry in Malaysia [34].

Big data analytics also supports root cause analysis, enabling manufacturers to identify and address the underlying causes of recurring warranty claims. Techniques such as failure mode and effects analysis (FMEA) and root cause analysis (RCA) can be enhanced with big data analytics, providing more accurate and comprehensive insights. These insights help manufacturers develop effective corrective actions, reducing the incidence of warranty claims and enhancing overall product quality. Additionally, the integration of big data analytics with quality management tools can enhance the predictive capabilities of these tools, allowing companies to identify potential failures before they occur and implement preventive measures [35] [36].

Despite its many benefits, the implementation of big data analytics in warranty management poses challenges. Ensuring data privacy and security is a significant concern, as handling large volumes of sensitive information requires robust safeguards. Additionally, the complexity of big data analytics demands substantial investment in technology and skilled personnel. Manufacturers need to develop comprehensive strategies to address these challenges, ensuring that they can fully leverage the potential of big data analytics to improve warranty management. Overcoming these challenges requires a commitment to continuous improvement and innovation, as well as a willingness to invest in the necessary resources and capabilities [36].

Real-time monitoring and decision-making processes within the automotive industry can be significantly enhanced by big data analytics. By implementing advanced data analytics platforms, companies can continuously monitor product performance and detect anomalies before they lead to failures. This proactive approach helps in reducing downtime, improving product reliability, and enhancing customer satisfaction. Real-time monitoring also allows companies to respond swiftly to emerging issues, preventing them from escalating and ensuring that quality issues are addressed promptly. This capability is critical for maintaining high levels of customer satisfaction and loyalty [37] [38]. By leveraging advanced analytics, companies can transform their warranty management practices, achieve cost savings, and gain a competitive advantage in the market. The strategic use of big data analytics is thus essential for the long-term success and sustainability of the automotive industry in Malaysia. By investing in the necessary resources and capabilities, companies can fully leverage the potential of big data analytics to improve their operations and enhance customer satisfaction [39].

2.6 Regression Model

Regression-based approaches, particularly Multiple Linear Regression (MLR), are widely adopted in warranty cost prediction due to their interpretability and effectiveness in modelling relationships between warranty outcomes and explanatory variables such as vehicle mileage and component failure rates. The performance of the MLR model is commonly assessed using the coefficient of determination (R^2), which represents the proportion of variance in the dependent variable that can be explained by the selected independent variables. In the context of automotive warranty prediction, the predictive model achieved $R^2 \geq 0.60$, reflecting a reliable level of predictive

accuracy given the inherent variability, data noise, and external influencing factors present in warranty datasets [31] [40] [41]. To address limitations associated with linear assumptions, more advanced regression techniques have also been explored. Support Vector Regression (SVR) is frequently applied to capture nonlinear relationships between operational factors and warranty costs, while Weibull analysis is particularly effective for modelling time-dependent failure behaviour and reliability characteristics of automotive components. The incorporation of these advanced methods has been shown to enhance prediction accuracy and provide deeper insight into failure trends, thereby strengthening data-driven warranty management strategies within the automotive industry [41], [42].

2.7 Validate The Model for Component Repair

Improving warranty management practices and operational efficiency requires a thorough investigation of warranty part recovery processes. Empirical research and root cause analysis can identify specific issues within the recovery process, enabling companies to develop targeted improvement strategies. For instance, understanding the root causes of component failures can help design more reliable products and reduce the incidence of warranty claims [43].

Preventive maintenance strategies during extended warranty contracts can significantly reduce warranty costs and enhance customer satisfaction. By adopting a proactive approach to maintenance, companies can prevent failures before they occur, thus improving product reliability and extending the lifespan of automotive components. This strategy not only reduces warranty costs but also boosts customer confidence in the brand. Efficiency in warranty part recovery processes is also critical. Streamlining these processes can lead to faster resolution of warranty claims and improved customer satisfaction. Implementing automated systems for tracking and managing warranty parts can reduce processing times and enhance the overall efficiency of the warranty management system. Such systems ensure that warranty parts are managed effectively, reducing delays and improving service quality [44].

Advanced technologies such as the Internet of Things (IoT) and blockchain can further enhance warranty part recovery processes. IoT devices can provide real-time data on product performance and condition, allowing companies to monitor and manage warranty parts more effectively. Blockchain technology ensures transparency and

traceability in the warranty part recovery process, reducing fraud and improving trust between customers and manufacturers [45] [46]. Collaboration with suppliers and other stakeholders is crucial for optimizing warranty part recovery processes. Working closely with suppliers ensures the timely availability of quality replacement parts and improves the overall efficiency of the recovery process. Effective communication and coordination among all stakeholders lead to better outcomes and higher customer satisfaction [47] [48].

Company A, a major automotive manufacturer in Malaysia, conducts annual supplier briefings to analyze and address design issues among its 124 suppliers. These briefings are critical for identifying key contributors to design-related problems that significantly impact warranty claims. The most recent briefing highlighted Company MNC and DATT as major contributors, with Company MNC responsible for 165 cases of design issues and DATT for 308 cases. This identification process is crucial for Company A strategy to improve product quality and enhance customer satisfaction by mitigating recurring design issues, as shown in Figure 2.6 [49].

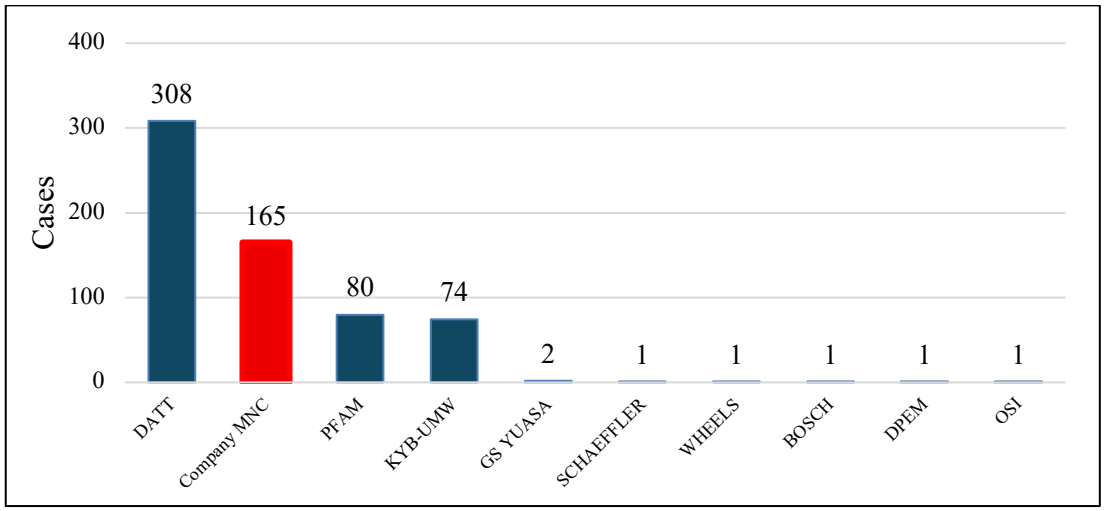


Figure 2.6 Company MNC Top 2 Supplier Contribute Design Issue in 2020 [49]

To address financial burdens and inefficiencies in its warranty claim process, Company A shifts from replacing entire assemblies to focusing on component parts replacements, reducing costs and resource wastage by substituting only defective components. High-quality component parts supplied by local supplier like Company MNC ensure reliable repairs, enhancing operational efficiency and sustainability, as shown in Figure 2.7 [41]. Analyzing five years of warranty claim data through the Global Warranty Management System (GWAMS), Company MNC identified abnormal

warranty claims associated with the Alternator Assembly, particularly noise complaints occurring during engine idling, as similarly issues illustrated in Figure 2.7. The recurring issue contributed to increased warranty expenditure due to the conventional practice of replacing the complete Original Failure Part (OFP) assembly during repair activities.

In response to expectations communicated during Company A 2023 supplier briefing, Company MNC proposed a component repair approach for the Alternator Assembly as an alternative to full assembly replacement. The proposed strategy focused on replacing only the affected child components responsible for the noise phenomenon, supporting cost reduction, improved repair efficiency, and better resource utilisation. This initiative aligns with the industry shift toward predictive maintenance and repair optimisation practices, where data driven failure analysis is used to improve warranty management effectiveness and minimise long term warranty cost impact.

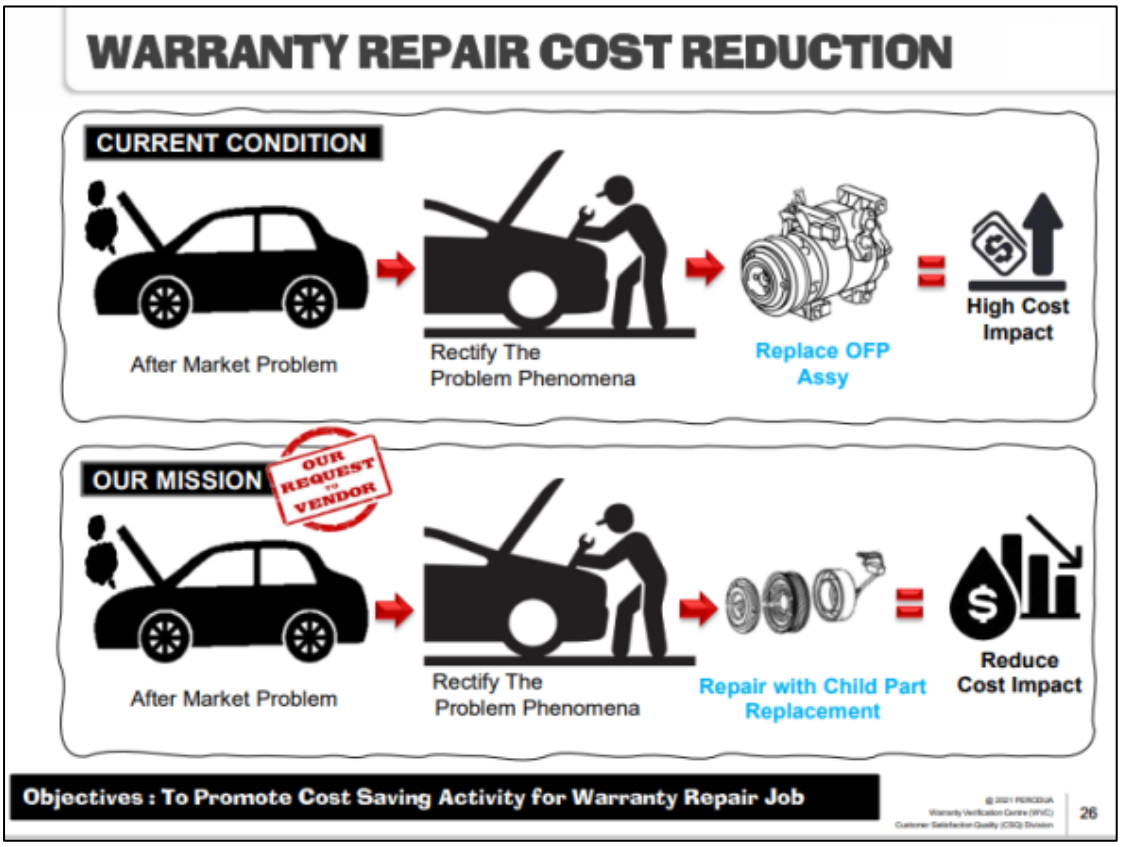


Figure 2.7 Company A Promotes Warranty Repair Cost Reduction with Model Concepts [49]

From a customer satisfaction perspective, transitioning to component parts replacements can have a profound impact. This approach minimizes repair times and

vehicle downtimes, which can significantly enhance customer satisfaction. The cost savings from this process can be reallocated towards customer-centric initiatives such as extended warranties or enhanced service packages, further bolstering customer loyalty. By improving the efficiency and quality of repairs, Company A not only reduces warranty claims but also reinforces its commitment to customer satisfaction [43].

The adoption of high-quality component parts supplied by partners like Company MNC ensures the reliability and performance of repairs, enhancing operational efficiency and sustainability. This approach minimizes repair times and vehicle downtimes, which can significantly enhance customer satisfaction. The cost savings from this process can be reallocated towards customer-centric initiatives such as extended warranties or enhanced service packages, further bolstering customer loyalty [50]. By improving the efficiency and quality of repairs, Company A not only reduces warranty claims but also reinforces its commitment to customer satisfaction.

Despite its importance, the warranty part recovery process faces several challenges. Ensuring timely and accurate data collection is essential but can be difficult to achieve consistently. Additionally, the analysis of defective parts requires specialized skills and equipment, which can be costly and resource intensive. Overcoming these challenges requires manufacturers to invest in advanced technologies and skilled personnel, as well as develop robust processes for data management and logistics. Ensuring data accuracy and completeness is crucial, as incorrect or missing data can lead to flawed analyses and ineffective decision-making. Additionally, integrating data from various sources, such as dealerships, repair shops, and customer feedback, can be complex and resource intensive. Overcoming these challenges requires robust data management practices and continuous investment in technology and skills development [51] [52].

2.8 Benchmarking Analysis

In the context of automotive warranty systems, it is crucial to analyze the suppliers and their respective contributions to various automobile manufacturers. This review focuses on the prominent players in the Malaysian automotive industry, namely Company B, Company A, Company C, and Company D, and their primary system suppliers.

Company B and Company D predominantly rely on internationally acclaimed suppliers such as Company MNC and BOSCH for critical systems like air conditioning, fuel, electrical, and brakes. This choice ensures high standards of quality and reliability, reflecting their commitment to superior product performance and customer satisfaction. Company A and Proton demonstrate a balanced mix of local and international suppliers. By integrating suppliers like APM and GEELY alongside Company MNC and SANDEN, these manufacturers achieve cost efficiencies while maintaining competitive performance standards [53]. This strategy supports local industry and leverages localized expertise as shown in Table 2.1.

Table 2.1
Benchmarking Based on Product Supply Chain [53].

System	Products	Company B	Company A	Company C	Company D
Air Conditioner	A/C Auto	Company MNC	Company MNC	Geely	Company MNC
	A/C Manual	Company MNC	Company MNC Sanden	Geely	Company MNC Sanden
Electrical	Starter	Company MNC	Company MNC	Bosch	Company F
	Alternator	Company MNC	Company MNC	Bosch	Company F
	Meter	Company MNC	Company MNC	Geely	Company F
Washer System	Washer Pump	Company A	Company MNC	APM	Company F
	Nozzle	Company A	Company MNC	APM	Company F
	Wiper Arm Blade	Company MNC	Company MNC	Geely	Company F

The diversified supplier strategy highlights tailored supply chain decisions to meet market needs and company goals. The reliance on renowned suppliers for key systems underscores a prioritization of reliability and customer trust across all brands. Local suppliers are strategically used for specific components to enhance cost-effectiveness. Overall, this analysis provides a comprehensive understanding of the automotive industry's approach to quality assurance and customer satisfaction in

Malaysia. It serves as a crucial reference point for further research into automotive warranty data analysis and its implications for enhancing customer satisfaction [54].

The reliance on renowned suppliers for key systems underscores a prioritization of reliability and customer trust across all brands. Local suppliers are strategically used for specific components to enhance cost-effectiveness. This strategy not only supports local industry but also ensures that the manufacturers maintain competitive performance standards. By balancing the use of local and international suppliers, these manufacturers can achieve cost efficiencies while maintaining high standards of quality and reliability. This approach highlights the importance of strategic supply chain decisions in the automotive industry and their impact on quality assurance and customer satisfaction [55] [56].

Company MNC consistently monitors competitor activities to facilitate continuous improvement. This practice allows the Company B to identify 'best practices' for mitigating warranty claims by analyzing competitor strategies and comparing them with their own. Company E contributes approximately 10% of the total business for Alternator Assemblies supplied to Company A. This benchmarking is a critical component of competition analysis, essential for securing future business opportunities, as illustrated in Figure 2.8.

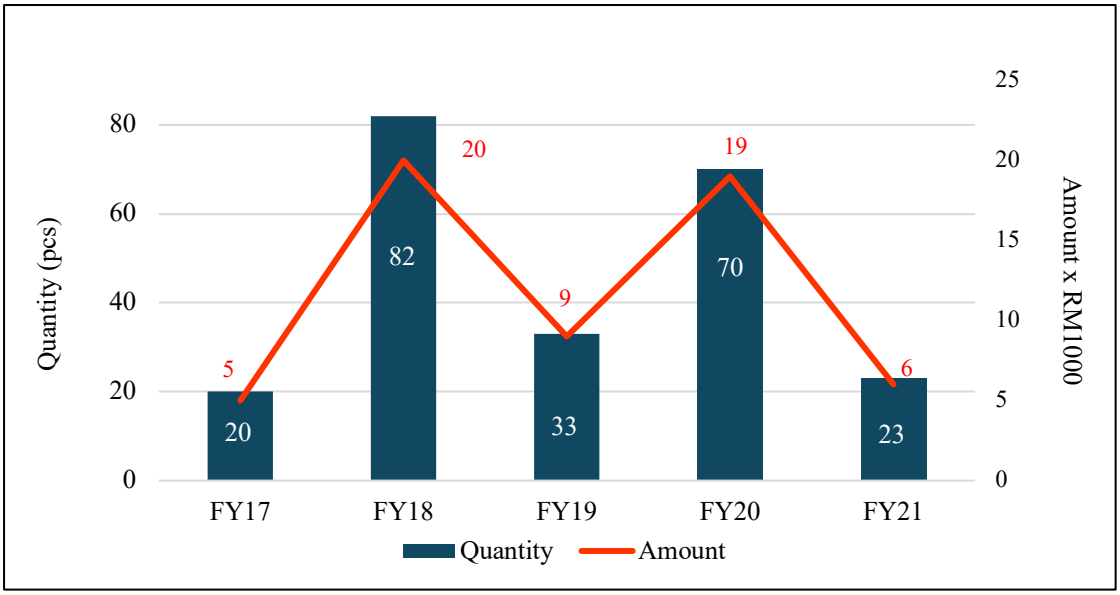


Figure 2.8 Total Warranty Claim Alternator for Company E

The analysis of suppliers and their contributions to various automobile manufacturers provides valuable insights into the automotive industry's approach to

quality assurance and customer satisfaction. By examining the strategies of prominent players in the Malaysian automotive industry, this review highlights the importance of selecting reliable suppliers and leveraging their expertise to enhance product performance and customer satisfaction. This comprehensive understanding of the automotive industry's approach to quality assurance serves as a crucial reference point for further research into automotive warranty data analysis and its implications for enhancing customer satisfaction [57] [58].

By leveraging the expertise of renowned suppliers, automotive manufacturers can ensure high standards of quality and reliability in their products. This approach not only enhances customer satisfaction but also supports the long-term success and sustainability of the automotive industry [59] [60]. The reliance on renowned suppliers for key systems underscores the importance of strategic supply chain decisions in achieving these goals. By balancing the use of local and international suppliers, manufacturers can achieve cost efficiencies while maintaining competitive performance standards [61].

The F1 score is a key performance metric for assessing predictive models in automotive warranty management, particularly in scenarios involving imbalanced datasets where warranty claims occur far less frequently than non-claim cases. By integrating precision and recall into a single measure, the F1 score provides a balanced evaluation of a model’s ability to accurately identify genuine warranty claims while limiting misclassification. This balance is essential for warranty-related applications, where both false positives and false negatives can lead to increased operational costs and suboptimal maintenance decisions.

In practice, an F1 score threshold of 0.70 is generally regarded as acceptable for warranty prediction models, while values approaching or exceeding 0.80 indicate a higher level of predictive reliability in identifying warranty claims and component failure events. The adoption of the F1 score as a primary evaluation metric supports more informed decision-making by improving claim prediction accuracy, reducing unnecessary corrective actions, and enabling proactive maintenance strategies, thereby contributing to effective warranty cost control and product quality improvement [62] [63].

2.9 Automotive Warranty Cost Forecasting Overview

2.9.1 Industry Context and Economic Impact

Warranty-related expenditure constitutes a significant financial burden within the automotive sector, with empirical studies indicating that warranty costs commonly represent between 2% and 5% of total annual revenue for major vehicle manufacturers [64]. As modern vehicles become increasingly sophisticated, incorporating thousands of interconnected mechanicals, electrical, and electronic components with varying reliability characteristics and warranty durations, the task of accurately estimating future warranty liabilities has become substantially more complex. Inaccurate forecasting directly affects financial provisioning, cost containment strategies, and long-term business sustainability [65].

This research is undertaken in collaboration with Company MNC, an automotive component manufacturer supplying parts to multiple Original Equipment Manufacturers (OEMs), including Company A. The scope of the research concentrates on forecasting warranty costs associated with alternator assemblies for selected vehicle platforms produced by Company A. A review of prevailing industry practices highlights significant structural limitations, where only approximately 45% of OEMs and 25% of tier-one suppliers maintain dedicated warranty management functions [66]. Consequently, many organisations continue to rely on basic spreadsheet-based historical extrapolation and manual accrual techniques rather than systematic, data-driven forecasting frameworks.

Company MNC reflects these industry-wide challenges through its continued dependence on Key Performance Indicator (KPI) oriented budgeting approaches, such as year-on-year warranty cost reduction targets, without the support of predictive analytical models. Such practices are labour intensive and increasingly prone to forecasting errors, particularly as warranty datasets expand in volume and complexity due to rising Total Industry Volume (TIV). Although the global automotive warranty management software market reached an estimated value of USD 2.58 billion in 2024 [67], most existing solutions primarily emphasise administrative efficiency and claim processing, rather than advanced predictive analytics or decision support [67].

As a result in Figure 2.9, inaccurate warranty forecasting remains a persistent economic risk for both component suppliers and vehicle manufacturers. Under-

provisioning exposes firms to unexpected financial liabilities, whereas excessive provisioning may undermine pricing competitiveness and profitability [68]. These risks are especially pronounced at the component level, where warranty datasets are often characterised by high volatility, relatively short historical time horizons, susceptibility to external disruptions such as supply chain instability and production interruptions [69].

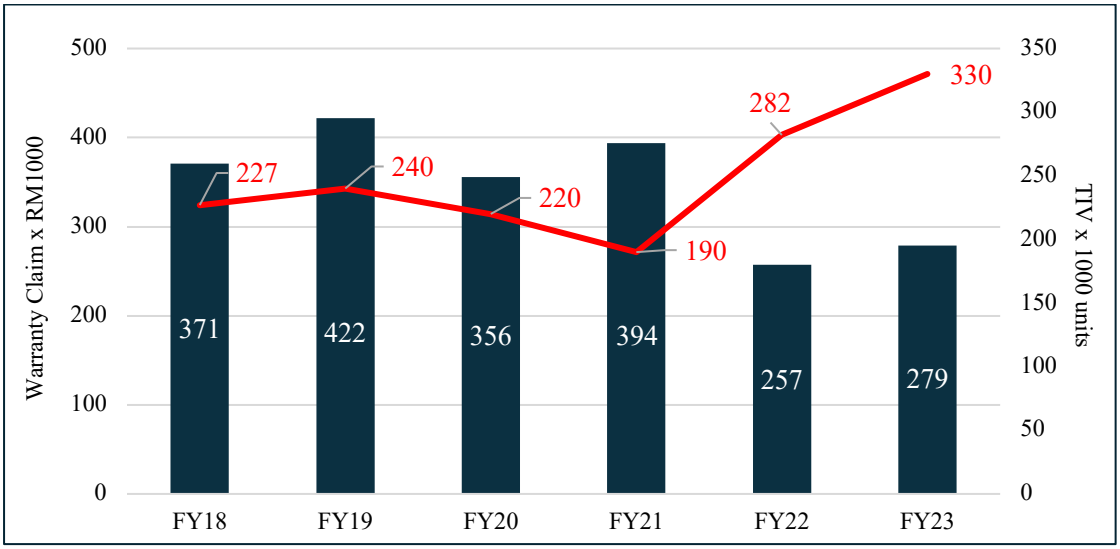


Figure 2.9 Company MNC Warranty Claim vs Vehicle Sales by Company A

Company MNC warranty claim behaviour and vehicle sales trend shown in Figure 2.9 a nonlinear relationship between market volume and warranty cost performance. Warranty claims increased from RM371K in FY2018 to RM422K in FY2019 before declining to RM356K in FY2020. Although vehicle sales reduced from 240 thousand units in FY2019 to 190 thousand units in FY2021, warranty claims remained relatively high at RM394K in FY2021. A more substantial improvement was observed in FY2022 and FY2023, where warranty claims reduced to RM257K and RM279K respectively, despite vehicle sales increasing sharply to 282 thousand and 330 thousand units.

The contrasting trend between sales growth and warranty claim reduction indicates that warranty performance was not solely dependent on vehicle sales volume. Instead, the results reflect the influence of factors such as component reliability improvement, corrective countermeasures, repair process enhancement, and warranty management effectiveness. The reduction in warranty claims during periods of higher sales volume demonstrates improved control of field quality performance and

supports the application of data driven warranty analysis to distinguish actual reliability improvement from market driven effects.

2.9.2 Traditional Forecasting Approaches

Conventional automotive warranty cost forecasting methods were originally derived from classical statistical and actuarial models developed within the insurance domain [70]. Prior to the emergence of advanced data analytics, these approaches formed the primary analytical foundation for warranty planning and cost estimation. Actuarial-based techniques, in particular, represent the earliest structured framework for warranty forecasting, drawing upon reliability theory, life-cycle analysis, and probability-based failure modelling [71].

In these models, expected warranty expenditure is typically estimated using historical failure frequencies, component reliability metrics, and predefined warranty coverage conditions. Probabilistic distributions are applied to project future claim volumes and associated costs, under the assumption that failure behaviour remains relatively stable over time [70]. While such assumptions simplify model formulation, they often fail to capture the dynamic nature of automotive product evolution, manufacturing improvements, and market-driven variability.

Within industry practice, basic trend extrapolation remains the most widely adopted forecasting method due to its minimal computational requirements and straightforward implementation. Spreadsheet-based tools, particularly Microsoft Excel, are commonly used to project warranty costs using linear trend lines, moving averages, and production volume adjustment factors linked to sales forecasts [72]. Although these techniques offer clarity and accessibility for managerial decision-makers, their effectiveness has diminished in contemporary automotive environments characterised by high data dimensionality, complex failure interactions, and rapidly evolving product architectures.

Traditional forecasting approaches are inherently constrained in their ability to address the complexity of modern vehicle systems. Linear trend-based methods are unable to capture nonlinear relationships between warranty outcomes and influencing factors such as design revisions, supplier changes, or process improvements [73]. Furthermore, these approaches lack robustness when confronted with structural breaks arising from external disruptions, including production interruptions or supply chain

disturbances, rendering historical patterns unreliable under changing conditions [69]. The limited capacity of traditional methods to incorporate multiple interacting variables further restricts their effectiveness in component-level warranty analysis.

Another fundamental limitation lies in the static nature of conventional forecasting techniques. Model updates typically require manual recalibration or adjustment of underlying assumptions, limiting responsiveness to evolving warranty patterns [72]. This shortcoming is particularly pronounced for component manufacturers supplying multiple Original Equipment Manufacturers (OEMs), where warranty behaviour can vary substantially across vehicle platforms, operating environments, and customer usage profiles.

2.9.3 Evolution Towards Advanced Analytics

The transition of the automotive industry towards advanced analytical methodologies has been largely motivated by the growing complexity of vehicle architectures and heightened competitive demands within the global market [74]. Contemporary vehicles integrate advanced electronic systems, software-controlled functions, and highly interdependent subsystems, resulting in failure mechanisms that are nonlinear and difficult to characterise using traditional forecasting approaches. Concurrently, manufacturers face increasing pressure to control warranty expenditure while sustaining high product quality and regulatory compliance, thereby requiring forecasting techniques that are more precise, adaptive, and responsive to changing operational conditions [75].

The widespread adoption of big data technologies has further transformed the landscape of warranty forecasting. Modern automotive products continuously generate extensive volumes of operational and condition-monitoring data through onboard diagnostics, telematics systems, and connected vehicle platforms [76]. When these data streams are integrated with manufacturing histories, supplier performance records, and usage environment information, they create a comprehensive data ecosystem capable of supporting more accurate and granular predictive analysis. For component suppliers such as Company MNC, access to such multi-source datasets enables the examination of warranty behaviour across diverse customers, vehicle models, and operating contexts, thereby enhancing understanding of component-level reliability and failure propagation patterns [77].

The practical deployment of advanced warranty analytics has been enabled by parallel developments in digital infrastructure and computational tools. Cloud-based computing platforms provide scalable and cost-effective resources for processing large and complex datasets, while contemporary machine learning libraries support the implementation of advanced predictive algorithms [78]. In addition, the increasing maturity of open-source programming environments, particularly Python and R, has significantly broadened access to statistical modelling, machine learning, and data mining techniques. These tools allow organisations to develop sophisticated forecasting solutions without reliance on high-cost proprietary software ecosystems [79], [80] [81]. Collectively, these technological advancements have substantially reduced both the technical complexity and financial barriers associated with adopting data-driven warranty forecasting frameworks.

Despite these advancements, widespread adoption of advanced analytics within the automotive industry remains constrained by organisational and technical challenges. Many firms lack the specialised expertise required to develop and maintain advanced analytical systems [81], while legacy information infrastructures often limit data integration and processing capabilities. Furthermore, organisational resistance to change and concerns over model interpretability among business stakeholders accustomed to traditional trend-based forecasts continue to impede adoption [82]. These barriers help explain the persistent reliance on conventional forecasting methods, despite the demonstrated advantages of advanced analytical techniques.

2.10 Statistical Methods for Warranty Prediction

2.10.1 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) constitutes a fundamental methodological stage in achieving the research objective of developing a predictive warranty management system. EDA is applied to systematically examine warranty data in order to identify underlying patterns, trends, and anomalies that are critical for determining the root causes of warranty claims. In this study, EDA is conducted at the initial stage of analysis prior to predictive model development to establish a comprehensive understanding of the data structure, data quality, and inherent relationships among variables.

Exploratory Data Analysis (EDA) was conducted as a foundational step to ensure the reliability and suitability of the warranty dataset prior to predictive model development. This stage focused on systematically examining data structure, quality, and variable behaviour using established statistical and visual analytical techniques. Descriptive statistical measures were applied to summarise central tendencies and variability, providing an initial assessment of data distribution and consistency. Complementary visualisations, including histograms, scatter plots, and box plots, were employed to identify distribution characteristics, detect outliers, and observe potential relationships among key variables. To assess linear dependency between warranty-related parameters, Pearson's correlation coefficient was utilised, where coefficients exceeding 0.5 were considered indicative of significant relationships, particularly between vehicle usage indicators such as mileage and warranty claim frequency. Furthermore, Principal Component Analysis (PCA) was implemented to address dimensionality challenges inherent in large-scale warranty datasets. By retaining the most influential components, PCA reduced data redundancy and improved model efficiency without compromising explanatory power. Consistent with findings reported in recent literature, the integration of structured EDA with subsequent machine learning analysis enhances prediction accuracy and supports proactive warranty and quality management strategies within the automotive industry [62] [83].

2.10.2 Process of Exploratory Data Analysis

The EDA process adopted in this study is structured into several key stages, each contributing to a robust understanding of the warranty dataset.

a) Data Collection and Cleaning:

The first stage of EDA involves the collection of relevant warranty-related data from multiple sources, including historical warranty claim records, vehicle usage information, and component-level failure data. The dataset comprises multiple variables such as vehicle age, mileage, model type, failed component, failure description, and claim cost. Data cleaning procedures are subsequently applied to address missing values, inconsistencies, and duplicate records. These steps ensure that the dataset is accurate, consistent, and suitable for subsequent statistical analysis and modelling.

b) Descriptive Statistics:

Descriptive statistical analysis is conducted to summarise the central tendency, variability, and distributional characteristics of the warranty data. Key statistical measures include mean, median, standard deviation, variance, skewness, and kurtosis. These metrics provide a high-level overview of warranty claim behaviour and facilitate the early identification of trends, dispersion patterns, and abnormal claim distributions that may indicate underlying quality or reliability issues.

c) Data Visualization:

Data visualisation techniques are employed to enhance interpretability and reveal relationships that may not be apparent through numerical summaries alone. Visual tools such as histograms, box plots, scatter plots, and correlation matrices are utilised to examine data distributions, identify trends, and detect potential anomalies. For instance, scatter plots are used to assess the relationship between vehicle mileage and warranty claim frequency, while box plots highlight variations in claim costs across different vehicle models or components.

d) Correlation Analysis:

Correlation analysis is applied to quantify the strength and direction of relationships between key variables in the warranty dataset. The Pearson correlation coefficient (*r*) is used to evaluate linear associations between continuous variables, such as mileage, claim frequency, and warranty cost.

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}} \quad (2.1)$$

Where:

- Xi and Yi represent the data points for variables X and Y,
- $\bar{X}$ and $\bar{Y}$ are the means of X and Y,
- n is the number of data points.

The coefficient ranges from −1 to +1, where values close to ±1 indicate strong correlations and values near zero indicate weak or no linear relationship. Identifying

significant correlations supports effective feature selection by highlighting variables that are most relevant for predictive model development.

e) Outlier Detection:

Outlier detection is a critical component of EDA, as extreme values can disproportionately influence model outcomes and distort forecasting accuracy. Outliers are identified using both visual methods (e.g., box plots) and statistical techniques such as z-score analysis.:

$$z = (Xi - \bar{X}) / \sigma \tag{2.2}$$

Where:

- Xi is the individual data point,
- $\bar{X}$ is the mean of the dataset,
- σ is the standard deviation.

A z-score beyond ±3 typically indicates an outlier.

Observations with z-scores exceeding ±3 are treated as potential outliers. In warranty data, such extreme values may represent unusually high-cost claims or abnormal failure events that require separate investigation rather than direct inclusion in model training.

2.10.3 Linear Regression Applications

Linear regression remains one of the most widely adopted statistical techniques in automotive warranty cost forecasting, primarily due to its transparency and ease of interpretation [84]. As a parametric modelling approach, it operates under the assumption of a linear association between independent variables and the dependent response, making it particularly suitable for exploratory analysis and baseline model development. Within warranty management contexts, linear regression is frequently employed to identify dominant cost contributors and to establish reference benchmarks

against which more complex predictive techniques can be evaluated. The simplest form of linear regression used for warranty cost estimation can be expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{2.3}$$

where Y represents the warranty cost, X_1 denotes an explanatory variable, β_0 is the intercept term, β_1 is the regression coefficient associated with the predictor, and ε represents the stochastic error component. Early warranty forecasting studies adopted this formulation due to its computational simplicity and its ability to quantify both the direction and relative influence of factors such as production volume, vehicle ageing, and operational exposure on warranty expenditure [85].

To better reflect the temporal characteristics inherent in warranty datasets, linear regression models have been extended to include time-dependent structures. Dynamic linear models incorporate autoregressive elements, enabling present warranty outcomes to be partially influenced by historical observations, thereby improving model responsiveness to evolving claim patterns [68].

Building upon this foundation, multiple linear regression expands the analytical framework by allowing the simultaneous inclusion of several explanatory variables, as represented by:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \tag{2.4}$$

This multivariate formulation enables analysts to assess the combined impact of multiple influencing factors while controlling for interdependencies among predictors [85], [86]. Empirical evidence from actuarial science and insurance forecasting literature indicates that regression models incorporating economic indicators—such as labour rates, wage indices, and material price fluctuations—consistently outperform univariate trend-based approaches. These findings are highly relevant to automotive warranty forecasting, where claim costs are similarly shaped by macroeconomic conditions, labour expenses, and component pricing dynamics.

Recent automotive focused research has further demonstrated the applicability of Multiple Linear Regression Analysis (MLRA) for structured warranty forecasting. A comprehensive study on warranty claim quantity prediction employed MLRA to

analyse automotive headlamp warranty data collected from January 2019, comprising 248 recorded claims [68]. The study developed a regression-based forecasting model to represent defect quantity trends over a standard 36-month warranty cycle. This work directly addressed a common industry limitation, where warranty budgeting and accrual practices often rely on rough estimation without a sound mathematical or statistical foundation, leading to either underestimation that erodes profitability or overestimation that weakens market competitiveness.

Through systematic variable validation procedures, including normality assessment, probability analysis, and multicollinearity testing, the researchers derived a statistically robust regression model, expressed as $Y_1 = -26.3 + 32.6X_3$, to predict cumulative defect quantities. The model forecasted an additional 952 warranty defects over the remaining warranty period, demonstrating the practical value of regression-based approaches in warranty planning. This study illustrates how structured statistical methodologies can elevate warranty management practices from heuristic estimation to evidence-based forecasting, thereby providing manufacturers with more reliable tools for warranty cost prediction and informed budgeting decisions across automotive supply chain operations.

2.10.3.1 Linear Regression

Linear regression was used as the baseline forecasting model to provide a clear, transparent, and interpretable reference for warranty cost prediction. This approach aligns with industry practice, where explainability is essential for financial and warranty-related decision-making. The model was developed using 48 monthly observations covering FY2019 and FY2021–FY2023, with FY2020 excluded due to disruptions caused by the COVID-19 pandemic. The response variable for the model was the total monthly alternator warranty cost (Total_RM).

a) Dataset Preparation and Feature Selection

The regression model utilised the basic regression feature set to maintain a conservative feature-to-observation ratio and reduce the risk of overfitting. Pure time identifiers were excluded, while predictive variables included linear time trends, fiscal

month indicators, fiscal year dummy variables, vehicle model proportions, and a COVID-19 impact indicator.

To further control model complexity, univariate feature selection was conducted using the SelectKBest method with f-regression scoring. The number of predictors was limited to a maximum of ten features, which is appropriate for the dataset size. Multicollinearity was assessed by examining pairwise correlations, with highly correlated feature pairs ($|r| > 0.8$) identified and reviewed. Individual correlations between each predictor and Total_RM were also analysed to support informed and transparent feature selection.

b) Model Training and Regularisation

Three linear regression variants were evaluated:

- All-features model,
- Selected-features model, and
- Regularised model.

Ridge regression was implemented using RidgeCV, with regularisation parameters $\alpha \in \{0.1, 1.0, 10.0, 100.0\}$ selected via cross-validation. StandardScaler preprocessing was applied to normalise feature scales and enable meaningful comparison of standardised coefficients for interpretability and feature importance assessment.

c) Validation and Performance Assessment

Model performance was evaluated using multiple metrics, namely R^2, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), to provide a balanced assessment of explanatory power and prediction accuracy. TimeSeriesSplit cross-validation with three folds was applied to preserve the temporal structure of the data while estimating out-of-sample performance. The number of folds was intentionally limited due to the small dataset size.

Residual diagnostics were performed to assess key linear regression assumptions. Autocorrelation was evaluated using the Durbin–Watson test, with acceptable values between 1.5 and 2.5. Heteroscedasticity was examined using the Breusch–Pagan test, while residual normality was assessed through the Shapiro–Wilk

test. In addition, diagnostic plots were generated to visually examine model fit, residual patterns, temporal prediction behaviour, and error distribution.

The final regression model was selected based on cross-validated RMSE performance. This model served as a statistical benchmark for comparison with more advanced forecasting approaches, while its limitations related to small sample size and sensitivity to modelling assumptions were explicitly acknowledged.

2.10.3.2 Random Forest

Random Forest regression was implemented as the first machine learning approach, leveraging ensemble learning to capture non-linear relationships while controlling overfitting through bootstrap aggregation.

a) Feature Set and Time-Series Validation

The Random Forest model utilised the enhanced regression feature set (23 features), incorporating temporal indicators, business composition metrics, and selected lag features. A custom time-series cross-validation framework with three folds and a minimum 60% training window preserved temporal causality while providing realistic out-of-sample evaluation.

b) Conservative Hyperparameter Optimisation

Hyperparameter tuning employed a restricted grid:

n_estimators: {50, 100, 200},
max_depth: {None, 10, 15, 20},
min_samples_split: {2, 5, 10},
min_samples_leaf: {1, 2, 4},
max_features: {sqrt, log2, 0.3}.

The search space was limited to 20 configurations, prioritising generalisation performance over training accuracy.

c) Feature Importance Analysis and Overfitting Control

Feature importance analysis using mean decrease impurity indicated strong dominance of temporal features (60–80%), with business features contributing 15–25%. This confirmed the primacy of time-based drivers in warranty cost behaviour. Ensemble

averaging and feature subsampling provided implicit regularisation, reducing overfitting risk relative to single-tree models.

2.10.3.3 XGBoost

XGBoost was implemented as the advanced gradient boosting method, with explicit emphasis on overfitting prevention informed by Random Forest findings.

a) Conservative Configuration and Preprocessing

Baseline parameters were intentionally conservative: 50 estimators, maximum depth of 3, learning rate of 0.1, subsampling ratios of 0.8, and combined L1/L2 regularisation ($\alpha = 1.0$, $\lambda = 1.0$). StandardScaler preprocessing was applied to stabilise gradient descent optimisation under mixed feature scales.

b) Regularisation-Focused Hyperparameter Tuning

Hyperparameter tuning prioritised regularisation effectiveness rather than complexity, with strict overfitting gap thresholds (training–validation R^2 gap < 0.5). Parameter ranges were deliberately narrow to ensure stability under small-sample conditions.

c) Gradient Boosting Methodology and Validation

Model evaluation emphasised Expected Out-of-Sample R^2, recognising that in-sample accuracy alone is insufficient for forecasting reliability. XGBoost demonstrated improved balance between predictive performance and generalisation relative to Random Forest under limited data conditions.

2.10.4 Phase 4: Model Walk-Forward Validation Framework

A three-iteration walk-forward validation strategy was implemented to evaluate forecasting reliability under realistic deployment conditions while accounting for data gaps and structural disruptions.

2.10.4.1 Iterative Validation Design

- Iteration 1: FY2019 → predict FY2021
- Iteration 2: FY2019 + FY2021 → predict FY2022
- Iteration 3: FY2019 + FY2021 + FY2022 → predict FY2023

This design evaluated minimum training requirements, learning gains from expanded datasets, and robustness to non-contiguous training periods.

2.10.4.2 Conservative Model Configurations

Validation employed conservative configurations across all models to prioritise generalisation:

- Linear models: 10–23 features,
- Random Forest: 20 estimators, max_depth = 5,
- XGBoost: 10 estimators, max_depth = 2, high regularisation.

2.10.4.3 Comprehensive Performance Assessment

Evaluation metrics included R^2, RMSE, MAPE, overfitting gap analysis, directional accuracy, and stability across iterations. Composite scoring balanced accuracy, robustness, and business relevance.

2.10.4.4 Cross-Iteration Analysis

Cross-iteration analysis tracked performance evolution as training data expanded, enabling evidence-based determination of optimal complexity and minimum data requirements for deployment.

2.10.5 Phase 5: Production Forecasting

The production forecasting phase implemented multi-model comparison methodology to generate FY2024 warranty claim cost forecasts, employing systematic retraining and ensemble techniques for comprehensive forecasting deployment.

2.10.5.1 Unified Training Dataset

All models were retrained using a unified dataset (FY2019, FY2021–FY2023), ensuring consistent temporal coverage. Model-specific feature sets were applied to maintain algorithm compatibility.

2.10.5.2 Production Model Training Implementation

Five production models were trained:

- Linear Basic,
- Linear Enhanced,
- ARIMA,
- Random Forest Conservative,
- XGBoost Ultra-Conservative.

2.10.5.3 FY2024 Forecast Generation Methodology

FY2024 forecasts were generated for April 2024–March 2025 using consistent temporal feature extrapolation. Ensemble forecasts were produced via simple averaging to reduce variance and improve robustness.

2.10.5.4 Comprehensive Comparison Framework

Comparison included monthly forecasts, annual totals, and expected out-of-sample performance. Visualisation panels illustrated historical continuation, annual comparisons, and monthly patterns.

2.10.5.5 Validation Actual Data and Identify the Best Performing Model

Final model evaluation compared FY2024 forecasts against actual warranty cost data using MAE, RMSE, MAPE, bias, total deviation, and directional accuracy. Models were ranked primarily by MAPE, with supporting qualitative assessment of trend and seasonality capture. This comprehensive evaluation identified the most reliable forecasting model for strategic warranty cost planning and continuous deployment.

2.10.6 ARIMA Time Series Models

Autoregressive Integrated Moving Average (ARIMA) models are widely used for forecasting time-dependent warranty trends in the automotive industry because they are capable of modelling temporal behavior and non-stationary data patterns [87]. In warranty analysis, ARIMA is suitable for evaluating monthly warranty cost trends, recurring claim behavior, and seasonal fluctuations associated with vehicle operation and component reliability. The Box Jenkins methodology provides a structured forecasting process involving data preparation, model identification, parameter estimation, diagnostic validation, and forecasting evaluation [88].

In this study, the ARIMA model was implemented using monthly total warranty claim cost data (Total_RM) collected over a 48-month period. The forecasting analysis focused on alternator assembly warranty claims; particularly recurring noise complaints associated with design related weaknesses involving bearing degradation and internal component instability during prolonged operating conditions. These recurring failures contributed to fluctuating monthly warranty trends and created temporal forecasting challenges within the dataset. Prior to modelling, the warranty data underwent data preparation and standardization processes, including datetime indexing, missing value treatment using linear interpolation, and trend decomposition analysis. Additive seasonal decomposition with a 12-month cycle was applied to separate the trend, seasonal, and residual components to better understand warranty behavior patterns. This process supported more stable parameter selection and reduced forecasting bias caused by irregular warranty fluctuations.

The ARIMA model is generally represented as ARIMA (p, d, q), where p represents the autoregressive order, d represents the differencing order used to achieve stationarity, and q represents the moving average order [23]. The general ARIMA equation is expressed as follows:

$$\Delta^d y_t = \delta + \sum_{i=1}^{p} \theta_i \, \Delta^d y_{t-i} + e_t - \sum_{j=1}^{q} \alpha_j \, e_{t-j} \tag{2.5}$$

where Δ^d denotes differencing of order d, y_t represents the warranty cost (or claim quantity) at time t, θ_i are the autoregressive coefficients capturing the influence of past

observations, α_j are the moving average coefficients representing the impact of past forecast errors, δ is a constant term, and e_t denotes a white noise error process [23]. This formulation enables ARIMA models to represent both the persistence effects inherent in warranty cost evolution and the influence of random shocks arising from unexpected quality issues or short-term operational disruptions.

Applications of ARIMA models in automotive warranty forecasting have demonstrated notable advantages over traditional trend-based approaches. Prior research implementing the Box–Jenkins methodology addressed a critical gap in warranty time-series analysis, where conventional practices often lack mathematical rigor and rely heavily on heuristic estimation [89]. In this study, 60 monthly observations of warranty claim cost data spanning January 2015 to December 2019 were analysed using the systematic Box–Jenkins procedure, including data preparation, model identification, parameter estimation, diagnostic validation, and forecasting. An optimal ARIMA (2,1,1) model was developed and validated through comprehensive diagnostic testing, including Anderson–Darling normality assessment, autocorrelation analysis, and residual white-noise verification.

The results demonstrated that ARIMA models are particularly effective in capturing seasonality effects inherent in warranty data without requiring external explanatory variables. The model successfully identified recurring seasonal patterns, such as increased engine start-related complaints during colder periods and higher occurrences of headlamp water penetration during rainy seasons. Based on the validated model, two-year warranty cost forecasts were generated for the period January 2020 to December 2021, with a total projected warranty cost of RM 4,475.83. This study established the ARIMA Box–Jenkins methodology as a mathematically robust and reliable forecasting tool for automotive warranty cost management, offering clear advantages over traditional rough estimation methods commonly employed in warranty budgeting and accrual processes.

To further address systematic seasonal variations, Seasonal ARIMA (SARIMA) models extend the standard framework by incorporating seasonal autoregressive and moving average components. Denoted as ARIMA(p, d, q) × (P, D, Q), these models enable explicit representation of calendar-based seasonality, dealer operational cycles, and periodic vehicle sales patterns, which are frequently observed in automotive warranty datasets [87]. As such, SARIMA models provide an enhanced analytical

capability for accurately modelling and forecasting warranty costs in environments characterised by strong seasonal effects.

2.10.6.1 Stationarity and Autocorrelation Testing

a) Overview of Stationarity Testing

Stationarity testing was conducted to determine whether the statistical properties of the automotive warranty cost time series remained stable over time. Stationarity is a fundamental prerequisite for reliable time-series forecasting, particularly for ARIMA-based models, as non-stationary data may lead to spurious relationships and unreliable forecasts. Inaccurate stationarity assumptions could adversely affect warranty cost prediction and compromise financial planning for Company A.

To provide robust and confirmatory evidence, two complementary econometric tests were applied: the Augmented Dickey–Fuller (ADF) test and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. The ADF test evaluates the presence of a unit root, indicating non-stationarity, while the KPSS test examines the null hypothesis of stationarity. The combined use of both tests strengthens inferential reliability, particularly in small-sample environments. These tests were applied to the 48-month aggregated warranty cost series using established econometric procedures.

b) Augmented Dickey-Fuller Test Implementation

The ADF test was employed to assess the null hypothesis that the warranty cost series contains a unit root. The test regression is expressed as:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \, \Delta y_{t-i} + \varepsilon_t \tag{2.6}$$

where Δy_t represents the first-differenced series, α is a constant, βt captures deterministic trend effects, and $\varepsilon_t$denotes the error term. The ADF test was selected due to its ability to account for autocorrelation in the residual structure, which is commonly observed in warranty cost data.

Multiple model specifications were evaluated, including models with a constant term and with both constant and trend, to ensure robustness across alternative data-generating processes. The maximum lag length was conservatively determined using

$$\min\left(12, \left\lfloor \left(\frac{n}{3}\right)^{\frac{1}{3}} \right\rfloor\right) \tag{2.7}$$

where n represents the sample size. This selection follows the small-sample guidance proposed by MacKinnon, ensuring an appropriate balance between retaining sufficient degrees of freedom and achieving reliable statistical inference. The criterion is particularly suitable for the present dataset, which comprises 48 observations, as it avoids over-parameterisation while maintaining robustness in hypothesis testing.

c) KPSS Confirmatory Testing

The Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test was applied as a confirmatory stationarity assessment to complement the Augmented Dickey–Fuller (ADF) test. Unlike ADF, which assumes a unit root as the null hypothesis, the KPSS test adopts stationarity as the null, thereby providing a balanced evaluation of time-series properties. The KPSS statistic is derived from the partial sum of regression residuals and evaluates deviations from level or trend stationarity using a long-run variance estimator.

Rejection of the null hypothesis indicates non-stationary behaviour, while non-rejection supports stationarity. In this study, joint interpretation of ADF and KPSS results ensured reliable stationarity classification prior to model development. This dual-testing strategy reduces the risk of spurious regression and strengthens the statistical validity of warranty cost forecasting models. The test statistic is defined as:

$$LM = \frac{1}{n^2\sigma^2}\sum_{t=1}^{n} S_t^2 \tag{2.8}$$

where S_t represents the cumulative residuals and σ^2 is the long-run variance. Both level-stationary and trend-stationary specifications were evaluated. Automatic lag selection was employed to appropriately account for serial correlation in the residuals. Agreement between ADF and KPSS results provided strong confirmation of the stationarity properties of the warranty cost series.

d) Autocorrelation Function Analysis

Temporal dependency structures in the warranty cost time series were examined using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) analysis. The sample autocorrelation at lag k was computed as $\rho_k = \gamma_k / \gamma_0$, where γ_k represents the autocovariance at lag k and γ_0 denotes the variance. ACF and PACF plots were used to identify persistence behaviour, cutoff patterns, and gradual decay characteristics, which are essential for determining appropriate autoregressive and moving average terms in ARIMA model specification. Statistical significance bounds were defined at $\pm 1.96/\sqrt{n}$ to account for sampling variability in the limited dataset. The maximum lag length was restricted to maintain robustness and prevent overfitting, ensuring reliable identification of temporal dependence prior to model estimation.

$$\min\left(20, \frac{n}{4}\right) \tag{2.9}$$

allowing investigation of short-term dependencies and potential seasonal effects associated with the 12-month fiscal cycle.

e) Ljung-Box Autocorrelation Testing

The Ljung–Box test was employed to formally evaluate the presence of residual autocorrelation in the warranty cost time series. The test statistic is defined as:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\rho_k^2}{n-k} \tag{2.10}$$

where n denotes the sample size, h represents the maximum lag, and ρ_k is the sample autocorrelation at lag k. The test was conducted at lag lengths of 5, 10, 15, and 20 months to assess short-term dependencies and potential seasonal effects. This procedure complemented ACF-based diagnostics by providing exact significance levels, thereby supporting robust validation of ARIMA model suitability for warranty cost forecasting.

f) First Difference Analysis

When required, first differencing was applied using the transformation:

$$\Delta y_t = y_t - y_{t-1} \tag{2.11}$$

to achieve stationarity. The differenced series was subjected to the same ADF and KPSS testing procedures to confirm stationarity and determine the appropriate integration order (d) for ARIMA specification. This ensured that any differencing applied was both statistically justified and methodologically consistent.

2.11 Machine Learning Applications in Warranty Forecasting

2.11.1 Random Forest Methods

Random Forest is a powerful ensemble machine learning technique that has gained increasing attention in automotive warranty forecasting due to its ability to model complex, non-linear relationships while maintaining robust predictive performance across heterogeneous datasets. By aggregating the outputs of multiple decision trees through bootstrap aggregating (bagging), Random Forest produces more stable and accurate predictions than individual tree-based models [90].

The Random Forest algorithm constructs a large number of decision trees using random subsets of both training observations and input features, following the principle that ensemble learning can outperform single estimators through collective decision-making [91]. Each decision tree is trained on a bootstrap sample drawn from the original dataset, while only a randomly selected subset of features is considered at each node split. This dual randomness—across observations and features—effectively reduces overfitting, a common limitation of standalone decision trees, while preserving strong predictive capability.

From a mathematical perspective, for a Random Forest consisting of B trees, the final prediction is obtained by averaging the individual tree predictions: [92]

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} f_b\,(x) \tag{2.12}$$

where $f_b(x)$ represents the prediction from the b-th decision tree and x denotes the input feature vector [92]. This ensemble averaging mechanism reduces prediction variance while maintaining low bias, which is particularly advantageous in warranty forecasting

applications where component failure behaviour may vary across products, suppliers, and usage conditions.

Empirical studies have demonstrated that Random Forest models can significantly outperform traditional human-based forecasting approaches in warranty cost estimation. Using proprietary data from a major European manufacturing firm, researchers reported that Random Forest-based models achieved prediction errors that were 18.4–25.8% lower than conventional managerial estimates. In addition, these models reduced both the frequency and magnitude of warranty cost overestimations. Notably, the Random Forest approach maintained superior predictive accuracy even under challenging conditions such as new product launches, where historical data availability is limited. These findings highlight the practicality and robustness of Random Forest models as reliable tools for improving financial forecasting accuracy in manufacturing environments where warranty provisions represent critical business estimates [93].

2.11.2 Gradient Boosting Techniques

Gradient boosting represents an advanced ensemble learning methodology that has demonstrated strong performance in automotive warranty forecasting through its sequential learning strategy and optimisation-based framework. In contrast to the parallel ensemble construction used in Random Forest, gradient boosting builds models sequentially, with each new model focusing on correcting the prediction errors of the preceding models [94].

The fundamental principle of gradient boosting involves iteratively adding weak learners typically shallow decision trees to minimise a predefined loss function using gradient descent optimisation. The process begins with an initial prediction, after which successive models are trained to approximate the negative gradient of the loss function associated with the previous iteration's residual errors [95]. This additive and corrective learning process enables gradient boosting models to capture subtle, high-order interactions among variables, making them well suited for warranty forecasting scenarios where small deviations can lead to significant cost implications.

The gradient boosting prediction can be mathematically expressed as:

$$F(x) = F_0(x) + \sum_{m=1}^{M} \gamma_m \, h_m(x) \tag{2.13}$$

where $F_0(x)$ denotes the initial prediction, $h_m(x)$ represents the m-th weak learner, γ_m is the learning rate or step size, and M is the total number of boosting iterations [95]. By iteratively refining predictions, gradient boosting achieves high forecasting accuracy while maintaining flexibility in modelling complex data structures.

Among gradient boosting implementations, Extreme Gradient Boosting (XGBoost) has emerged as a leading algorithm for warranty forecasting applications due to its computational efficiency and enhanced algorithmic design [95]. XGBoost incorporates advanced features such as regularised objective functions, optimised tree construction methods, and effective handling of missing values, all of which address common challenges encountered in large-scale warranty datasets [96].

The XGBoost objective function integrates both prediction accuracy and model complexity through regularisation terms and is defined as:

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2.14}$$

where $l(y_i, \hat{y}_i)$ denotes the loss function measuring prediction error, $\Omega(f_k)$ represents the regularisation penalty for the k-th decision tree, and K is the total number of trees in the ensemble [96]. The inclusion of L1 and L2 regularisation terms mitigates overfitting, which is particularly important in warranty forecasting contexts where historical patterns may not fully represent future cost behaviour due to product changes, supplier variation, or market disruptions.

2.11.3 Ensemble Methods

Ensemble methods represent an advanced class of machine learning techniques that integrate predictions from multiple heterogeneous models to achieve higher forecasting accuracy and robustness than individual models operating in isolation [97]. Within the context of automotive warranty forecasting, ensemble approaches are particularly valuable as they mitigate the limitations inherent in single-model strategies

by exploiting the complementary strengths of diverse analytical methods, including classical statistical models, tree-based learners, and gradient boosting algorithms [98]. The theoretical basis of ensemble learning is grounded in the principle that aggregating predictions from models with partially uncorrelated error structures can reduce overall prediction variance while preserving, or even enhancing, predictive accuracy [99]. This variance reduction effect is especially beneficial for warranty forecasting, where individual models may respond differently to nonlinearity, seasonality, noise, or structural shifts in warranty claim behaviour caused by product updates, supplier changes, or market disruptions.

Ensemble learning theory further indicates that performance gains are maximised when the combined models exhibit algorithmic diversity [100]. In automotive warranty applications, such diversity can be achieved by integrating fundamentally different modelling paradigms. For example, linear regression models are effective in capturing global linear cost trends, ARIMA models address temporal dependence and non-stationary patterns, Random Forest algorithms model complex nonlinear interactions among variables, and XGBoost models optimise predictive performance through iterative error correction and regularisation. The integration of these distinct approaches allows ensemble systems to capture a broader range of warranty cost dynamics than any single model alone.

Empirical evidence from related application domains supports the effectiveness of ensemble-based forecasting strategies. Studies in transportation and logistics cost prediction have shown that ensemble models incorporating Random Forest components significantly outperform standalone machine learning models by capturing complex nonlinear relationships that traditional forecasting techniques often fail to detect [95]. These findings highlight the potential of ensemble frameworks to enhance forecasting reliability and accuracy, reinforcing their suitability for warranty cost estimation where precise predictions are essential for effective budgeting, accrual planning, and cost optimisation

2.12 Comparative Studies: AI vs Statistical Methods

2.12.1 Performance Comparison Literature

Comparative studies investigating the relative performance of artificial intelligence (AI) and traditional statistical methods in warranty forecasting present a nuanced and context-dependent landscape, where model superiority is strongly influenced by data characteristics, forecasting horizons, and practical application requirements [31]. The evolving body of literature indicates that while machine learning techniques often achieve higher statistical accuracy metrics, their practical advantages are not universally consistent across all warranty forecasting scenarios.

Existing comparative research spans a broad range of methodologies, encompassing conventional statistical approaches and advanced machine learning models. Statistical techniques such as linear regression, ARIMA, and econometric models offer well-established theoretical foundations, transparency, and interpretability, which remain valuable for business decision-making and regulatory contexts. In contrast, machine learning approaches including Random Forest, gradient boosting methods, and neural networks demonstrate superior capability in modelling complex, nonlinear relationships and high-dimensional interactions commonly present in modern warranty datasets [101]. As a result, recent studies increasingly emphasise the importance of selecting forecasting approaches based on application context rather than relying on a single universally optimal method.

2.12.2 Performance Metrics in Comparative Studies

To enable objective comparison of forecasting accuracy across different methodological approaches, contemporary studies employ a set of standardised evaluation metrics. The most commonly adopted metrics in warranty forecasting research include the following:

a) Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \times 100\% \tag{2.15}$$

where A_i denotes the actual observed value and F_i represents the forecasted value. MAPE provides a scale-independent measure, enabling comparison across warranty datasets with differing cost magnitudes [102].

b) Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (A_i - F_i)^2} \tag{2.16}$$

RMSE assigns greater weight to larger prediction errors due to the squaring operation, making it particularly relevant for warranty applications where large forecasting errors can result in disproportionate financial consequences [49].

c) Coefficient of Determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (A_i - F_i)^2}{\sum_{i=1}^{n} (A_i - \bar{A})^2} \tag{2.17}$$

where $\bar{A}$ represents the mean of the observed values. R^2 measures the proportion of variance in the dependent variable explained by the model and is commonly used to assess overall model fit quality [102].

d) Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |A_i - F_i| \tag{2.18}$$

MAE provides a robust error measure that is less sensitive to extreme values, which is advantageous for warranty datasets that may contain sporadic high-cost claim events [102].

2.12.3 Expected Out-of-Sample Performance

Assessment of expected out-of-sample performance is essential for evaluating model generalisation capability, particularly in warranty forecasting contexts where limited historical data restricts conventional validation approaches [50]. This evaluation framework enables estimation of likely predictive performance for future warranty periods beyond the available training data, thereby supporting more informed model selection decisions in small-sample environments.

Expected R^2 provides a statistical mechanism for adjusting in-sample performance metrics to account for optimistic bias arising from model fitting and complexity [103]. By incorporating penalties for overfitting, this approach offers a more realistic estimate of a model's true predictive capability when applied to unseen warranty data, which is critical for practical deployment in warranty budgeting and accrual planning.

2.12.4 Cross-Validation for Time Series

Model validation for time-series warranty data requires specialised techniques that preserve temporal ordering while enabling reliable performance assessment under data constraints. Conventional k-fold cross-validation is unsuitable for time-series applications, as it violates the chronological structure of warranty data and introduces information leakage from future observations [104]. Consequently, validation strategies specifically designed for sequential data are required.

Walk-Forward Validation

Walk-forward validation is widely regarded as the benchmark methodology for evaluating time-series forecasting models. This approach systematically expands the training dataset over time while maintaining strict temporal integrity between training and testing phases [98]. The process begins with an initial training window, generates

forecasts for a defined horizon, and subsequently incorporates newly observed data into the training set before repeating the forecasting procedure.

For a time series of length T with an initial training window of size W and forecast horizon h, the validation process is defined as follows:

- **Training Set 1:** $t = 1,2, \dots, W$
- **Test Set 1:** $t = W + 1, W + 2, \dots, W + h$
- **Training Set 2:** $t = 1,2, \dots, W + h$
- **Test Set 2:** $t = W + h + 1, W + h + 2, \dots, W + 2h$

This sequential evaluation framework ensures that future information is never used in model training, thereby providing a realistic assessment of out-of-sample forecasting performance that closely mirrors real-world warranty forecasting deployment conditions.

2.13 Big Data Analysis for Warranty and Quality Improvement

Big data analysis enables the extraction of actionable insights from large-scale and complex datasets, playing a critical role in modern automotive quality and warranty management. The analysis of product quality warranty data using data mining techniques supports improvements in product performance and customer satisfaction by systematically identifying fault patterns and underlying root causes. These insights allow manufacturers to implement targeted design improvements, reduce defect occurrences, and enhance overall product quality and reliability.

The integration of analytical findings into structured continuous improvement frameworks further strengthens their practical impact. The Plan–Do–Check–Act (PDCA) cycle provides a systematic methodology for continuous quality enhancement by guiding organisations through iterative stages of planning objectives, implementing corrective actions, evaluating outcomes, and refining processes to sustain performance improvements [105]. Within warranty management, insights derived from big data analytics inform the planning stage, support evidence-based implementation, and enable objective performance evaluation, thereby reinforcing data-driven decision-making.

Overall, big data analysis serves as a foundational enabler for these improvement efforts by allowing organisations to process and analyse vast volumes of structured and unstructured data efficiently. The primary objective is to transform raw warranty and quality data into meaningful, results-oriented insights through the application of advanced statistical techniques and analytical methods.

2.14 Product Design Specification

A Product Design Specification (PDS) is a fundamental document in engineering product development, serving as the formal foundation for all subsequent design activities. It systematically defines the scope, requirements, constraints, and performance expectations of a product or system before detailed design solutions are generated. Rather than prescribing a specific solution, the PDS establishes a structured problem framework, ensuring that design outcomes align with both customer expectations and engineering feasibility [44].

A well-defined PDS typically encompasses key elements such as functional requirements, performance targets, cost limitations, safety considerations, environmental and regulatory requirements, physical constraints (including size and weight), aesthetic considerations, and manufacturing or production limitations. By clearly articulating these requirements, the PDS provides a common reference point for designers, engineers, and stakeholders, thereby minimising the risks of scope creep, excessive redesign, budget overruns, and customer dissatisfaction.

Importantly, the PDS is not a static document. As product development progresses and new technical, regulatory, or market-related information emerges, the specification may require periodic review and refinement. This iterative nature allows the PDS to remain relevant throughout the development lifecycle while maintaining alignment with evolving project objectives and constraints [106]. In this project, the design methodology follows a structured four-step engineering design process, as illustrated in Figure 2.10 This approach reflects established engineering design practices and provides a systematic pathway from problem definition to validated design outcomes. While Figure 2.11 presents a customized PDCA cycle designed specifically for warranty data evaluation.

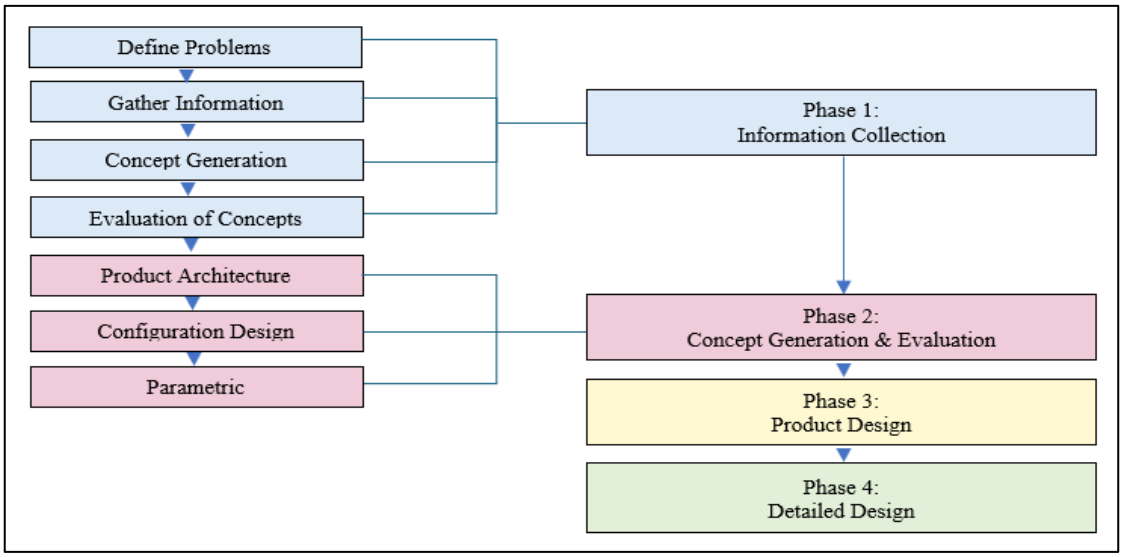


Figure 2.10 Framework Product Design Specification Approach

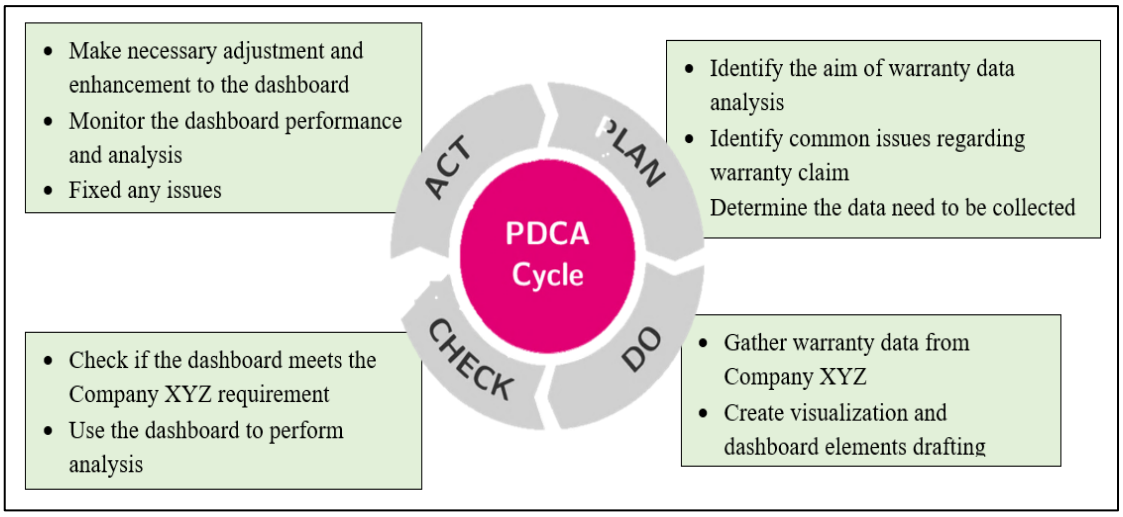


Figure 2.11 PDCA Cycle for Warranty Data Analysis

2.14.1 Data Exploration, Manipulation, And Cleaning Techniques

Effective data exploration, manipulation, and cleaning are essential prerequisites for ensuring the accuracy, reliability, and validity of analytical outcomes in warranty forecasting research. This study emphasises the importance of systematic data preparation processes to address data inconsistencies, missing values, and formatting issues commonly present in large-scale warranty datasets.

Versatile analytical tools such as Microsoft Excel and R are utilised to support these processes. Excel provides flexible and user-friendly capabilities for importing data from multiple sources, including CSV files and spreadsheet formats, making it suitable

for preliminary data handling and validation tasks. Its built-in functions facilitate efficient data cleaning operations, such as duplicate removal, data standardisation, and correction of formatting inconsistencies, thereby streamlining the preparation of warranty datasets for further analysis [107]. These preparatory steps establish a reliable foundation for subsequent advanced analytics and predictive modelling.

2.14.2 Advanced Analytical Techniques For Warranty Data Analysis (WDA)

Advanced analytical techniques for Warranty Data Analysis (WDA) play a pivotal role in improving product reliability and enhancing customer satisfaction in the automotive industry. A comprehensive review of these techniques highlights their importance in enabling the early identification of reliability issues, allowing manufacturers to implement corrective actions before failures escalate into costly repairs or extended downtime. By systematically analysing warranty data, organisations can gain valuable insights that support timely interventions and proactive quality management.

A key focus of WDA lies in leveraging warranty information to inform design improvements. Insights derived from warranty claims enable engineers to identify recurring failure modes and incorporate targeted design modifications, thereby improving system robustness, product durability, and overall quality. These enhancements contribute to improved field reliability and enable more accurate estimation of warranty liabilities, which is particularly critical for products with long warranty coverage periods.

Recent research has also emphasised the potential of online monitoring and real-time analytical techniques, such as statistical control charts, for continuous performance tracking and rapid detection of emerging trends or anomalies. Such approaches support timely decision-making and facilitate early corrective actions in dynamic operating environments. However, the increasing complexity of modern products, coupled with the growing adoption of wireless and connected technologies, presents significant challenges in processing and analysing large-scale, real-time warranty data. Addressing these challenges requires more advanced data handling and analytical strategies to ensure effective extraction of actionable insights and sustained improvement in product performance [23] [108].

2.14.3 Enhance Decision-Making with Business Intelligence (BI)

Business Intelligence (BI) tools play a key role in supporting data-driven decision-making by integrating and visualising data from multiple sources. Microsoft Power BI provides a comprehensive BI platform that consolidates heterogeneous data into interactive and visually intuitive dashboards, enabling effective monitoring of performance metrics and trends [109]. In this research, Power BI is used to integrate data from sources such as Excel and transform them into dynamic visualisations that support timely and informed managerial decisions.

Data visualisation enhances analytical understanding by presenting complex information in simplified graphical formats, allowing stakeholders to rapidly identify patterns, relationships, and anomalies that may be obscured in raw datasets. This visual approach improves communication, facilitates problem identification, and supports performance improvement initiatives. Overall, BI tools such as Microsoft Power BI convert raw data into actionable insights, contributing to improved organisational efficiency and competitive advantage [110].

2.14.4 Data Preparation and Extraction, Transformation and Loading Process

The Extract, Transform, and Load (ETL) process and data preparation phase constituted a critical foundation for the overall analytical framework developed in this study. A total of more than 71,000 raw warranty claim records were extracted from the company's internal warranty database. These records were subsequently imported into Microsoft Excel, where systematic data cleaning and standardisation procedures were performed to ensure data consistency, reliability, and analytical suitability.

Data cleaning involved the removal of duplicate records, correction of inconsistent date and currency formats, and appropriate handling of missing values where feasible. Excel functions such as *Trim*, *Clean*, and *Replace* were employed to standardise text fields and eliminate formatting irregularities. Microsoft Excel was selected for this stage due to its flexibility in handling structured datasets, ease of manual verification, and capability for rapid data manipulation. The adopted procedures followed recommended data preparation practices outlined by Microsoft for data initialisation and quality assurance in Excel environments [111].

Following data cleaning, metadata attributes were defined for each key field, including claim date, vehicle model, failed component, failure description, and total accessories cost. The establishment of metadata enhanced data interpretability and supported structured analysis by clarifying the role and meaning of each variable within the dataset [112]. Once organised, the refined dataset was imported into Microsoft Power BI using its built-in data connector. Relationships between tables were established using key identifiers such as model code, part identification, and claim date to enable efficient filtering, aggregation, and reporting during subsequent analysis.

2.14.5 Dashboard Development and Visualization with Power BI

Microsoft Power BI was utilised to develop dynamic and interactive dashboards for visualising warranty data. The dashboards were designed to present key performance indicators (KPIs), including total warranty claims, average claim cost, and failure frequency by component and vehicle model. Various visual elements, such as bar charts, pie charts, and time-series plots, were employed to support intuitive interpretation of warranty trends and patterns [113].

Interactive filtering capabilities were incorporated to allow users to analyse warranty claims by vehicle model, failure cause, geographic region, and claim period. These interactive features enhanced usability and enabled stakeholders from different functional departments to explore the data according to their specific informational needs. Microsoft Power BI was selected for this application due to its strong data integration capabilities, real-time update functionality, and user-friendly interface, making it accessible even to non-technical users [110].

2.14.6 Application of Plan Do Check Act Cycle in Data Analytics

The overall analytical methodology was structured using the Plan–Do–Check–Act (PDCA) cycle to support continuous improvement as shown in Figure 2.11. During the *Plan* phase, the objectives and system requirements for the warranty analysis platform were defined. The *Do* phase involved the implementation of the data processing pipeline, ETL procedures, and dashboard development.

The *Check* phase focused on evaluating dashboard performance and collecting user feedback related to visual clarity, functionality, and analytical effectiveness. In the

Act phase, iterative improvements were introduced based on evaluation outcomes, including refinements to dashboard layout, enhancement of filtering options, and optimisation of navigation. The application of the PDCA cycle within this digital analytics context aligns with continuous quality improvement principles and is consistent with Industry 4.0 practices [114].

2.14.7 Time-Based Analysis and Predictive Failure Modelling

Time-based trend analysis was conducted using a warranty matrix approach to examine changes in warranty claim behaviour over time. This analysis enabled the mapping of claim frequencies across different time intervals following product delivery, thereby highlighting variations in product reliability throughout the lifecycle. Such temporal analysis facilitated the identification of potential design weaknesses or manufacturing-related issues that manifest at specific stages of product usage.

In addition, predictive modelling techniques were applied to historical warranty claim data to support forward-looking failure estimation [113]. Forecasting methods, including time-series analysis and logistic regression, were employed to predict future failure occurrences and identify components with elevated risk levels. These predictive insights enable engineers to implement preventive actions in advance, contributing to reduced warranty costs and improved product reliability [115].

2.14.8 Warranty Cost Analysis and Financial Trends Evaluation

A comprehensive warranty cost analysis was conducted to assess the financial implications of warranty claims. This analysis included the calculation of total warranty cost by vehicle model, average cost per failure event, and identification of components contributing the highest cumulative costs [113]. The Alternating Geometric Process was applied to estimate expected cost escalation over the product lifecycle, providing insight into long-term warranty expenditure trends.

To support integrated decision-making, financial analysis results were incorporated into the Power BI dashboards alongside technical warranty indicators. This integration enabled managers and engineers to evaluate the economic impact of design or quality issues and develop informed strategies for cost control and warranty management [112], [116].

2.14.9 Integration of Quality 4.0 and Business Intelligence

Quality 4.0 principles were integrated into this study by enhancing traditional quality management practices with advanced business intelligence capabilities. The deployment of Microsoft Power BI enabled near real-time monitoring of warranty performance, allowing engineers and decision-makers to detect anomalies promptly and respond proactively [113].

The dashboard-based approach also facilitated cross-functional collaboration by translating complex technical warranty data into accessible visual formats. This integration supported informed decision-making, continuous process improvement, and product enhancement initiatives. Ultimately, the adoption of BI-driven analytics contributed to a cultural shift towards a data-driven organisation focused on continuous improvement, operational excellence, and customer satisfaction [117], [118].

2.15 Literature Review Summary

The literature review demonstrates that warranty management within the automotive industry has progressively evolved from conventional reactive approaches toward predictive, data driven, and technology enabled systems. As summarised in Table 2.2, prior studies have emphasised Malaysian automotive market growth, quality management methodologies, and analytical techniques relevant to warranty analysis. The Malaysian Automotive Association (MAA) highlighted increasing vehicle sales and market expansion, while DFMEA and CAM methodologies were recognised as important approaches for defect prevention and quality improvement. In addition, Exploratory Data Analysis (EDA) has been widely applied to identify warranty claim patterns, trends, and relationships within industrial datasets.

The analytical and predictive dimensions of warranty management are further consolidated in Table 2.3, which outlines key indicators and validation approaches adopted in previous studies. Component repair strategies have been reported as effective mechanisms for reducing warranty claims and replacement cost, supporting the transition from complete part replacement toward targeted repair activities. Predictive modelling studies have also demonstrated the use of regression-based approaches with performance indicators such as $R^2 > 0.6$ for model accuracy evaluation and F1 score > 0.7 for predictive validation. These indicators establish practical benchmarks for

assessing forecasting capability and analytical performance within warranty management applications.

Supporting technologies and measurement approaches identified in Table 2.4 further strengthen the development of modern warranty systems. Reliability assessment through Cronbach's Alpha (>0.7) has been used to validate measurement consistency, while technologies such as IoT, blockchain, and real time monitoring have been explored to enhance warranty part recovery, traceability, and operational visibility. Competitive benchmarking practices have also been recognised as important tools for evaluating warranty performance against industry peers.

Despite these developments, a research gap remains in integrating these analytical methods, validation measures, and supporting technologies into a unified warranty management framework tailored to the Malaysian automotive environment. Limited studies have combined market analysis, predictive modelling, reliability validation, component repair strategy, and business intelligence-based visualisation within a single system. Therefore, this research addresses the gap by developing a data driven warranty management model that integrates EDA, predictive forecasting, model validation metrics, reliability assessment, and interactive dashboard analytics to strengthen warranty decision making, cost optimisation, and operational monitoring in the Malaysian automotive industry.

Table 2.2
Domain Parameter

Source	Year	Study Focus	Indicator	Min Value
Malaysian Automotive Association (MAA)	2023	Malaysian automotive market growth	Vehicle sales	N/A
Wu, Shaomin, Muhammad Jawad Akhtar, Afshan Naseem, Faaiz Ahsan	2012, 2024	Quality management and defect prevention methodologies	DFMEA and CAM	N/A
F. Arena et al.	2022	Identifying patterns and trends in warranty data	EDA	≥ 0.3 (Pearson's)

Table 2.3
Correspondence Index

Source	Year	Study Focus	Indicator	Validation Test
Salizanah, Muhamad Lutfi,	2023	Warranty claim management and part replacement strategies	Component repair	5% cost reduction in warranty claims
J. Liu, G. Liu, Y. Zhang	2023	Predicting warranty claim occurrences and costs	Regression models	$R^2 > 0.6$ (model accuracy)
V. Krivtsov	2023			
S. Wu, A. Akbarov	2011			
F. Arena et al.	2022	Model performance in predicting warranty claims	F1-score	> 0.7 (validation cutoff)
Dnyanesh Rajpathak, Yiming Xu, Ian Gibbs	2020			

Table 2.4
Measurement Model

Source	Year	Study Focus	Indicator	Validation Test
C. Lee Kwee Yoon	2023	Reliability testing of questionnaire/feedback data	Cronbach's Alpha	> 0.7 (acceptable reliability)
Z. M. Mohd Hanafi Azman Ong	2019			
T. M. F.-C. Paula Fraga-Lamas	2019	Real-time monitoring and warranty part recovery processes	IoT and blockchain	N/A
A. Liao	2022			
Hairi, Mohd Hamsani Mohd	2023	Competitor benchmarking in warranty claim management	Benchmarking against competitors	N/A

2.16 Research Gap

The existing literature highlights substantial advancements in the integration of big data analytics, predictive maintenance models, and quality management tools within the automotive industry. Techniques such as DFMEA, AHP, DEMATEL, and

regression models have been effectively utilized in various contexts to improve warranty management practices, enhance product reliability, and optimize operational efficiency. However, a significant gap remains in the application of these methodologies specifically within the Malaysian automotive market.

While general approaches have been explored, there is a distinct lack of research that addresses the unique challenges and dynamics of the Malaysian automotive industry, particularly in terms of localized supplier performance, customer behaviour, and the specific regulatory and market conditions in Malaysia. The integration of these advanced methodologies into a predictive warranty management system that is tailored to the needs of Malaysian automotive manufacturers is underexplored.

Furthermore, the existing studies often focus on isolated aspects of warranty management, such as predictive analytics or quality control, without a holistic approach that combines various methodologies to create a comprehensive predictive system. Specifically, there is a need for research that integrates Exploratory Data Analysis (EDA) with Regression Models and evaluates the system's performance using metrics like the F1-Score and Cronbach's Alpha to ensure both accuracy and reliability.

This gap suggests that current models may not fully capture the complexities of the Malaysian automotive market, leading to suboptimal warranty management strategies. Addressing this gap through targeted research could result in the development of a more robust and effective predictive warranty management system that not only reduces warranty costs but also improves customer satisfaction and enhances the overall competitiveness of Malaysian automotive manufacturers on a global scale. This research would provide valuable insights into how advanced data-driven strategies can be localized and applied to improve warranty management practices in a specific regional context.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

According to Chapter Two, the analysis and visualization of warranty data are crucial for mitigating warranty costs. The analysis begins with identifying the primary cost factors, which are consistently highlighted in the annual report of Company MNC. Company MNC conducts this analysis based on diverse data sourced from OEM customers. Successful countermeasures will be implemented in other car assemblers as YOKOTEN activities to prevent the recurrence of similar issues in the market. All feasibility study data and monitoring results will be shared after the project's initiation and conducted by the internal organization. The entire process and Standard Operating Procedures related to this project are depicted in the flowchart shown in Figure 3.1.

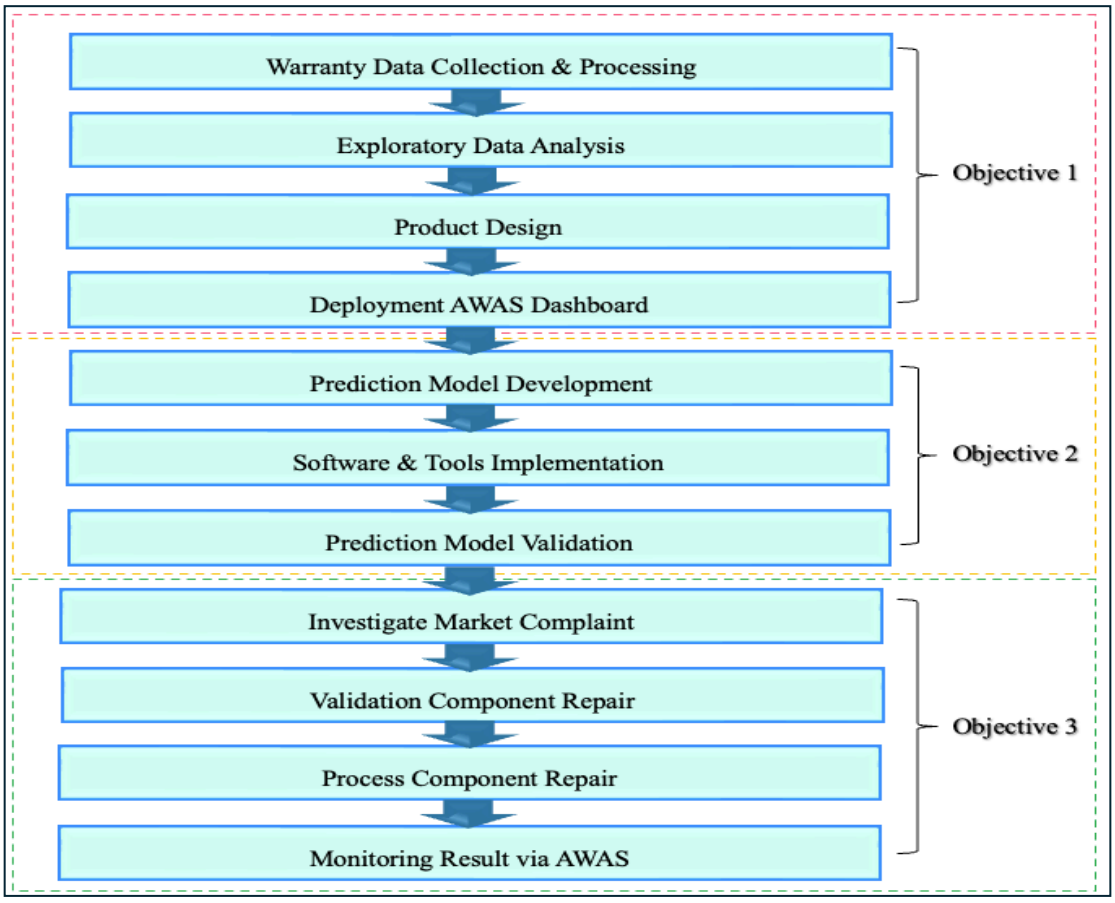


Figure 3.1 Flowchart of Warranty Management System Model System

3.2 Product Design Specification

This research employs a structured Product Design Specification (PDS)–based methodology comprising four sequential phases. Phase 1 (Information Collection) focuses on defining the problem through systematic gathering and analysis of warranty data, field reports, and operational feedback to establish design requirements and constraints. Phase 2 (Concept Generation and Evaluation) involves the development and assessment of multiple solution concepts using technical feasibility, cost impact, and reliability criteria to ensure alignment with the PDS. Phase 3 (Product Design) translates the selected concept into a structured system architecture, including configuration and parametric design, enabling effective analysis and visualisation of warranty behaviour. Finally, Phase 4 (Detailed Design) refines the proposed solution through detailed implementation and validation, ensuring the final system meets functional, analytical, and operational objectives. This phased approach ensures traceable, data-driven decision-making and supports robust warranty cost reduction and product quality improvement.

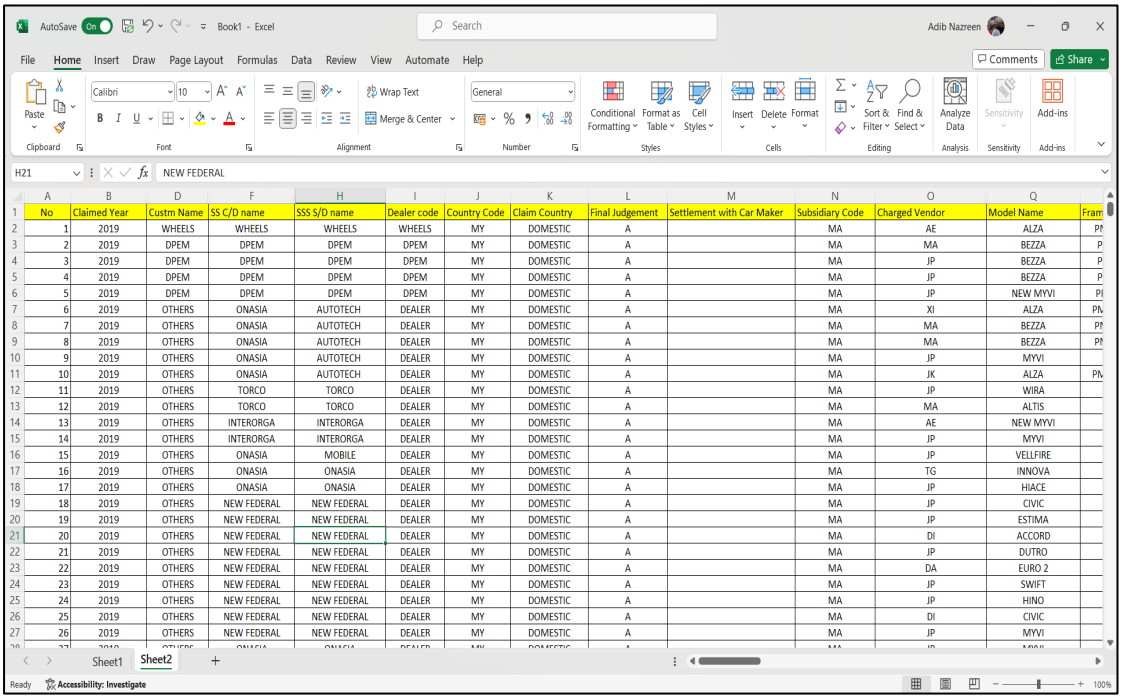
3.2.1 Phase 1: Information Collection

The primary challenge faced by automotive manufacturers in warranty management lies in the existence of large volumes of unstructured and fragmented data, with no single integrated platform to monitor recurring failures or cost-intensive warranty patterns. This lack of consolidated visibility hampers effective identification of common component failures, inefficient repair practices, and geographically suboptimal service operations. Addressing these issues forms the central motivation of this project, which focuses on transforming historical warranty data into actionable insights for decision-making and predictive analysis.

To establish a clear problem definition, raw warranty claim data were initially explored using Microsoft Excel, as illustrated in Figure 3.2. This preliminary investigation aimed to identify high-frequency claims, high-cost components, and recurring failure modes. Key data fields including claim year, vehicle model, failed component, cause description, mileage, and total claim cost were analysed to reveal early patterns in failure distribution and cost concentration. Descriptive analytics techniques such as frequency analysis, cost comparison, and basic trend visualization

were employed to highlight components or product lines with disproportionately high warranty exposure. Literature supports this approach, noting that warranty behaviour analysis is an effective means of localising system vulnerabilities and enabling predictive maintenance planning [119], [120].

This phase also incorporated business best practices that emphasise a closed-loop warranty management approach, whereby historical warranty data are continuously analysed to reduce failure rates and improve product design reliability. Supplementary inputs including customer complaints, service centre feedback, and historical recall data were reviewed to ensure that the analytical scope captured both technical and operational perspectives [121]. The absence of structured failure pattern definitions and the lack of effective visualization tools were identified as core root causes, thereby justifying the need for an integrated, data-driven warranty analysis system supported by Business Intelligence (BI) tools.



No	Claimed Year	Custm Name	SS C/D name	SSS S/D name	Dealer code	Country Code	Claim Country	Final Judgement	Settlement with Car Maker	Subsidiary Code	Charged Vendor	Model Name
1	2019	WHEELS	WHEELS	WHEELS	WHEELS	MY	DOMESTIC	A		MA	AE	ALZA
2	2019	DPEM	DPEM	DPEM	DPEM	MY	DOMESTIC	A		MA	MA	BEZZA
3	2019	DPEM	DPEM	DPEM	DPEM	MY	DOMESTIC	A		MA	JP	BEZZA
4	2019	DPEM	DPEM	DPEM	DPEM	MY	DOMESTIC	A		MA	JP	BEZZA
5	2019	DPEM	DPEM	DPEM	DPEM	MY	DOMESTIC	A		MA	JP	NEW MYVI
6	2019	OTHERS	ONASIA	AUTOTECH	DEALER	MY	DOMESTIC	A		MA	XI	ALZA
7	2019	OTHERS	ONASIA	AUTOTECH	DEALER	MY	DOMESTIC	A		MA	MA	BEZZA
8	2019	OTHERS	ONASIA	AUTOTECH	DEALER	MY	DOMESTIC	A		MA	MA	BEZZA
9	2019	OTHERS	ONASIA	AUTOTECH	DEALER	MY	DOMESTIC	A		MA	JP	MYVI
10	2019	OTHERS	ONASIA	AUTOTECH	DEALER	MY	DOMESTIC	A		MA	JK	ALZA
11	2019	OTHERS	TORCO	TORCO	DEALER	MY	DOMESTIC	A		MA	JP	WIRA
12	2019	OTHERS	TORCO	TORCO	DEALER	MY	DOMESTIC	A		MA	MA	ALTIS
13	2019	OTHERS	INTERORGA	INTERORGA	DEALER	MY	DOMESTIC	A		MA	AE	NEW MYVI
14	2019	OTHERS	INTERORGA	INTERORGA	DEALER	MY	DOMESTIC	A		MA	JP	MYVI
15	2019	OTHERS	ONASIA	MOBILE	DEALER	MY	DOMESTIC	A		MA	JP	VELLFIRE
16	2019	OTHERS	ONASIA	ONASIA	DEALER	MY	DOMESTIC	A		MA	TG	INNOVA
17	2019	OTHERS	ONASIA	ONASIA	DEALER	MY	DOMESTIC	A		MA	JP	HIACE
18	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	CIVIC
19	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	ESTIMA
20	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	DI	ACCORD
21	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	DUTRO
22	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	DA	EURO 2
23	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	SWIFT
24	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	HINO
25	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	DI	CIVIC
26	2019	OTHERS	NEW FEDERAL	NEW FEDERAL	DEALER	MY	DOMESTIC	A		MA	JP	MYVI

Figure 3.2 Warranty Claim Raw Data.

Preliminary data filtering and structuring were conducted using Microsoft Excel, while Microsoft Power BI was employed for data integration, relational mapping, and interactive filtering. Additional contextual data such as product specifications, production locations, and customer service reports were examined to improve traceability of claim origins and to identify technical vulnerabilities [122]. The

integration of warranty claims with supporting metadata is widely recognised in the literature as a critical enabler for effective root cause analysis and failure prediction, encompassing hardware-related faults, software issues, human errors, and organisational process failures [123], [124]. Establishing a centralised data environment with embedded warranty logic is therefore considered the foundational step in developing an optimal warranty analysis system [125] [126].

Consistent with the requirements defined in the Product Design Specification (PDS) and established engineering design methodologies, Phase 1 focused on a structured approach to problem formulation and information acquisition. This stage encompassed systematic identification of the core problem, comprehensive data collection, exploratory idea generation, and preliminary evaluation of potential conceptual solutions. Warranty claim data obtained from Company MNC were examined within a Plan–Do–Check–Act (PDCA) framework to facilitate iterative analysis and support continuous improvement practices [127]. Particular emphasis was placed on organising and standardising the data environment to ensure analytical robustness and the reliability of subsequent insights. The overall data acquisition and preparation process is illustrated in Figure 3.3, while the adapted PDCA cycle demonstrates the integration of continuous improvement principles into warranty data assessment.

To assess the practicality and robustness of the proposed warranty data visualisation system, a series of structured validation activities was undertaken. Functional validation was conducted using actual warranty datasets to verify the correctness of visual representations, filtering operations, and overall system performance. In parallel, usability assessment was carried out through expert evaluation by quality engineers from Company MNC, with attention given to interface readability, clarity of information presentation, and effectiveness in supporting decision-making. System reliability was further examined through repeated data loading and analytical execution to confirm output consistency and identify potential irregularities. Feedback obtained from these validation activities guided iterative system enhancements and confirmed the suitability of the developed solution for operational warranty analysis applications.

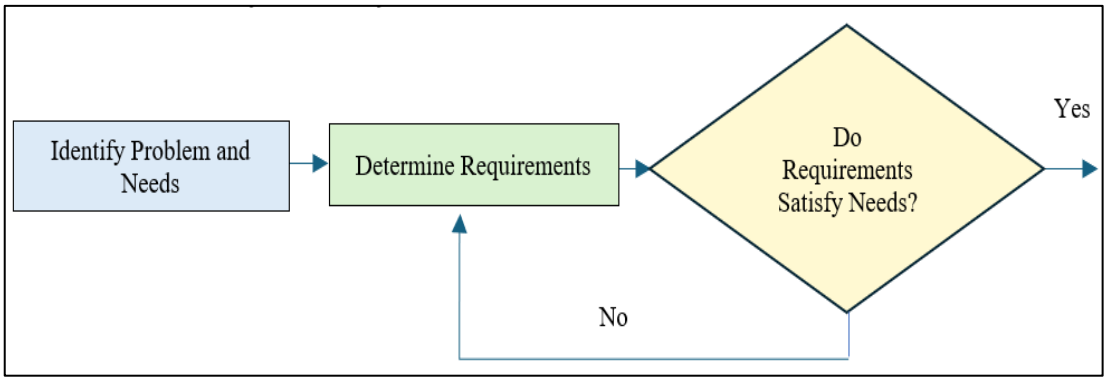


Figure 3.3 Data Collection Process

Following problem framing and data consolidation, conceptual design progressed towards the evaluation of suitable data visualization platforms. A comparative software assessment, summarised in Table 3.1, examined Microsoft Power BI against alternative solutions such as Tableau and Google Data Studio. Selection criteria included ease of use, dashboard interactivity, data model flexibility, scalability, and compatibility with existing Excel-based and SQL data environments at Company MNC. Microsoft Power BI demonstrated superior integration capabilities and operational alignment with the organisation's existing infrastructure [128].

Prototype dashboards were subsequently developed using sample warranty data exported directly from the AWAS system. These prototypes tested the platform's ability to visualise key performance indicators (KPIs), identify trends by vehicle model, region, and component type, and support responsive interaction across desktop and mobile platforms. Evaluation criteria included dashboard load time, data refresh reliability, clarity of KPI representation, and user navigation efficiency. Power BI's advanced capabilities such as Power Query for data transformation, DAX for complex relationship modelling, and Row-Level Security (RLS) for controlled data access further reinforced its suitability. Prior studies similarly highlight the effectiveness of such BI tools in improving warranty management performance by enabling early fault detection and targeted corrective actions [129].

Overall, this phase establishes a robust foundation for the subsequent system development stages. By consolidating fragmented warranty data, validating data integrity, and selecting a proven BI platform, the project ensures that the development phase is built upon a reliable, scalable, and industry-aligned analytical infrastructure.

Table 3.1
Decision Software

Criteria	Weightage	Current solution	Looker studio	Power BI	Microsoft Excel
Compatibility with large data	3	0	2	3	3
User friendly interface	3	0	2	3	1
Data modelling flexibility	2	0	1	2	1
Dashboard interactivity	1	0	1	3	0
Total			6	11	5

3.2.2 Phase 2: Concept Generation and Evaluation

Phase 2 focuses on concept generation and evaluation, corresponding to the embodiment design stage in engineering design methodology. At this stage, the emphasis is placed on defining the overall system architecture, data flow structure, and analytical framework to ensure that the proposed solution is technically feasible, scalable, and aligned with the problem definition established in Phase 1. This phase is critical for validating that the selected concept is not only theoretically sound but also practically implementable within the operational constraints of Company MNC. Following the identification of key challenges such as inefficient processing, fragmented storage, and limited visualization of large-scale warranty data, this phase aimed to generate and evaluate multiple solution concepts for designing a warranty data analysis and visualization system. The objective was to identify analytical and architectural approaches capable of transforming raw warranty data into structured, interpretable, and decision-oriented insights.

Several alternative concepts were explored during this stage. These included a fully manual Excel based analysis approach, a hybrid approach combining Excel with Business Intelligence tools, and a fully Business Intelligence driven solution. Each concept was evaluated against predefined technical and operational criteria including ease of use, capability to handle large datasets, compatibility with existing Excel based data formats, dashboard interactivity, scalability, and accessibility for users with

varying levels of technical expertise. These criteria ensured that both engineering robustness and end user usability were considered during concept selection.

Among the evaluated options, Microsoft Power BI emerged as the most suitable platform for warranty data visualization and analysis. This selection was driven by its intuitive user interface, advanced data visualization capabilities, and native support for relational data modelling across multiple tables. Power BI also enables dynamic dashboard updates, real time filtering, and drill down analysis, which are essential for monitoring warranty trends and identifying recurring failure patterns. In addition, its support for role-based access control and secure report sharing is particularly important in managing confidential warranty data across multiple organisational levels.

To systematically explore possible system configurations, a morphological matrix as shown in Table 3.2 was employed to generate combinations of system features including data preparation tools, analytical methods, visualization platforms, and security mechanisms. This structured approach facilitated the development of a hybrid solution in which Microsoft Excel was used for initial data preparation and cleansing, while Power BI served as the primary platform for analysis, visualization, and reporting. The hybrid configuration leveraged the strengths of both tools, ensuring flexibility in data handling while maintaining analytical depth and scalability.

Subsequently, a weighted decision matrix was applied to quantitatively compare the proposed concepts. Evaluation criteria included data accuracy, processing speed, visualization flexibility, system scalability, implementation cost, and ease of maintenance. Each criterion was assigned a relative weight based on its importance to the project objectives, and each concept was scored accordingly. The results of this evaluation demonstrated that the Excel and Power BI hybrid solution achieved the highest overall score, reflecting its ability to meet both technical requirements and user expectations.

The structured concept generation and evaluation process ensured that the selected solution complied with engineering best practices and user centred design principles. Prior studies emphasise that formalised concept selection methods are essential in data driven system development to avoid misalignment with operational needs and to support scalable and maintainable system architectures. Accordingly, the outcome of this phase established a validated system concept that serves as a robust foundation for subsequent detailed design and implementation stages [130].

Table 3.2
Morphological Chart Comparison of Software

Feature Category	Excel	Power BI	Looker Studio
Data Handling Capacity	Limited to medium datasets	Handles large and complex datasets	Medium datasets (depends on source)
Ease of Use	High (widely familiar)	Moderate (requires learning visuals, DAX)	Moderate (needs integration knowledge)
Visualization Capability	Basic charts, pivot tables	Advanced interactive dashboards	Good dashboarding, limited customization
Real-time Data Update	Manual refresh	Supports automatic scheduled refresh	Supports real-time from Google sources
Data Source Integration	Manual import (CSV, Excel only)	Broad range: Excel, SQL, cloud, APIs	Strong for Google ecosystem, API needed
Security / Access Control	File-level password protection	Role-based security (Row-Level Security)	Limited, mostly view-only permissions
Interactivity	Low (basic filters/slicers)	High (slicers, drill-throughs, buttons)	Moderate (basic filters)
Cost / Licensing	Free with MS Office	Free & Pro versions (license for full use)	Free (within Google account)
Collaboration	File sharing via email or cloud	Real-time sharing, cloud workspace	Web-based sharing (limited user control)
Scalability	Not scalable for big data	Highly scalable for enterprise use	Limited scalability

In this phase, attention was directed towards the conceptual development of a warranty data analysis system that emphasises usability, analytical efficiency, and decision-support capability. The primary objective was to design a system that simplifies warranty claim evaluation while remaining accessible to users with diverse technical backgrounds. To achieve this, the proposed system architecture incorporates predefined configurations and an intuitive user interface, thereby minimising reliance on advanced technical expertise and reducing operational complexity [127].

To enhance the accuracy and consistency of claim assessment, a two-dimensional warranty evaluation framework was introduced. This framework integrates key operational factors, namely product age and usage intensity, enabling a more structured interpretation of warranty behaviour and improving the reliability of claim classification [131]. An overview of the proposed conceptual interface is illustrated in

Figure 3.4, highlighting core system components including data input mechanisms, analytical processing modules, and result visualisation features.

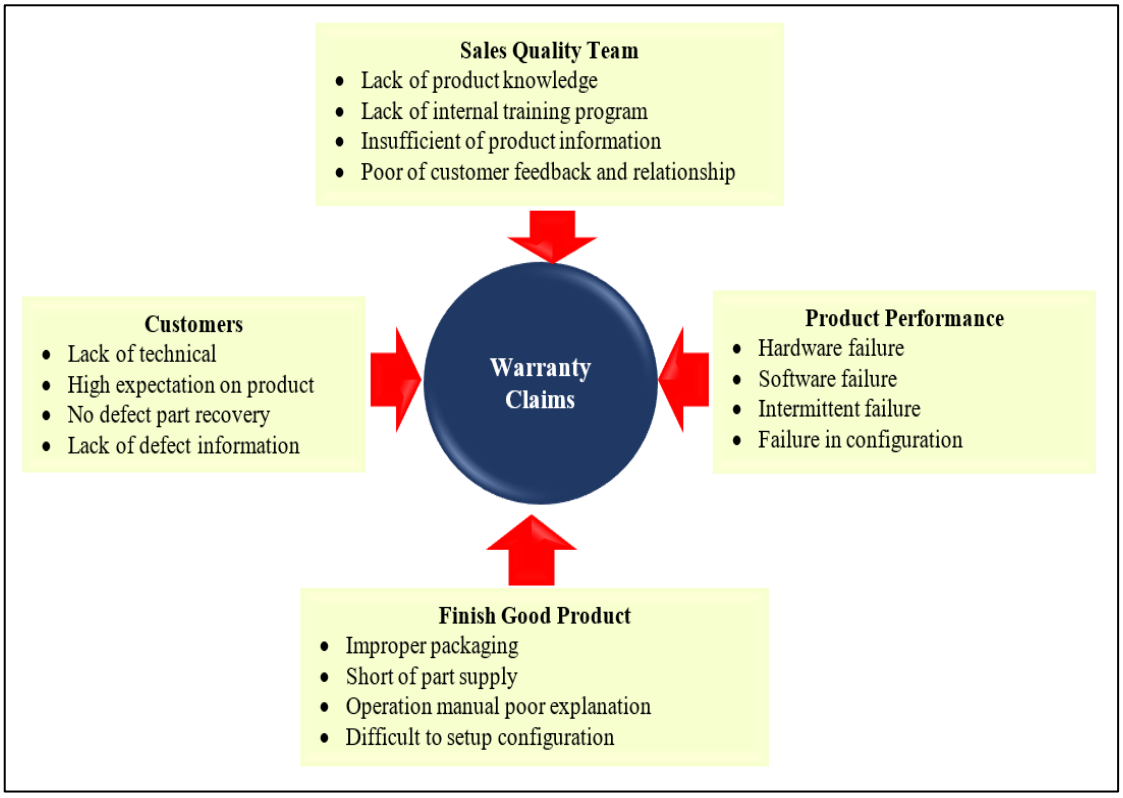


Figure 3.4 Conceptual Interface Design

To identify the most appropriate analytical platform for warranty data processing, a comparative evaluation matrix was developed, as presented in Table 3.1. The assessment framework considered multiple selection criteria, including resource availability, ease of use, visual design capability, and computational performance. Each criterion was assigned a weighted score based on its relative importance to system functionality and long-term operational efficiency.

a) Availability of Resources:

This criterion evaluated the accessibility of technical and organisational resources required for tool implementation and maintenance. The proposed system and Power BI both achieved high scores of 3, reflecting strong resource availability and institutional support. Looker Studio received a moderate score of 2, while Microsoft

Excel was assigned a score of 1, indicating limited scalability for advanced analytical deployment.

b) User-Friendliness:

Ease of interaction was a critical consideration to support adoption among non-technical users. Power BI achieved the highest usability score of 3 due to its intuitive navigation and interactive features. The proposed system and Looker Studio recorded moderate usability scores of 2, while Microsoft Excel demonstrated similar limitations.

c) Design and Graphical Interface:

This criterion assessed the effectiveness of visual representation and dashboard design. The proposed system achieved a score of 2, reflecting a structured but moderately constrained visual capability. Power BI, Looker Studio, and Microsoft Excel each scored 1, indicating relatively basic visual presentation when evaluated against customised analytical requirements.

d) Performance:

System performance was evaluated based on the ability to manage, process, and visualise large warranty datasets efficiently. Both the proposed system and Power BI achieved the highest score of 3, demonstrating strong data handling capabilities. In contrast, Looker Studio and Microsoft Excel were assigned scores of 1 due to performance limitations under large data volumes.

Based on the aggregated scores, Power BI achieved the highest overall rating with a total score of 10, followed by Looker Studio with 8 points and Microsoft Excel with 6 points. These results provide a structured justification for the selection of Power BI as the primary analytical platform for warranty data analysis. The subsequent section presents a detailed evaluation of the selected tool's technical capabilities, focusing on scalability, usability, and performance alignment with system requirements.

3.2.3 Phase 3: Product Design

Phase 3 focuses on the detailed product design of the Automotive Warranty Analysis System, translating the selected concept into a structured and implementable system architecture. At this stage, the embodiment design defines how the data pipeline, analytical models, and visualization layers are organised and interconnected to support effective warranty analysis and decision making [131]. The system is systematically decomposed into four main layers, namely data acquisition, data transformation, data modelling, and visualization, following established principles of modular analytics system design [132].

The product design begins with the data acquisition layer, where raw warranty claim data are imported from Microsoft Excel spreadsheets and supporting databases. These datasets contain essential attributes such as claim cost, failed component, vehicle mileage, vehicle model, and claim date. The imported data are processed in the Power BI data transformation layer using Power Query, where data cleansing and standardisation are performed. This process includes the removal of duplicate records, handling of missing values, normalisation of cost and mileage fields, and conversion of date formats into a consistent structure. These procedures ensure that the dataset is reliable and suitable for downstream analysis [133]. Figure 3.5 illustrates the Power Query Editor environment used during this data preparation stage.

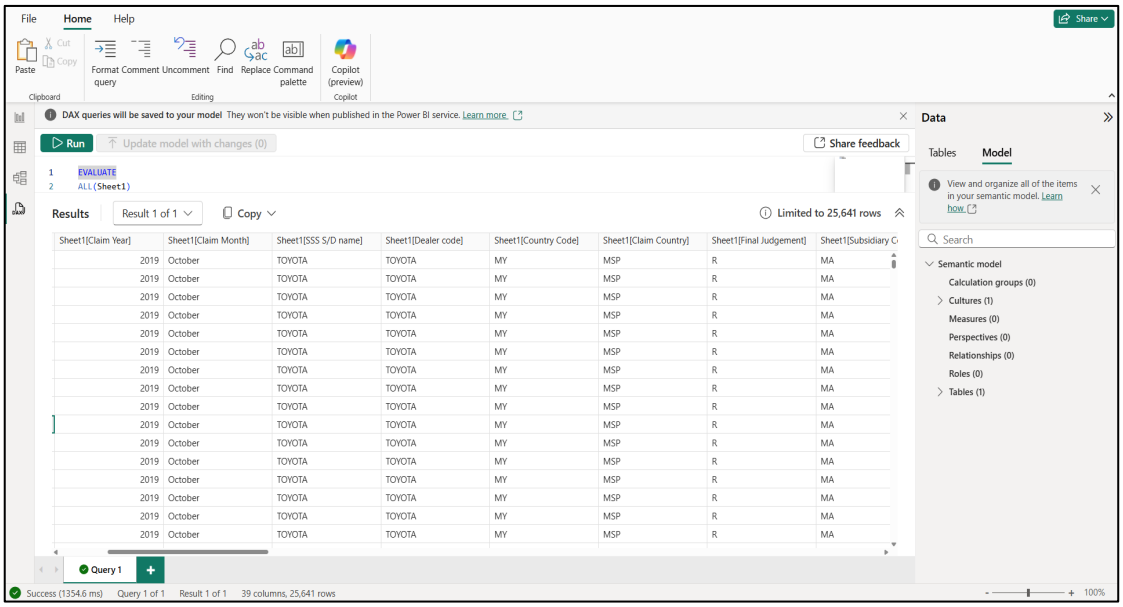


Figure 3.5 Power Query Editor

Following data transformation, the data modelling layer defines structured relationships between tables using key identifiers such as Vehicle Identification Number, Dealer Code, Part Number, and Model Name. One to many and many to one relationships are established within Power BI to support efficient aggregation, filtering, and slicing of data across multiple analytical views. This relational data modelling approach is essential for enabling flexible querying and multi-dimensional analysis across different report pages and is consistent with best practices in warranty analytics system development [62].

A core component of the system is the Key Performance Indicator engine developed using Data Analysis Expressions. This engine computes critical metrics including total number of claims, average claim cost, claim frequency per 71000 vehicles, and regional claim distribution. These indicators form the analytical backbone of the system and provide quantitative insights that support managerial decision making and quality improvement initiatives [129]. The final architectural layer is the visualization and reporting module, where interactive dashboards are constructed. Visual elements include bar charts, line charts, pie charts, scatter plots, and slicers that allow users to explore warranty data by vehicle model, fiscal year, dealer, region, and failure category. The modular design ensures responsiveness across desktop and mobile platforms while maintaining scalability and adaptability [132].

This multi layered architectural approach allows individual system components to be reused, modified, or extended without affecting the overall system integrity. Previous studies have shown that modular analytics architectures improve maintainability, enhance real time decision support, and strengthen user interaction in industrial data systems [132], [133].

The configuration design phase further defines how the system components interact from data import to final visualization. In this project, Microsoft Excel is used as the initial preprocessing environment, while Microsoft Power BI serves as the central platform for data modelling, analysis, and visualization. Warranty claim data are first structured into standardised Excel tables containing fields such as claim date, claim year, model name, dealer code, part name, failure description, mileage, and total claim amount. These tables form the foundational datasets that are subsequently imported into Power BI [134].

Within Power BI, a relational data model is established by linking tables such as claim details, dealer information, and part specifications through primary and foreign

keys. This structure enables dynamic filtering and cross dimensional analysis. Visualization types are selected based on analytical objectives. Bar charts are used to compare total claims across vehicle models or regions, pie charts illustrate the distribution of failure causes, line charts track claim trends over time, and scatter plots are used to examine relationships between mileage and failure occurrence. Interactive features such as slicers, drop down filters, and drill through functions allow users to navigate the data across different time periods, models, and failure categories. Automated refresh schedules are also configured to ensure that dashboards are updated whenever new data are added, enabling near real time monitoring of warranty trends [133].

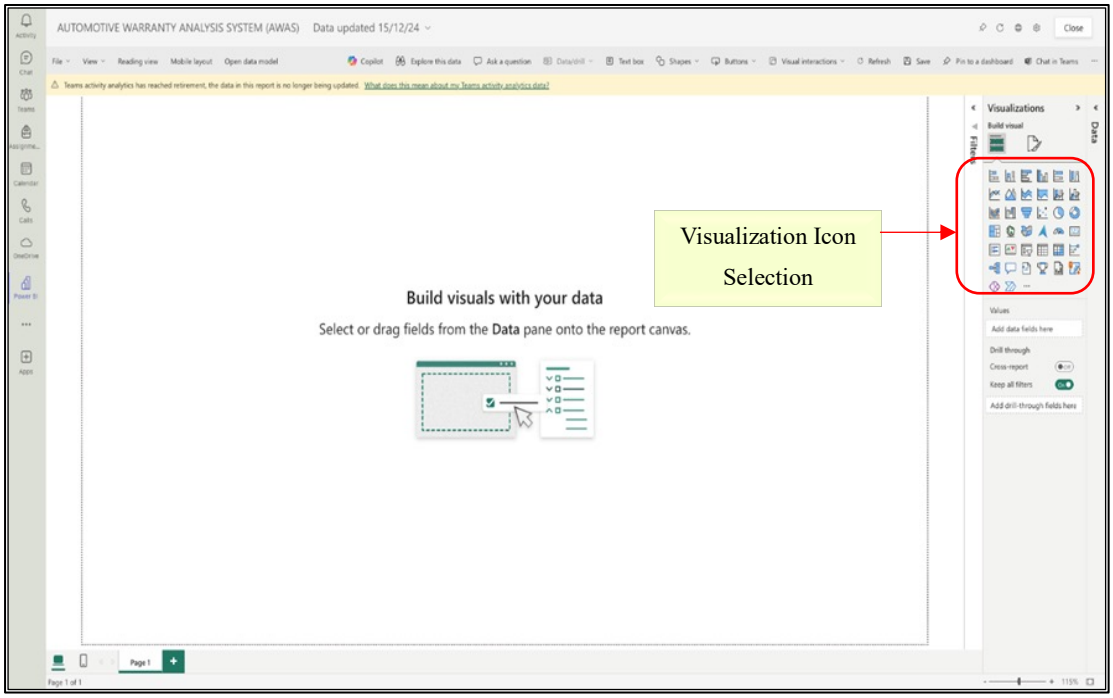


Figure 3.6 Visualizations in Microsoft Power BI to Improve Interactive of Dashboards

Figure 3.6 presents examples of Power BI visualisations developed to enhance dashboard interactivity and analytical clarity. The integration of structured data processing with intuitive visualisation bridges the gap between raw warranty data and decision ready information, ensuring that both technical and user requirements are addressed [134].

The interactive dashboard illustrated in Figure 3.7 demonstrates how warranty claim data can be analysed longitudinally by linking production dates, sales volumes,

and service timelines. This structure enables the assessment of time gaps between production and failure occurrence, supporting trend analysis and predictive modelling. By examining claim frequencies across service months, the system facilitates identification of recurring failure patterns and abnormal deviations, which are essential for parametric failure analysis and warranty forecasting [113], [134].

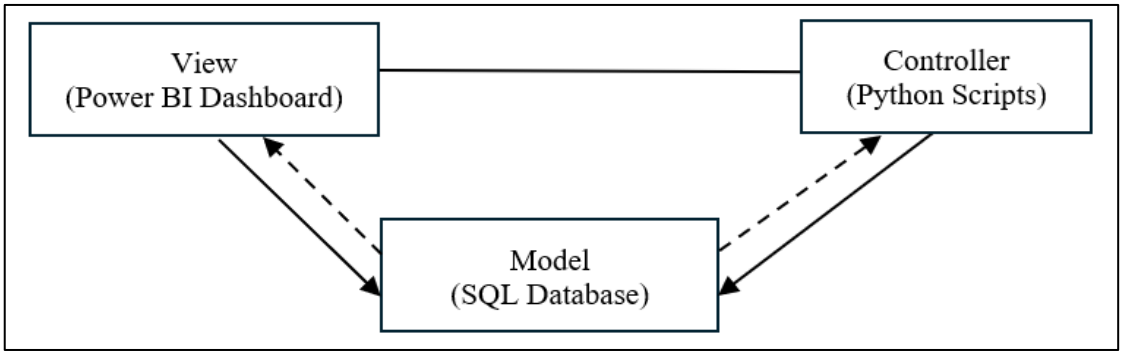


Figure 3.7 MVC Framework for System Architecture

To support long-term scalability, modularity, and ease of maintenance, the proposed warranty analysis system was designed using the Model–View–Controller (MVC) architectural framework. This structure enables clear separation between data management, system logic, and user interaction, thereby improving system robustness and adaptability.

Within this architecture, the Model layer is responsible for data handling and analytical computation. Structured warranty records are stored in SQL-based databases to ensure data integrity and efficient querying, while Python scripts are utilised for data preprocessing, transformation, anomaly detection, and predictive analysis [77] [135]. This layer provides a reliable analytical foundation for downstream processing and visualisation.

The Controller layer, implemented using Python, serves as the core processing unit of the system. It coordinates data retrieval from the Model layer, manages business logic, executes analytical workflows, and prepares processed outputs for presentation. In addition, the Controller handles user-driven requests and ensures that analytical results are generated consistently and accurately before being transmitted to the visual interface.

The View layer is implemented using Microsoft Power BI, which functions as the primary user interface for the system [127], [131]. Power BI presents analytical

outputs through interactive dashboards that allow users to explore warranty trends, failure patterns, and cost distributions intuitively. Connectivity between the Controller layer and Power BI is established through data gateways or scheduled refresh mechanisms, ensuring that dashboard content remains up to date and responsive to newly processed data.

This MVC-based design framework enables the integration of advanced analytics with interactive visualisation while maintaining flexibility for future system expansion, making it suitable for real-world automotive warranty management applications. The clear separation of concerns between the Model, View, and Controller layers allows backend enhancements to be implemented independently of frontend design changes. This modularity supports isolated testing, efficient debugging, and continuous system improvement [136], [137].

The adoption of the MVC framework enables seamless integration of predictive analytics and real time visualisation, enhancing system reliability and adaptability. Such an architecture is particularly suitable for cloud based and real time analytics applications in the automotive aftermarket. By combining structured data storage, advanced analytics, and interactive visualisation within a unified framework, the proposed system strengthens warranty management capabilities, supports proactive decision making, and contributes to improved product quality, cost control, and customer satisfaction [77], [134], [137].

3.2.4 Phase 4: Parametric Analysis and Dashboard Design

Phase 4 focuses on parametric analysis and dashboard design, forming a critical bridge between structured warranty data representation and actionable analytical insights. At this stage, warranty claims are analysed using clearly defined measurable parameters to enable systematic diagnosis of failure trends and support forecasting of potential failures over time. The parametric design approach provides a structured representation of warranty behaviour, enabling both temporal failure analysis and integration with interactive dashboards.

3.2.4.1 Parametric Warranty Data Structuring

The parametric analysis was guided by a warranty data matrix concept adapted from established warranty analysis literature [138]. The matrix aligns sales volume against warranty claims received over successive time intervals. Each row represents a cohort of products sold within a specific period, while each column captures the number of warranty claims received in subsequent months. This structure enables longitudinal tracking of product reliability across the warranty lifecycle.

The key parameters defining the warranty data matrix are as follows [123], [139]:

a) Di: Sales date
b) Mi: Number of units sold on date Di
c) rij: Number of claims received in the j-th month after sales date Di
d) rj: Total number of claims received in month j across all sales dates
e) M: Total number of units sold

This parametric formulation is essential for detecting time-dependent failure patterns, identifying high-risk service periods, and estimating product reliability. Warranty data from Company MNC were cleaned, grouped, and formatted in Microsoft Excel to conform to this matrix structure. Sales and claim records were aggregated by sale month and claim month, allowing calculation of monthly failure rates (rij / Mi) for each production batch. This enabled clustering of failure behaviour and identification of critical claim accumulation periods.

The matrix-based approach is particularly effective for large, longitudinal datasets such as Company MNC's warranty database, where claims are accumulated over extended periods and linked to specific vehicle models and components. By structuring warranty data parametrically, defect trends become more visible, supporting preventive maintenance planning and data-driven warranty policy formulation [140].

Table 3.3
Parameter of Warranty Data

Date of sales	**Sold volume**	**Claim received**							
		1	**2**	**…**	**m_o**	**$m_o + 1$**	**…**	**n_o-1**	**n_o**
D_1	M_1	r_{11}	r_{12}	…	$r_{1,mo}$	$r_{1,mo+1}$	…	$r_{1,no-1}$	$r_{1,no}$
D_2	M_2		r_{21}	…	$r_{2,mo-1}$	$r_{2,mo-1}$	…	$r_{2,no-2}$	$r_{2,no-1}$
…	…			…	…		…	…	…
D_{m0-1}	M_{mo-1}			…	$r_{mo-1,2}$	$r_{mo-1,3}$	…	$r_{mo-r1,no-mo+1}$	$r_{mo,no-mo+2}$
D_{mo}	M_{mo}			…	$r_{mo,1}$	r_{mo-2}	…	$r_{mo,no-mo}$	$r_{mo,no-mo+1}$
Total	**M**	**r_1**	**r_2**	**…**	**r_{mo}**	**r_{mo+1}**	**…**	**r_{no-1}**	**r_{no}**

Table 3.3 presents the structured warranty data matrix used in this study, organising production volumes and claims received across the product lifecycle. The first columns capture manufacturing and sales information, while subsequent columns represent claims received at increasing service durations, enabling lifecycle-based reliability assessment.

3.2.4.2 Dashboard Design and Analytical Implementation

The detailed design stage translates the parametric warranty structure into a functional and interactive analytical system implemented in Microsoft Power BI. Power BI was selected due to its capability to process large datasets, establish relational data models, and support real-time interactive visualization. The dashboard was designed to support multiple stakeholders, including quality engineers, managers, and service analysts, enabling effective exploration of warranty trends.

Key dashboard elements include:

- Total warranty claims by month and year
- Top failing parts and failure frequency
- Warranty cost trends over time
- Failure rates by vehicle model, component, and region
- Claim lead time between sales and claim occurrence
- Dynamic filtering by model, time period, dealer, and failure category

Visual components such as bar charts, line charts, KPI cards, scatter plots, and matrix tables were employed to present insights clearly and intuitively. Power BI's Data Analysis Expressions language was used to compute claim rates, failure ratios, and aggregated cost metrics. Relationships between vehicle data, claims data, and part categories were established using a star schema to ensure high performance and efficient drill-down analysis [141].

To ensure replicability and reliability, the dashboard development process began with importing cleaned datasets from Excel, each containing essential fields such as vehicle model, claim date, part name, failure code, mileage, and claim cost. The dashboard was iteratively tested using historical warranty data, and feedback from the Supplier Quality Assurance department at Company MNC was incorporated to refine usability, clarity, and analytical relevance. This iterative refinement ensured that the dashboard met user requirements while supporting real-time decision-making [140].

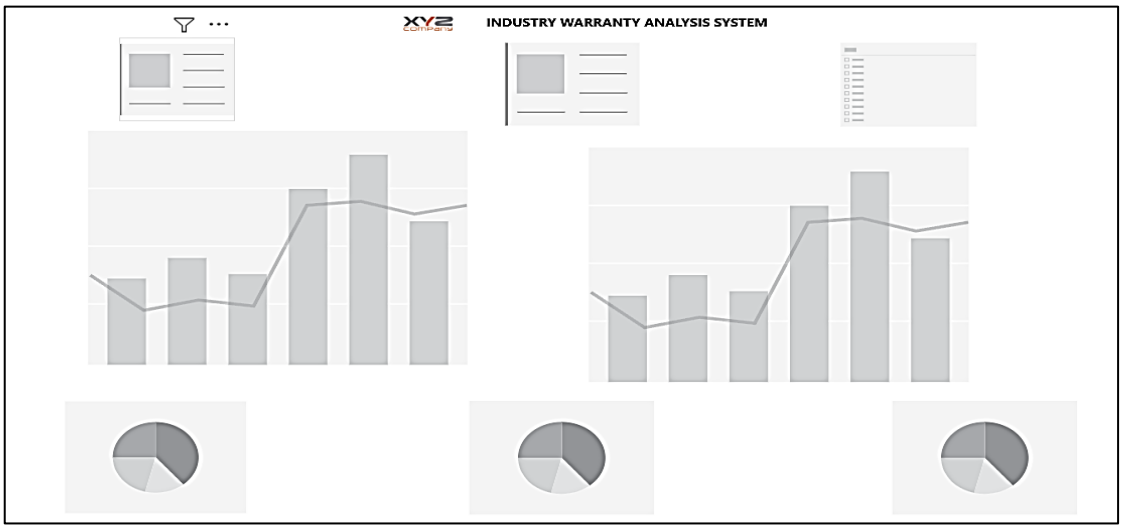


Figure 3.8 Layout ff Automotive Warranty Analysis System Dashboard

Figure 3.8 illustrates the overall layout of the Automotive Warranty Analysis System dashboard, while Figure 3.9 presents the detailed graphical user interface design. The interface follows a modular layout with structured navigation panels, interactive visual elements, and clearly defined analytical sections. Navigation menus allow users to switch between warranty claims, cost analysis, failure trends, and customised reports, enhancing accessibility and analytical efficiency.

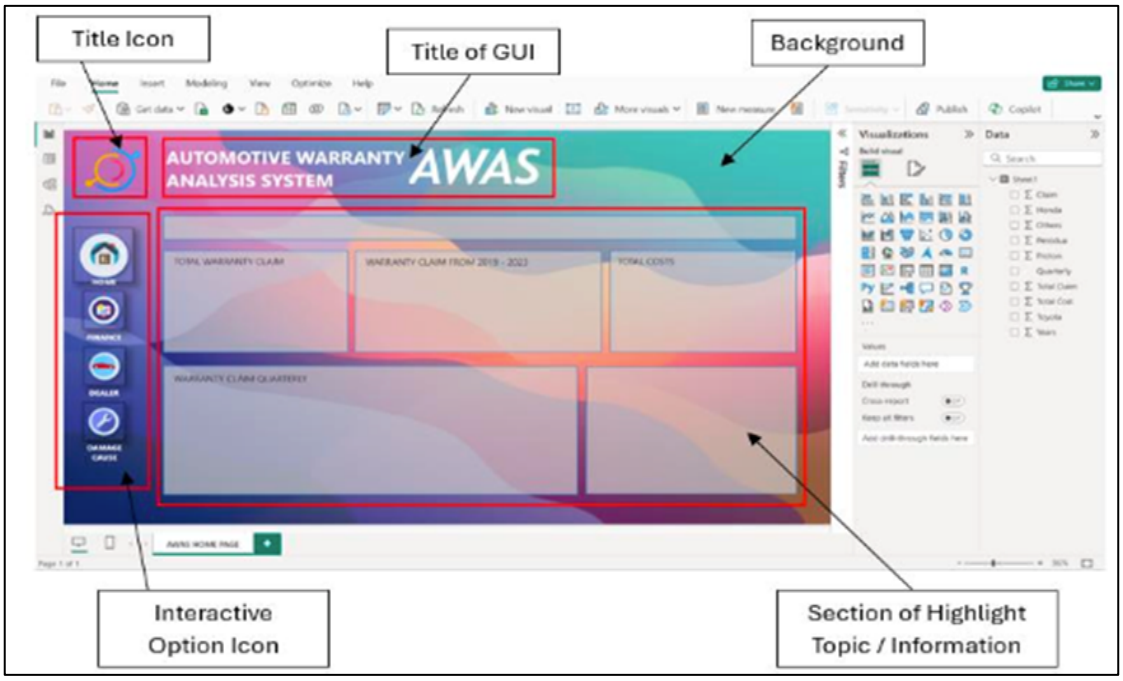


Figure 3.9 Detailed GUI Design for Warranty Management System

3.2.4.3 Power BI Integration and Predictive Capability

Microsoft Power BI functions as the primary visualisation and reporting layer within the proposed warranty analysis system, providing interactive and near real-time insight into warranty performance. Its capabilities include dynamic filtering, drill-down exploration, slicer-based segmentation, scheduled data refresh, and rule-based alerting, all of which support timely identification of abnormal claim behaviour and emerging quality concerns. These features allow users to explore warranty trends at varying levels of detail and respond proactively to deviations from expected performance. Previous studies have demonstrated that Power BI–enabled dashboards improve information transparency, analytical accessibility, and decision-making effectiveness within enterprise-scale analytics environments [136].

The system architecture has been deliberately designed to support scalability and future extensibility. Additional data sources enhanced analytical modules, and more advanced forecasting techniques can be incorporated without requiring fundamental changes to existing system workflows. This modular integration approach ensures that the platform remains adaptable to evolving business and analytical requirements.

Predictive capability is strengthened through the integration of AI-driven analytical models, enabling forward-looking assessment of warranty claim behaviour, failure pattern evolution, and cost exposure under different operational scenarios. By

combining business intelligence with predictive analytics, the system supports not only descriptive and diagnostic analysis but also anticipatory decision-making. Prior research confirms that real-time analytics platforms augmented with predictive modelling significantly enhance forecasting reliability and operational efficiency, particularly in data-intensive industrial applications [137].

3.2.4.4 Strategic Impact

The integration of parametric warranty analysis, interactive visualisation, and predictive analytics enables the proposed system to function as a comprehensive decision-support platform rather than a purely descriptive reporting tool. By combining structured analytical methods with real-time dashboard insights, the system facilitates timely interpretation of warranty trends and supports evidence-based managerial and engineering decisions. This integrated approach enhances the effectiveness of warranty management by improving visibility into failure behaviour, enabling proactive intervention, and supporting systematic quality improvement initiatives. Consequently, the system contributes to improved product reliability, more efficient warranty cost control, and sustained alignment with long-term organisational performance and competitiveness objectives.

3.3 Model Validation

This study adopts a quantitative research paradigm to systematically evaluate and compare forecasting models for automotive warranty cost prediction. An empirical comparative analysis is employed using real-world warranty claim data obtained from Company MNC to assess the relative performance of multiple forecasting techniques. The quantitative framework enables objective evaluation through standardised performance metrics, including Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). These metrics allow consistent and statistically meaningful comparison across models while maintaining strong practical relevance for industrial warranty management applications.

The validation methodology contrasts traditional statistical approaches—namely Basic Linear Regression, Enhanced Linear Regression, and ARIMA—with artificial intelligence–based techniques, including Random Forest, XGBoost, and

Ensemble models. This comparative structure is designed to identify the most effective forecasting approach for component-level warranty cost prediction, particularly under small dataset constraints, which are common in industrial warranty environments.

3.3.1 Software and Tools Implementation

The research was implemented using the following software and tools:

a) Development Environment: Google Colab cloud-based platform with Python 3.8+ for accessibility and computational resource availability
b) Data Processing Libraries: pandas for data manipulation and analysis, NumPy for numerical computations and array operations
c) Statistical Analysis: statsmodels for ARIMA implementation and statistical testing procedures
d) Machine Learning Frameworks: scikit-learn for Linear Regression and Random Forest models, XGBoost for gradient boosting implementation
e) Visualization Tools: matplotlib and seaborn for data visualization and results presentation

3.3.2 Methodological Framework Overview

A four-phase structured methodological framework was adopted to systematically develop, validate, and deploy warranty cost forecasting models while preserving temporal integrity.

1. **Phase 1: Data Preparation and Structuring**
   This phase involved data extraction from Company MNC, data quality assessment, feature filtering, and temporal structuring, including treatment of disruption periods.
2. **Phase 2: Exploratory Data Analysis (EDA)**
   EDA established the time-series foundation, performed statistical testing, and identified business-relevant patterns to guide model selection and feature engineering.

3. **Phase 3: Model Development**

   Multiple forecasting models were developed in parallel following systematic feature engineering. Core models included Linear Regression, ARIMA, Random Forest, and XGBoost.

4. **Phase 4: Model Walk-Forward Validation Framework**

   An iterative walk-forward validation approach was applied across three progressive training cycles. Enhanced Linear Regression and Ensemble models were developed and refined during this phase.

5. **Phase 5: Production Forecasting**

   The forecasting phase applied a multi-model comparative approach to estimate warranty claim costs, incorporating systematic model retraining and ensemble methods to support robust forecast generation.

All models were retrained using consolidated datasets and validated against FY2024 data. Comparative performance analysis was conducted to identify the most suitable model for real-world deployment. This structured framework ensures methodological rigor, preserves chronological data integrity, and addresses the challenges associated with warranty forecasting under limited data conditions.

3.3.2.1 Phase 1: Data Preparation and Structuring

The dataset was obtained from Company MNC, a vehicle component manufacturer supplying multiple automotive customers. Warranty claim records were provided as separate Microsoft Excel files for each fiscal year from FY2019 to FY2024. Data from FY2019 to FY2023 were consolidated into a single dataset for model development, while FY2024 data were reserved exclusively for out-of-sample validation to ensure unbiased performance assessment.

The dataset contained confidential industrial records, including claim dates, component details, monetary values, and internal reference identifiers. To align with the research scope, the data were filtered to include only alternator components supplied to a single customer, referred to as Company A.

The original Excel files contained non-essential elements, including non-data header rows and empty or internal-reference columns. These elements were removed to

retain analytically relevant information only. The cleaned dataset was subsequently exported to CSV format to ensure compatibility with Python-based analytical tools.

The final dataset comprised 2,514 accepted warranty claims, representing a total claim cost of RM 373,702.16, with individual claim values ranging from RM 322.38 to RM 22,544.85. The data covered a 50-month period from April 2019 to March 2024, consistent with the Japanese fiscal year calendar. A total of 49 distinct features were retained, including temporal identifiers, monetary values, and component specifications. Only approved claims were included to ensure accuracy in financial forecasting, with rejected or pending claims excluded. All records were anonymised under confidentiality agreements to protect proprietary information while preserving analytical validity for model development and validation.

a) Environment Setup and Data Loading.

The computational environment was configured using Google Colab with Python 3.8 or later, ensuring reproducibility and access to adequate computational resources. Essential Python libraries were systematically imported, including pandas, NumPy, matplotlib, seaborn, statsmodels, scikit-learn, and XGBoost.

Visualisation parameters were standardised using seaborn styling for consistency, and non-critical warning messages were suppressed to streamline output readability. The dataset was securely uploaded to the Colab environment using the built-in file upload mechanism. Multiple character encoding formats (UTF-8, Latin-1, CP1252, ISO-8859-1) were tested to ensure accurate data interpretation. The dataset was successfully loaded using UTF-8 encoding, preserving data integrity and international character compatibility. Initial structural verification confirmed the presence of 2,514 records across 49 features, with a memory footprint of approximately 2.5 MB, indicating efficient computational handling. Data type classification procedures were applied to distinguish numerical and categorical variables. Key variables including *Claimed Month*, *Acceptance Status*, *XYZ_RM* (warranty cost), and temporal identifiers were identified and validated for forecasting model development. Particular attention was given to the target variable *XYZ_RM*, which served as the primary output for warranty cost prediction.

b) Data Quality Assessment & Filtering.

Data quality assessment was conducted to ensure the reliability and analytical validity of the warranty dataset prior to forecasting model development. An initial screening was performed using the *Accept/Reject* status attribute to distinguish between realised and non-realised warranty claims. Only accepted claims (ACC), which represent actual financial liabilities incurred by Company MNC, were retained for analysis. Rejected claim categories, including *INS*, *TFC*, *Over Warranty*, and *Over Mileage*, were excluded, as these entries do not contribute to actual warranty expenditure and would distort cost forecasting outcomes.

Validation of the target variable was subsequently carried out by comparing two available cost-related fields: *RM* (claimed warranty cost) and *XYZ_RM* (approved warranty cost). A detailed data integrity assessment revealed that the *RM* field exhibited complete numerical formatting and consistent values across all records. In contrast, the *XYZ_RM* field contained multiple computational inconsistencies, including unresolved reference errors (e.g., #REF!), which compromised its suitability for reliable quantitative analysis. Based on this comparative evaluation, the *RM* variable was selected as the primary target variable for warranty cost forecasting due to its superior completeness, consistency, and numerical reliability.

Following target variable selection, additional filtering procedures were applied to retain only accepted warranty claims, thereby reducing the dataset from the initial 2,514 records to a refined subset representing actual financial exposure for Company MNC. Duplicate detection was performed using available *Claim ID* identifiers, and any duplicated records were systematically removed to prevent analytical bias and overestimation of warranty costs.

Final data validation confirmed full completeness of the selected target variable, absence of duplicate entries, and sufficient temporal coverage to support robust time-series analysis and predictive modelling. These data quality assurance steps ensured that the refined dataset provided a reliable foundation for subsequent exploratory analysis, model development, and validation phases.

c) Fiscal Year Mapping and Temporal Structure Creation.

Temporal structuring procedures were implemented to establish an appropriate time-series framework aligned with Company MNC's fiscal calendar and warranty reporting practices. The original dataset utilised a six-digit date format (YYYYMM) in the *Claimed Month* field, which required systematic transformation to reflect Company

MNC's fiscal year cycle, defined from April to March, in accordance with its automotive component supply operations.

Custom fiscal year mapping logic was developed to convert calendar months into corresponding fiscal year classifications. Under this mapping scheme, months from April to December were assigned to the current fiscal year, while months from January to March were allocated to the preceding fiscal year to preserve fiscal reporting consistency. In addition, a fiscal month index was introduced to support seasonal analysis, with April designated as fiscal month 1 and March as fiscal month 12 within each fiscal year cycle.

The application of fiscal year mapping resulted in the creation of structured temporal variables,including *Calendar_Year*, *Calendar_Month*, *Fiscal_Year*, *Fiscal_Month*, and *Fiscal_Period*. Verification of the transformed dataset confirmed warranty data coverage spanning fiscal years FY2019 to FY2023. Variations in monthly data availability were observed across fiscal years, reflecting operational conditions and external disruptions affecting Company MNC's automotive supply chain.

A business-informed assessment was conducted to identify and classify disruption periods associated with the COVID-19 pandemic. The analysis confirmed that FY2020 (April 2020 to March 2021) experienced significant manufacturing and supply chain interruptions, resulting in only two months of recorded warranty data. These records primarily reflected claims accumulated during post-shutdown restart activities rather than normal operational patterns. In contrast, FY2019, FY2021, FY2022, and FY2023 each exhibited complete 12-month data coverage. To prevent bias and distortion in forecasting model training, FY2020 was excluded from model development, while remaining fiscal years were retained to ensure representative and stable temporal patterns. This exclusion was grounded in operational context specific to Company MNC and ensured accurate treatment of disruption effects in warranty forecasting.

Subsequently, monthly time-series aggregation was performed to construct forecasting-ready datasets with suitable temporal granularity for component-level warranty cost prediction. Individual warranty claim records were aggregated into monthly observations, generating derived variables such as monthly claim counts, total monthly warranty cost (*Total_RM*), average monthly claim cost (*Mean_RM*), and related statistical measures. Throughout the aggregation process, fiscal year classifications and chronological order were preserved. Data integrity checks confirmed

full retention of total claim volumes and financial values, resulting in a continuous 50-month time series suitable for robust time-series and machine learning–based forecasting in support of Company MNC's warranty management operations.

3.3.2.2 Phase 2: Exploratory Data Analysis (EDA)

Phase 2 of the methodology focuses on Exploratory Data Analysis (EDA) to develop a comprehensive understanding of the structure, behaviour, and temporal characteristics of Company MNC's warranty cost data. This phase aims to identify underlying patterns, trends, and anomalies that inform subsequent forecasting model selection and configuration. EDA serves as a critical bridge between data preparation and model development, ensuring that predictive techniques are grounded in empirical characteristics of the warranty dataset.

Following preliminary data handling and validation in Microsoft Excel, EDA is applied to provide a deeper and more granular understanding of warranty claim behaviour across multiple fiscal years. This process supports the identification of focus components, such as recurring issues associated with the Alternator Assembly, thereby guiding targeted quality improvement initiatives. By enabling early insight into claim dynamics, EDA forms a critical foundation for developing predictive models that anticipate warranty occurrences, optimise resource allocation, and enhance operational efficiency.

a) Time Series Foundation Analysis

Time series foundation analysis was conducted to establish a detailed understanding of warranty cost dynamics and temporal behaviour relevant to Company MNC's warranty management operations. Individual warranty claim records were aggregated into monthly time-series observations to provide an appropriate temporal resolution for forecasting applications. The resulting dataset spans a total of 50 months, covering the period from April 2019 to March 2024, and reflects Company MNC's fiscal reporting structure.

The preparation of the time series involved transforming individual claim-level data into monthly summary metrics, including total warranty cost, claim frequency, and average claim value. These transformations preserved fiscal year classifications and chronological order to ensure temporal integrity. This structured aggregation enabled

systematic examination of long-term trends, short-term fluctuations, and potential seasonal effects associated with Company MNC's automotive component warranty obligations.

Figure 3.10 presents the consolidated time-series visualisation, illustrating monthly warranty cost evolution, claim count trends, and variations in average claim amounts over time. Periods affected by COVID-19–related business disruptions are clearly identifiable as gaps or irregular patterns in the time series, highlighting their impact on warranty activity. Data completeness assessment confirmed full 12-month coverage for FY2019, FY2021, FY2022, and FY2023, while FY2020 exhibited limited data availability due to pandemic-induced manufacturing and supply chain interruptions. These observations provided critical contextual input for subsequent model development and validation decisions.

Figure 3.10 Line Graphs Showing Monthly Warranty Claim Trends with COVID-19 Impact.

Figure 3.11 presents the monthly warranty claims analysis for Company MNC over the period from April 2019 to March 2024. The figure illustrates (top) total monthly warranty costs in RM, (middle) monthly warranty claim counts, and (bottom) the evolution of average claim amounts. The red shaded region denotes the COVID-19–related business disruption period spanning April 2020 to January 2021, while dashed vertical lines indicate fiscal year boundaries based on Company MNC's April–March reporting cycle. The visualisation highlights pronounced volatility in warranty activity, alongside an overall declining trend in warranty obligations across the observed period.

Statistical characterisation of the time series indicates substantial temporal variability in Company MNC's monthly warranty costs. The coefficient of variation exceeds 0.6, reflecting significant month-to-month fluctuations and confirming the presence of instability that cannot be adequately captured using simple trend-based forecasting methods. Such variability underscores the necessity for advanced statistical and machine learning–based modelling approaches capable of handling non-stationary and volatile warranty data.

Further insight into the temporal dynamics is provided through rolling mean and rolling variance analysis, as illustrated in Figure 3.10. The rolling analysis reveals distinct periods of heightened volatility, which coincide with major business transitions and operational changes affecting Company MNC's warranty activities. These findings provide critical empirical justification for the selection of forecasting models that explicitly account for changing variance structures, temporal dependence, and external disruptions within the warranty cost timeseries.

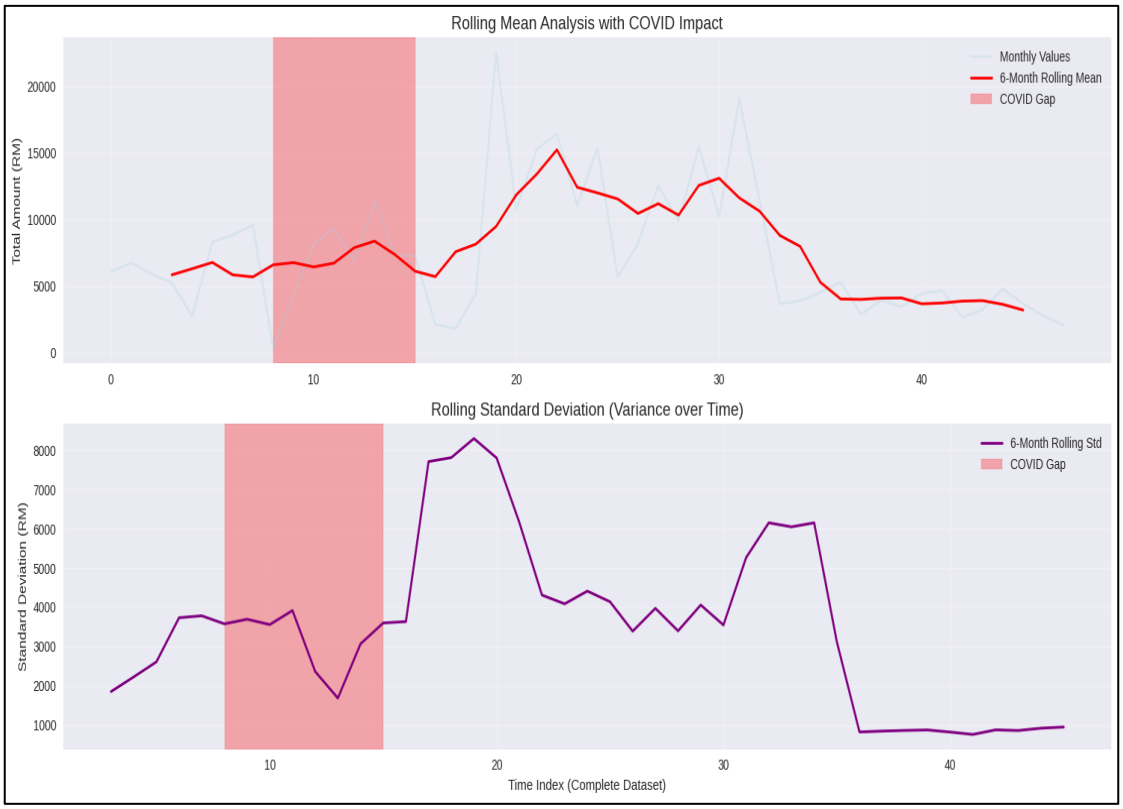


Figure 3.11 Line Graphs Showing Rolling Mean and Rolling Standard Deviation of Monthly Warranty Claim Amounts with COVID-19 Impact

Figure 3.11 illustrates the six-month rolling mean and rolling standard deviation of monthly warranty costs to examine temporal volatility and trend characteristics in Company MNC's warranty data. The upper panel presents the rolling mean analysis, where individual monthly warranty cost observations are shown as a light background series, while the six-month rolling average is depicted as a solid red line. This representation highlights the underlying trend movements by smoothing short-term fluctuations, thereby revealing longer-term changes in warranty cost behaviour. The lower panel displays the rolling standard deviation, which captures changes in variability over time and identifies periods of elevated volatility. The COVID-19–related business disruption period is explicitly highlighted, demonstrating its pronounced impact on warranty cost variability within Company MNC's operations.

Seasonal pattern assessment was further conducted through fiscal month aggregation to identify recurring temporal effects in warranty cost obligations. Figure 3.12 presents the calendar month analysis, revealing systematic seasonal variations across the fiscal year. The results indicate that November exhibits the highest average warranty cost (RM 13,605), while August records the lowest (RM 4,798), reflecting a strong seasonal structure. The calculated seasonality strength exceeding 1.0 confirms

the presence of pronounced seasonal effects in the warranty cost time series. These findings provide essential empirical justification for incorporating seasonality-aware forecasting techniques during model development and validation, ensuring that predictive models adequately capture the temporal dynamics inherent in Company MNC's warranty cost behaviour.

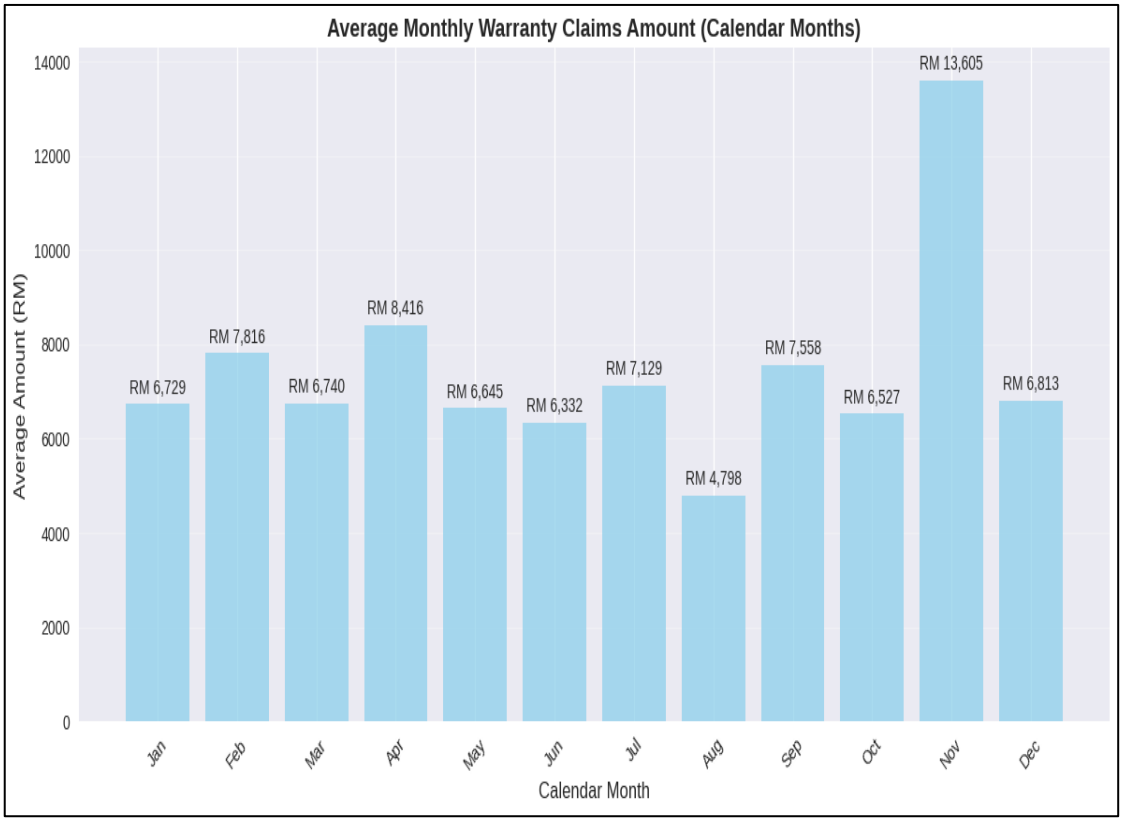


Figure 3.12 Average Monthly Warranty Claim Amounts by Calendar Month

Figure 3.12 presents the average monthly warranty costs aggregated by calendar month across the full dataset, excluding FY2020 to avoid distortion from pandemic-related operational disruptions. The bar chart illustrates clear seasonal patterns in Company MNC's warranty obligations, with bar heights representing average monthly costs and corresponding values annotated above each bar. The results indicate that November exhibits the highest seasonal warranty expenditure (RM 13,605), while August records the lowest (RM 4,798). This pronounced variation confirms the presence of strong seasonal effects within the warranty cost time series, underscoring the necessity of incorporating seasonality-aware mechanisms in the development and validation of forecasting models.

b) Business Pattern & Feature Analysis

Statistical findings were interpreted in conjunction with Company MNC's operational context, including fiscal year cycles, seasonal component failure behaviour, and external disruptions such as the COVID-19 pandemic. The exclusion of incomplete FY2020 data was justified on both statistical and business grounds, ensuring temporal balance and reliability for walk-forward validation. This integration of econometric testing with business insight ensured that model selection aligned with practical warranty forecasting requirements. Business pattern and feature analysis was conducted to identify meaningful predictors for regression-based and machine learning models while maintaining statistical validity under small-sample constraints. The analysis focused on key business dimensions influencing warranty claim behaviour and cost variability.

Vehicle model distribution was examined to identify primary contributors to warranty claim volume and cost variation. Claim frequencies were analysed across all vehicle models using both absolute counts and percentage distributions to assess concentration effects. As summarised in Table 3.4, the analysis revealed a high concentration of warranty activity, with two dominant models—AR and BZ—accounting for 83.9% of total warranty claims. These models exhibited distinct cost profiles, with average claim costs of RM 120.69 for AR and RM 184.05 for BZ, highlighting the importance of incorporating model-specific effects into warranty cost forecasting.

Table 3.4
Vehicle Model Distribution and Claim Characteristics

Vehicle Model	Total Claims	Percentage (%)	Mean Cost (RM)	Std Dev (RM)	Total Cost (RM)	Claim Period
AR	1,145	47.6	120.69	86.22	138,196	202102–202403
BZ	873	36.3	184.05	102.8	160,676	201904–202403
MV	127	5.3	125.72	94.25	15,966	201905–202403
AX	121	5	184.02	57.69	22,267	201904–202403
NEW MV	121	5	255.39	123.82	30,903	201904–202205

Vehicle Model	Total Claims	Percentage (%)	Mean Cost (RM)	Std Dev (RM)	Total Cost (RM)	Claim Period
AT	11	0.5	398.48	110.55	4,383	202201–202309
NEW MV	3	0.1	274.19	0	823	202310–202403
NEW AL	2	0.1	244.29	0	489	202308–202311
Total	2,403	100	155.61	98.37	373,703	201904–202403

Statistical summaries including mean warranty claim cost, standard deviation, minimum and maximum values, and total warranty expenditure were computed for each vehicle model to evaluate cost variability and distributional differences. As presented in Table 3.4, substantial variation in warranty costs is observed across vehicle models. Newer models, such as AT, exhibit markedly higher average claim costs (RM 398.48) compared to more established models such as AR (RM 120.69), indicating differing reliability profiles and cost exposure across the vehicle portfolio.

To assess temporal consistency in warranty behaviour, a temporal stability analysis was conducted by examining model-specific claim patterns across fiscal years. The coefficient of variation (CV) was calculated for each major vehicle model to quantify the stability of warranty cost patterns over time, with results summarised in Table 3.5. CV values below 0.5 were interpreted as indicating stable warranty behaviour, values between 0.5 and 1.0 representing moderate variability, and values exceeding 1.0 reflecting highly volatile claim patterns. The observed variability across models provided empirical justification for incorporating vehicle model indicators as categorical explanatory variables in subsequent regression and machine learning–based forecasting models.

Table 3.5
Seasonal Pattern Strength Analysis by Vehicle Model

Vehicle Model	Seasonal Strength	Pattern Classification	Coefficient of Variation (FY)	Temporal Stability
AR	1.059	Strong	1.006	Volatile
BZ	0.893	Moderate	0.5	Moderate
MV	1.134	Strong	1.225	Volatile

AX	1.289	Strong	0.577	Moderate
NEW MV	1.983	Strong	0.867	Moderate

An analysis of part category distribution was performed to evaluate the structural composition and cost behaviour of warranty claims within the alternator-related dataset. As presented in Table 3.6, warranty cases were predominantly associated with Electrical and Electronic (E/E) components, which accounted for 86.2% of the total claims. In contrast, Air Conditioning (A/C) components contributed 13.8% of the recorded cases. From a cost perspective, the average claim value for E/E components was RM 153.42, while A/C related claims exhibited a marginally higher mean cost of RM 168.61, indicating slightly greater cost exposure per incident for A/C failures despite their lower frequency.

Further investigation into failure characteristics was conducted using original failed part identifiers and associated condition descriptions to determine dominant failure modes. The analysis demonstrated a highly concentrated failure pattern, where a single part number, TG1042112560, was responsible for 87.2% of the total warranty claims. This strong concentration suggests that overall warranty expenditure was primarily driven by a specific component issue rather than distributed across multiple part types or categories.

The results of this analysis provided critical insight into the key contributors to warranty claims and supported the formulation of targeted, failure-specific predictive features. However, the limited variability across part categories and the overwhelming dominance of a single failed part reduced the explanatory value of part category variables in the modelling process. Consequently, these variables were assigned a lower priority during feature selection compared to vehicle model attributes and temporal indicators, as summarised in Table 3.7. This prioritisation approach ensured that model inputs remained aligned with observable data variation and contributed meaningfully to predictive performance.

Table 3.6
Part Category Analysis Summary

Part Category	Total Claims	Percentage (%)	Mean Cost (RM)	Std Dev (RM)	Total Cost (RM)	Models Affected
E/E	2,072	86.2	153.42	101.46	317,892	8
A/C	331	13.8	168.61	80.63	55,810	7
Total	2,403	100	155.61	98.37	373,703	8

Table 3.7
Feature Engineering Priority Matrix

Feature Category	Feature Type	Priority Level	Business Justification	Statistical Significance
Temporal	Fiscal_Month	High	Strong seasonality observed (>1.0 seasonal strength)	Chi-square $p < 0.001$
Temporal	Time_Index	High	Clear linear trend in warranty costs	$R^2 = 0.64$
Categorical	Model_Dummies	Medium	8 manageable categories, significant cost differences	ANOVA $p < 0.001$
Temporal	COVID_Dummy	Medium	Structural break identified in FY20	t-test $p < 0.05$
Categorical	Part_Category	Low	Only 2 categories (limited variability)	t-test $p < 0.01$
Business	Vehicle_Age	Medium	Component lifecycle dependency	Correlation $r = 0.34$
Business	Claim_Volume	Medium	Operational capacity indicator	Correlation $r = 0.28$

Temporal pattern analysis was performed to examine seasonal and inter-year variations in warranty claim behaviour. An assessment based on fiscal months was conducted to analyse claim distributions across the 12-month fiscal cycle, allowing systematic identification of recurring seasonal effects in both claim frequency and magnitude.

The degree of seasonality for major vehicle models was quantified using a seasonality strength index defined as the ratio between the range of monthly claims and the corresponding mean monthly claims, as expressed below:

$$\text{Seasonality Strength} = \frac{\text{Maximum Monthly Claims} - \text{Minimum Monthly Claims}}{\text{Mean Monthly Claims}} \tag{3.2}$$

As presented in Table 3.5, several vehicle models demonstrated strong seasonal characteristics, with index values exceeding unity. In particular, models AR (1.059), MV (1.134), AX (1.289), and NEW MV (1.983) exhibited pronounced seasonal fluctuations, indicating that warranty claim behaviour for these models is significantly influenced by temporal factors. These results confirm the necessity of explicitly incorporating seasonality-related variables in subsequent forecasting models.

To assess the stability of warranty claim patterns across fiscal years, a chi-square test of independence was employed to evaluate cross-year consistency. The analysis produced a chi-square statistic of 566.68 with a corresponding p-value below 0.0001, indicating statistically significant differences in claim distributions across fiscal periods. This finding suggests the presence of temporal structural changes, which may be attributed to product revisions, manufacturing process improvements, or external influences such as market or supply chain disruptions.

The outcomes of this temporal analysis directly informed the feature engineering strategy. Variables capturing fiscal timing and seasonal effects were prioritised as key model inputs, while identified structural variations were explicitly addressed during model training and validation. The resulting feature prioritisation is summarised in the feature matrix provided in Table 3.7. Based on the results of the business pattern analysis and exploratory data assessment, a structured feature engineering framework was established. Candidate predictors were systematically grouped into three principal categories to ensure adequate representation of warranty cost determinants, namely temporal features, categorical features, and business-related indicators.

Feature prioritisation was determined through a combination of statistical relevance and domain-driven considerations. Temporal variables, particularly fiscal month indicators, were assigned high priority due to the strong seasonal patterns identified in Table 3.5, alongside linear time indices reflecting long-term trends in warranty costs. Medium-priority features included vehicle model indicators and dummy variables representing the COVID-19 period, reflecting their moderate but meaningful influence on warranty behaviour. In contrast, part category variables were assigned lower priority owing to the limited number of distinct categories within the alternator-focused dataset and their relatively low explanatory variability, as previously demonstrated. This structured approach ensured that model complexity remained

proportionate to data variability while preserving the most informative predictors for forecasting performance.

The consistency of identified business patterns was examined through formal statistical validation to determine their suitability for forecasting applications. Chi-square tests of independence were employed to assess whether critical relationships—particularly the distribution of vehicle models across fiscal years—exhibited stability over the study period. A significance level of $p < 0.05$ was adopted, and statistically significant outcomes were interpreted as indicators of temporal pattern variation that required explicit treatment during model formulation.

This validation step ensured that only relationships supported by sufficient statistical evidence were retained for predictive modelling, thereby mitigating the risk of overfitting or dependence on coincidental correlations associated with the relatively limited dataset size. By differentiating persistent structural patterns from those influenced by temporal changes or external disruptions, the analysis guided appropriate feature handling strategies, including the introduction of time-based indicators or interaction effects where necessary. The adoption of a conservative pattern validation approach reinforced the integrity of the walk-forward validation framework applied in subsequent modelling stages. This ensured that engineered features remained stable, generalisable, and appropriate for out-of-sample warranty cost forecasting, thereby strengthening the reliability of the overall predictive framework.

The outcomes of the business pattern analysis were systematically integrated into the model development strategy to ensure that feature selection and model configuration were both statistically justified and aligned with observed warranty behaviour. Insights obtained from the temporal analysis, feature prioritisation, and pattern validation stages directly informed the design of forecasting models applied in this study. For linear regression–based models, a parsimonious feature set was intentionally adopted, prioritising high-impact temporal variables and key categorical indicators identified in Table 3.7. This configuration supported model transparency and interpretability while reducing susceptibility to overfitting, which is particularly critical when modelling warranty costs using a limited historical dataset. In contrast, tree-based models, including Random Forest and XGBoost, were structured to accommodate a richer representation of predictor variables. These models incorporated extended feature sets, including interaction terms and derived business metrics, to capture non-linear relationships and complex pattern interactions identified during the exploratory

analysis. This differentiation in feature usage reflected the inherent modelling capabilities of each algorithm and ensured consistency between analytical findings and model design.

The integration of business pattern insights further justified the application of conservative regularisation and complexity control strategies during model training. Particular emphasis was placed on managing multi-level categorical variables and representing seasonal effects within the constrained temporal scope of the dataset. By grounding model configuration decisions in validated business patterns rather than purely statistical optimisation, the modelling framework achieved improved robustness, generalisability, and practical relevance for out-of-sample warranty cost forecasting.

3.3.2.3 Phase 3: Model Development

Phase 3 focuses on the systematic development of forecasting models for automotive warranty claim cost prediction. Multiple modelling paradigms were implemented to balance interpretability, predictive accuracy, and robustness under small-sample constraints. The modelling strategy intentionally contrasts traditional statistical methods with machine learning approaches to evaluate their relative suitability for component-level warranty forecasting.

Feature engineering was carried out to convert the cleaned monthly warranty data into structured input variables that could be used by different forecasting models. As the dataset consisted of only 48 monthly observations, the feature engineering process focused on simplicity, correct time ordering, and suitability for each type of model. Different feature sets were created to support both easy-to-interpret statistical models and more flexible machine learning models.

The process started with the cleaned monthly time-series dataset prepared in Phase 2. The dataset covered the period from April 2019 to March 2024 and contained 48 observations. The target variable was *Total_RM*, which represents the total warranty claim cost for each month. Before generating any features, the data were arranged in chronological order and checked to ensure there were no missing months or inconsistencies in time sequence.

Temporal features were created to capture long-term trends, seasonal effects, and major disruptions in warranty costs. A linear time index was introduced to represent

overall growth or decline over time, while a quadratic time term was added to allow for non-linear trend behaviour.

Seasonal patterns were represented using twelve fiscal month dummy variables (FM_1 to FM_12), enabling the models to learn differences between months. Fiscal year dummy variables (FY_2019, FY_2021, FY_2022, FY_2023) were also included to reflect year-to-year structural changes. Fiscal year 2020 was excluded due to incomplete data caused by the COVID-19 pandemic. In addition, a binary COVID-19 indicator was introduced to explicitly represent the disruption observed during August–September 2021. In total, 29 temporal features were created.

Lag features were generated to capture the influence of past warranty outcomes on current warranty costs. To reduce the risk of overfitting, the maximum lag length was limited to six months. Lagged values were created for both warranty cost (Total_RM_lag1 to Total_RM_lag6) and claim volume (Claims_Count_lag1 to Claims_Count_lag6), resulting in 12 lag-based features. Missing values created at the beginning of the time series were intentionally kept as NaN to preserve the correct time sequence and avoid using future information in earlier periods.

Rolling window features were used to describe short-term behaviour and variability in warranty costs. Rolling statistics were calculated using 3-month and 6-month windows, based on the dataset size and observed seasonal patterns. For each window, the rolling mean, standard deviation, minimum, and maximum values were computed. To ensure statistical reliability, rolling calculations required at least half of the window length to contain valid data. Missing values produced by rolling calculations were handled using forward and backward filling to maintain continuity. This step produced eight rolling window features.

Growth rate features were added to capture changes in warranty behaviour over time. Percentage changes were calculated for monthly (1-month), quarterly (3-month), and annual (12-month) intervals for both warranty cost and claim count. These features helped the models detect sudden increases or decreases that might not be visible from raw values alone. Missing values at the start of the series were retained to avoid distorting early data points.

Business-related features were created by aggregating transaction-level data into monthly summaries. Vehicle model composition features were generated for the four main models identified during exploratory analysis (AR, BZ, MV, AX). Both absolute

counts and proportional values were included to represent how the mix of vehicle models influenced warranty costs.

Part category indicators were also created for major subsystem groups, namely Electrical/Electronic and Air Conditioning components. All business features were grouped by claim month and merged into the main dataset using left joins to preserve the time structure. In total, 74 features were generated from 48 observations, resulting in a feature-to-observation ratio of 2.06. This high ratio indicates a strong risk of overfitting, especially for regression-based models. Correlation analysis revealed many highly correlated features, confirming the need for careful feature selection and regularisation during model development. Nineteen features contained missing values, mainly due to lag, rolling, and growth rate calculations. The proportion of missing values ranged from approximately 2% to 25%, depending on feature type.

To match feature complexity with model capability, three feature sets were defined, as summarised in Table 3.8:

1. **Basic Regression Feature Set**
   Consisted of 14 features, including a linear time trend, fiscal month indicators, and the COVID-19 indicator. This set was designed for stable and interpretable linear regression modelling.
2. **Enhanced Regression Feature Set**
   Expanded to 23 features by adding quadratic time terms, fiscal year indicators, and vehicle model proportions. This set provided improved explanatory power while remaining suitable for regression models.
3. **Full Feature Set**
   Included all 74 features and was intended for tree-based and ensemble models. Although the feature-to-observation ratio was high, this set was considered appropriate due to the built-in regularisation and feature selection mechanisms of Random Forest and XGBoost.

Missing values were handled according to feature type. Missing values in lag and growth rate features at the beginning of the series were kept preventing time leakage during walk-forward validation. Rolling window features were filled using forward and backward methods to maintain continuity. Categorical and business features were zero-

filled where appropriate to ensure compatibility across models. Overall, this structured feature engineering approach produced modelling-ready datasets of different complexity levels without requiring further preprocessing. Each forecasting model received a feature set tailored to its learning behaviour and sensitivity to overfitting, as detailed in Tables 3.8 and 3.9.

Table 3.8
Feature Engineering Summary by Model Type

Feature Category	Basic Regression	Enhanced Regression	Full Features	Description
Temporal Features	14	19	29	Time trends, seasonal dummies, fiscal year indicators
Lag Features	0	0	12	1–6-month lags for warranty costs and claim counts
Rolling Window	0	0	8	3-month and 6-month rolling statistics
Growth Rate	0	0	5	Monthly, quarterly, and annual growth rates
Business Features	0	4	20	Vehicle model proportions and part categories
Total Features	14	23	74	
Observations	48	48	47	Monthly records used in training
Feature Ratio	0.29	0.48	1.57	Features per observation
Overfitting Risk	Low	Moderate	High	Based on feature-to-observation ratio

Table 3.9
Feature Group Summary and Missing Value Treatment

Feature Category	Count	Key Examples	Missing Values	Treatment Strategy
Time Trends	2	Time_Linear, Time_Squared	0 (0%)	None required
Seasonal Dummies	12	FM_1, FM_2, ..., FM_12	0 (0%)	None required
Fiscal Year Indicators	5	FY_2019, FY_2021, FY_2022, FY_2023	0 (0%)	None required
COVID Impact	1	COVID_Impact (binary)	0 (0%)	None required
Warranty Cost Lags	6	Total_RM_lag1, Total_RM_lag3, Total_RM_lag6	1–6 (2.1–12.5%)	Preserved as NaN

Claim Count Lags	6	Claims_Count_lag1, Claims_Count_lag3, Claims_Count_lag6	1–6 (2.1–12.5%)	Preserved as NaN
Rolling Statistics	8	Total_RM_rolling_mean_3m, Total_RM_rolling_std_6m	2–5 (4.2–10.4%)	Forward/backward fill
Growth Rates	5	Total_RM_growth_1m, Total_RM_growth_12m	1–12 (2.1–25%)	Preserved as NaN
Vehicle Model Features	8	Count_ARUZ, Prop_BEZZA, Count_MYVI, Prop_AXIA	0 (0%)	Zero fill
Part Category Features	4	Count_E_E, Prop_E_E, Count_A_C, Prop_A_C	0 (0%)	Zero fill
TOTAL FEATURES	74	–	19 features affected	Multiple strategies

The complete feature list is available in Appendix A. Features were stratified into Basic (14), Enhanced (23), and Full (74) sets based on model complexity requirements and overfitting risk assessment.

3.4 Investigation and Determine Root Cause

The process of investigation within Company MNC involves a systematic and collaborative approach to identifying and resolving issues related to warranty claims. This process includes conducting training and joint inspection activities with car manufacturing company and service centers. Service support is provided to address end-user complaints at OEM service centers through remote website consultations or face-to-face interactions, ensuring prompt and effective resolution.

Company MNC Sales Quality team plays a crucial role in this process by collecting Original Failure Parts (OFP) every two weeks from customers. This collection is essential for verifying the root cause of part replacements and assessing part conditions. The collected OFPs are then segregated and sent to the relevant Company MNC Quality Assurance (QA) plants for further investigation. The lead time for OFP investigation by the QA plant is within 30 days of part receipt, ensuring timely identification and resolution of issues.

The QA team employs Fault Tree Analysis (FTA) to finalize the root cause and potential problem causes. In cases where the root cause remains unclear, QA members propose conducting joint verification activities on the actual vehicle. This involves

gathering more vehicle data, understanding the phenomenon, and identifying other contributing factors. The measurement data process utilizes various tools based on the problem's characteristics, such as SDR2 for noise recording, DST3 for vehicle data recording, a multimeter for checking wire harness continuity and resistance, and belt tension tools for measuring belt tension.

To enhance the investigation process, Company MNC forms a collaborative team involving Quality Assurance (QA). This team utilizes several analytical tools, including fishbone diagrams, Fault Tree Analysis (FTA), and comparison data with both compliant and non-compliant vehicles, to identify the root cause of the failure. If a primary quality issue is detected, Company MNC employs creative problem-solving techniques to generate new ideas and strategies. These techniques include brainstorming sessions where all team members freely suggest ideas without criticism, and brainwriting or free association using the 6-3-5 formula, where six participants each write down three ideas within five minutes under the supervision of a leader.

Following these ideation sessions, Company MNC proceeds with recovering OK parts from end-user or company vehicles for comparison with OFPs at QA plants. This comparison process is critical for verifying the root cause and gaining a deeper understanding of the failure mechanism. Once the root cause is identified, Company MNC conducts a feasibility study to determine the viability of conducting component repairs on the identified product. This study ensures that the proposed repair activities align with the expectations of Company A or other OEM customers, ultimately aiming to reduce warranty costs and enhance overall customer satisfaction.

By integrating these comprehensive investigation and resolution processes, Company MNC ensures that warranty issues are promptly and effectively addressed, thereby maintaining high standards of product quality and customer satisfaction. This approach not only mitigates warranty costs but also supports the company's commitment to continuous improvement and excellence in the automotive industry [75].

3.4.1 Customer Feedback Activity

Annually, customers conduct supplier briefings for suppliers to communicate their goals and objectives [142]. As a leading supplier, Company MNC incorporates these activities into the company's Key Performance Indicators (KPIs) to monitor

project progress aligned with customer expectations. At the end of each year, Company MNC evaluates its performance through customer satisfaction feedback, which measures technical service support, product reliability, and the effectiveness of conducted activities.

Cronbach's Alpha is used to measure the reliability and internal consistency of the feedback. For a sample size of 50 (n=50), the table below presents Cronbach's Alpha values for the feedback items across three fiscal years: FY2021, FY2022, and FY2023.

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N-1) \cdot \bar{c}} \tag{3.1}$$

α = Cronbach's alpha coefficient

N = the number of items

$\bar{c}$ = average covariance between item-pairs

$\bar{v}$ = average variance

Table 3.10
Rule of Thumb for Results [143]

Cronbach's Alpha	Internal Consistency
$\boldsymbol{\alpha \geq 0.9}$	Excellent
$\boldsymbol{0.9 > \alpha \geq 0.8}$	Good
$\boldsymbol{0.8 > \alpha \geq 0.7}$	Acceptable
$\boldsymbol{0.7 > \alpha \geq 0.6}$	Questionable
$\boldsymbol{0.6 > \alpha \geq 0.5}$	Poor
$\boldsymbol{0.5 > \alpha}$	Unacceptable

Table 3.11
Cronbach's Alpha for 3 Years

Fiscal Year	Cronbach's Alpha	Interpretation
FY2021	0.91	Excellent reliability
FY2022	0.93	Excellent reliability
FY2023	0.92	Excellent reliability

The performance of four prominent automotive manufacturers Company B, Company A, Company D, and Company A is illustrated over three fiscal years:

FY2021, FY2022, and FY2023. The following analysis links the feedback data with Cronbach's Alpha values as shown at Table 3.12:

Table 3.12
Analysis Links the Feedback Data by Customers

Customer	FY2021 (%)	FY2022 (%)	FY2023 (%)	Trend Analysis
Company B	100	100	92	Consistent high performance with a decline in FY2023 due to quality
Company A	88	88	90	Steady improvement, reflecting effective strategic measures for design issue
Company D	87	92	95	Significant progress due to quality enhancement initiatives
Company C	91	97	99	Exceptional growth, indicating robust quality control measures

a) Reliability Test

The high Cronbach's Alpha values (above 0.90) indicate excellent reliability, suggesting that the feedback items consistently measure the intended constructs (technical service support, product reliability, and effectiveness of activities) across different fiscal years. This high reliability ensures that the feedback results are dependable and repeatable over time.

b) Validity

1. Content Validity: The feedback comprehensively covers all aspects of customer satisfaction, ensuring all relevant dimensions are included.
2. Construct Validity: The high consistency of feedback items supports their ability to measure the theoretical construct of customer satisfaction.
3. Criterion-Related Validity: The feedback results correlate with other measures of customer satisfaction, such as warranty claims and repeat purchases, further validating the instrument.

The Rasch Measurement Model further validates the feedback by ensuring that the feedback responses are invariant across different groups of respondents. This model transforms the raw scores into a linear scale, providing a precise measure of latent traits.

The histogram of Rasch Measure scores demonstrates the distribution and supports the reliability of the feedback instrument [144].

The chart data reflects the performance outcomes of the manufacturers, supported by the high reliability of the feedback instrument. The implementation of advanced strategies by Company MNC, such as quality control measures, real-time data analysis, and cutting-edge manufacturing technologies, has contributed significantly to minimizing warranty claims and enhancing customer satisfaction.

The annual customer satisfaction feedback conducted by Company MNC is a crucial tool for measuring and improving performance. The high Cronbach's Alpha values confirm the feedback's reliability, ensuring that the data collected is consistent and accurate. This project, led by the Sales Engineering Department, incorporates a cost-effectiveness model considering warranty claim costs, operational costs, and component repair costs, ultimately aiming to achieve high customer satisfaction and operational efficiency through the integration of manufacturing intelligence, industrial management, and World Class Maintenance principles.

Consequently, Company MNC has successfully minimized warranty claims, thereby substantially enhancing overall customer satisfaction. This achievement is underscored by the successful completion of the SIRIM audit under certification IATF 16949:2016, which is conducted every three years. The results of the feedback are shown in Figure 3.13. [145].

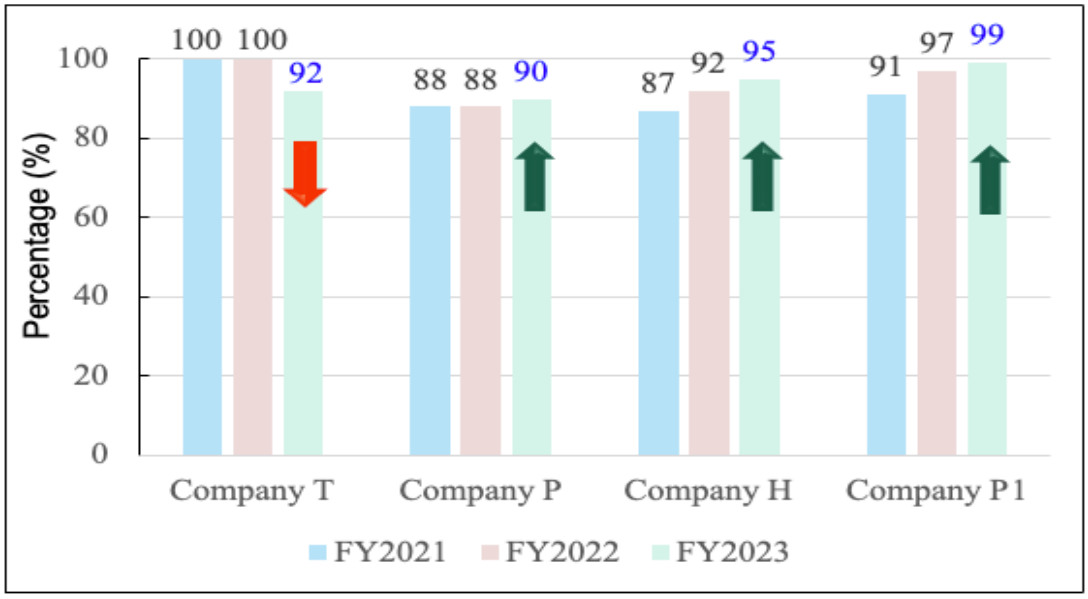


Figure 3.13 Customer Satisfaction Feedback for Market Claim

3.4.2 Conduct Component Repair

Component repair costs encompass component parts costs, profit margins, and taxes, excluding the OEM car assembler/service center's profit margin, administration, handling, labour, and additional taxes. Conducting a feasibility study is crucial for making a GO or NO-GO decision by Top Management regarding this project. Company MNC will evaluate the benefits and disadvantages of undertaking component repair. A risk assessment will also be conducted to avoid any unfavourable outcomes and potential loss of profitability for the company.

The feasibility study will involve team members from various departments, including the Service Department, Aftermarket Department, and OES Department. Representatives from each department will participate in weekly joint meetings to update the activity schedule, which will be presented to the Sales Top Management or Managing Director monthly. During this period, a repair manual process will be provided to train mechanics and technicians. Ultimately, Top Management will decide which business model to utilize based on data and feasibility studies conducted by the team members. The final proposal, once approved by Top Management, will be submitted to Company A for further discussion on the next action and activity schedules.

Training is essential for mechanics and technicians to ensure that the repair quality matches the standards of new Alternator Assembly units. Concurrently, Company MNC aims to build new skills and knowledge among its workforces. Currently, the plan is to collaborate with Quality Assurance (QA) Thermal to produce video tutorials on component repair. These videos will provide clear visualization for mechanics and technicians, enhancing their understanding of the repair process. Given the COVID-19 pandemic, face-to-face training and practical sessions are not feasible. Therefore, video tutorials are considered the best alternative to traditional repair manuals.

Countermeasure monitoring will commence from the implementation date of the repair. Any recurring issues reported by end-users will be addressed immediately, with the Original Failure Part (OFP) recovered for Quality Assurance (QA) plants to investigate the root cause. The repair component will initially be piloted at selected Company A Service Centers for easy monitoring before being expanded to other regions. In case of any issues or complaints, Company MNC will provide technical

support through remote diagnosis via platforms such as TEAMS, WhatsApp, or ZOOM due to the ongoing COVID-19 pandemic. During these sessions, Company MNC will involve specialists, including QA members and parent company, to ensure quick resolution of any field issues.

3.5 Warranty claim analysis

The analysis of warranty claim data, as depicted in the chart, provides a comprehensive overview of the warranty claims over five fiscal years, from FY18 to FY23. The data, analyzed using Excel, reveals both the amount and quantity of warranty claims. The fiscal year FY18 saw warranty claims amounting to RM616K with a quantity of 4690 pieces. In FY19, the claims slightly increased to RM648K, although the quantity marginally decreased to 4607 pieces. FY20 experienced a minor reduction in claims amounting to RM643K and a significant decline in quantity to 3452 pieces. The trend continued in FY21 with claims amounting to RM504K and 2977 pieces. FY22 saw a further reduction to RM445K and 2247 pieces, and FY23 recorded the lowest figures, with RM388K in claims and 1909 pieces as shown in Figure 3.14.

To identify the focus products based on the amount, a detailed analysis highlighted several key components. The Alternator Assembly emerged as a significant cost factor, alongside the EFI ECU and Motor Cooling Fan. The pie chart indicates in Figure 3.15 that the highest warranty claim amount, RM345K, is attributed to Company A, followed by Company B (Domestic) with RM131K, Company B (OVERSEAS) with RM91K, and others with RM40K.

Similarly, the focus products based on quantity reveal critical components contributing to warranty claims. The Alternator Assembly again surfaced as a prominent issue, along with the EFI ECU and Motor Cooling Fan. The pie chart shows that the highest quantity of claims, 2850 pieces, is associated with Company B (Domestic), followed by Company A with 2662 pieces, Company B (Oversea) with 314 pieces, and others with 273 pieces as shown in Figure 3.16.

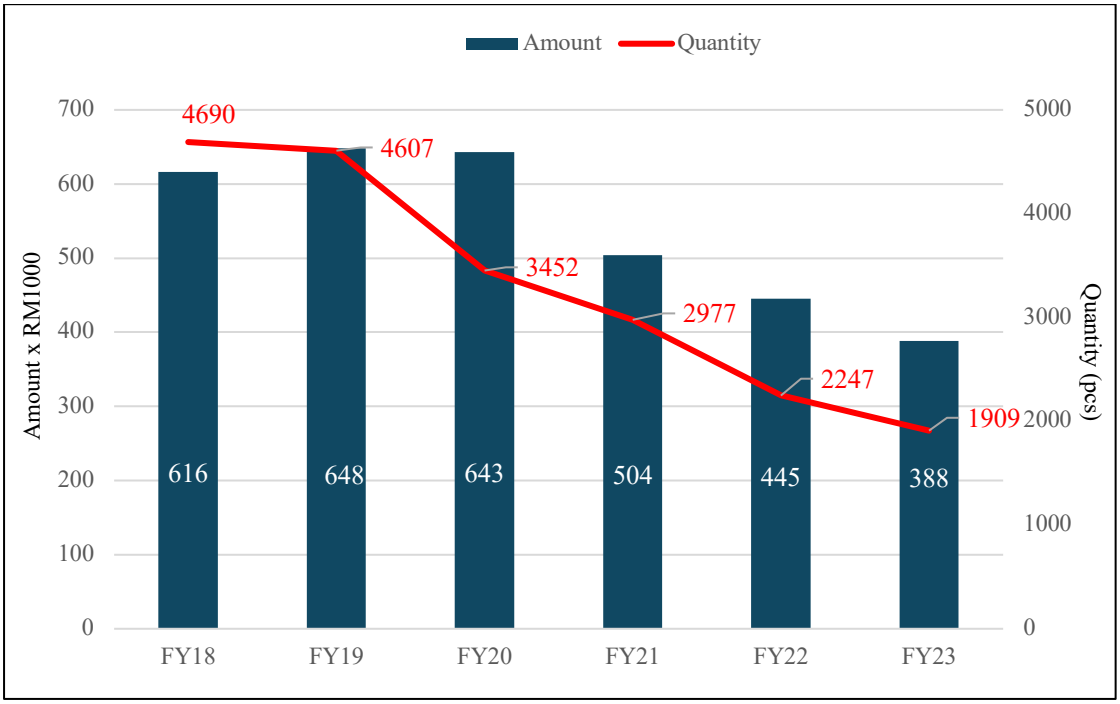


Figure 3.14 Warranty Claim Amount vs Quantity From FY18 ~ FY23

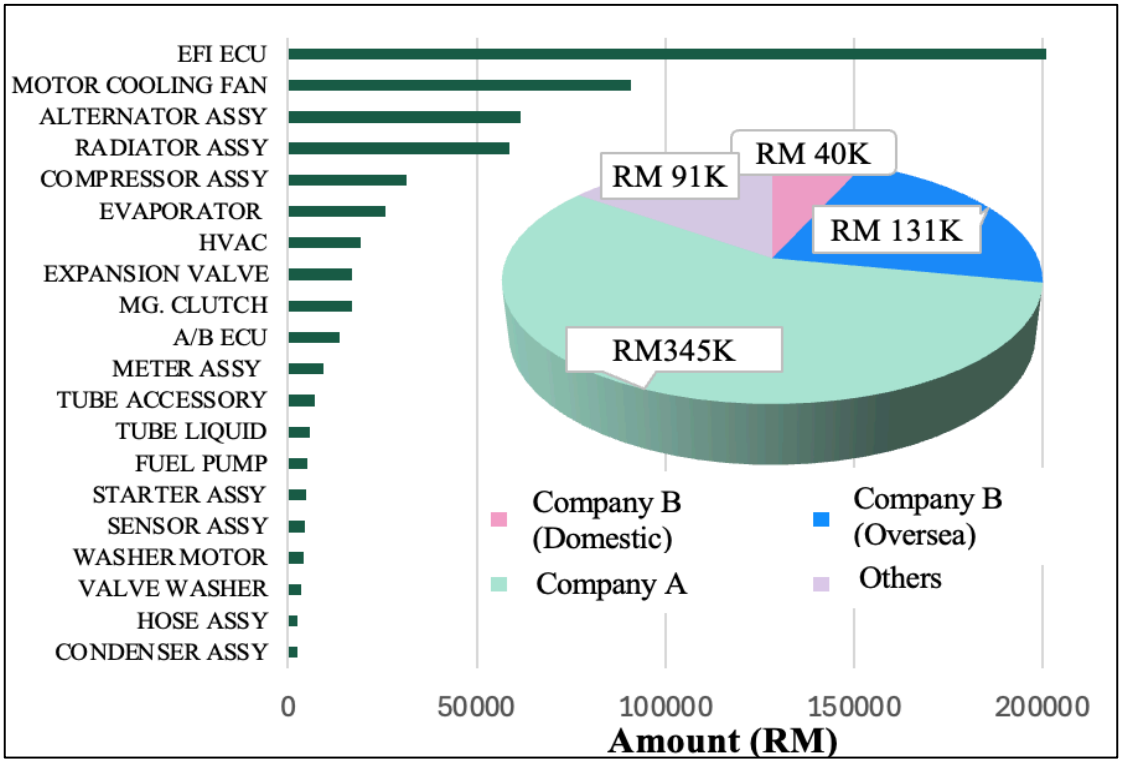


Figure 3.15 Identify Focus Product Based on Amount

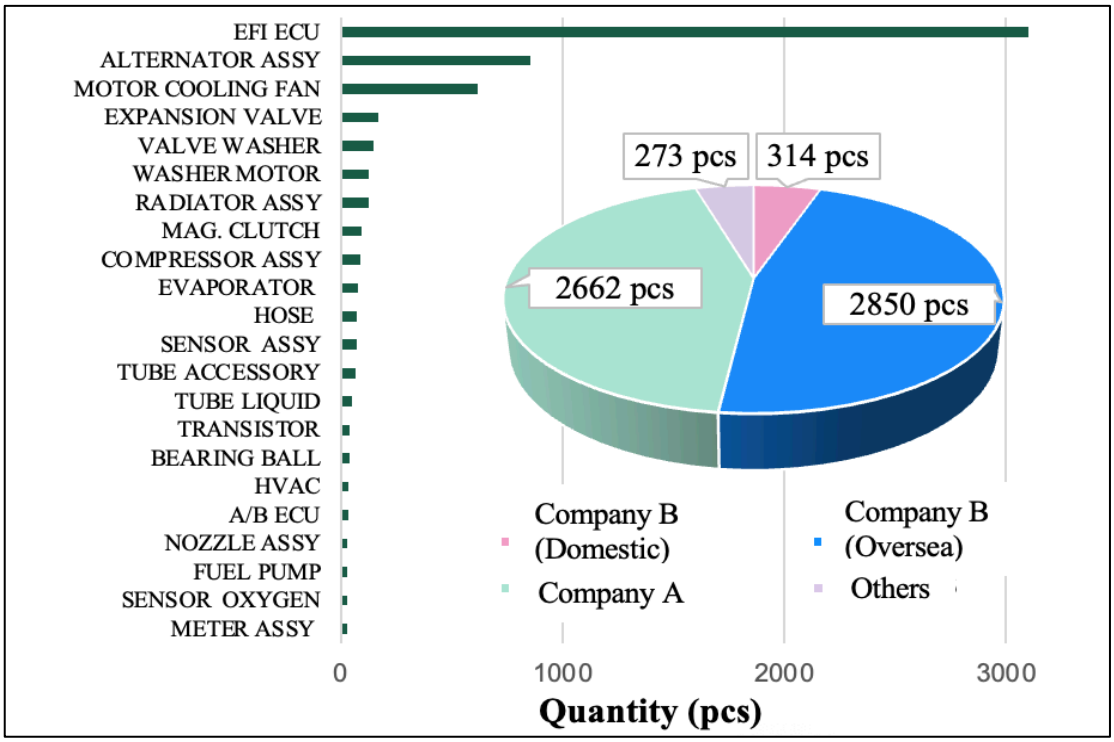


Figure 3.16 Identify Focus Product Based on Quantity

3.6 Develop a Predictive Warranty System Model

In the contemporary automotive industry, big data plays a pivotal role in enhancing warranty management strategies. The extensive utilization of warranty claim data from fiscal years 2018 to 2022 exemplifies this influence. However, the current methods employed for data processing are time-consuming and often lack accuracy, necessitating a more efficient approach. Forecasting analysis emerges as a significant solution, offering the capability to predict and pre-empt vehicle issues. By handling extensive data compilations for all customers, this method not only streamlines the analysis process but also enhances the reliability of the outcomes.

Furthermore, the integration of big data analytics leads to improved cost efficiency and elevated levels of customer satisfaction through insightful analytics. For instance, analyzing the Total Industry Volume (TIV) actual versus forecast, along with Original Equipment Manufacturer (OEM) sales within the domestic market, provides valuable insights into market performance. This, coupled with detailed assessments of warranty costs across fiscal years, underscores the potential for substantial financial optimization and strategic planning.

Adapting to market trends and securing future business success through data-driven approaches is imperative. The deployment of sample dashboards facilitates comprehensive data analysis, enabling organizations to respond dynamically to market demands. These dashboards exemplify the practical application of big data in formulating strategic decisions, thereby affirming the uniqueness and necessity of developing robust data systems in the automotive sector.

Through comprehensive data analysis and the development of sophisticated systems, the automotive industry can significantly enhance its warranty management, forecasting capabilities, cost efficiency, and customer satisfaction, thereby ensuring sustained business success.

3.7 Validate The Model for Component Repair

Upon part recovery activities, the investigation revealed an abnormal noise originating from the front bearing. Detailed analysis identified bearing failure due to ball flaking and internal electrical pitting. This issue was evident through various symptoms such as unsmooth rotation of the front bearing compared to the smooth rotation of the rear bearing. Electrical pitting was observed on the inner ring raceway, and flaking, along with electrical pitting, was found on the ball surface [75]. Additionally, a white structure had been detected at the flaking site, indicating significant material degradation as Figure 3.17.

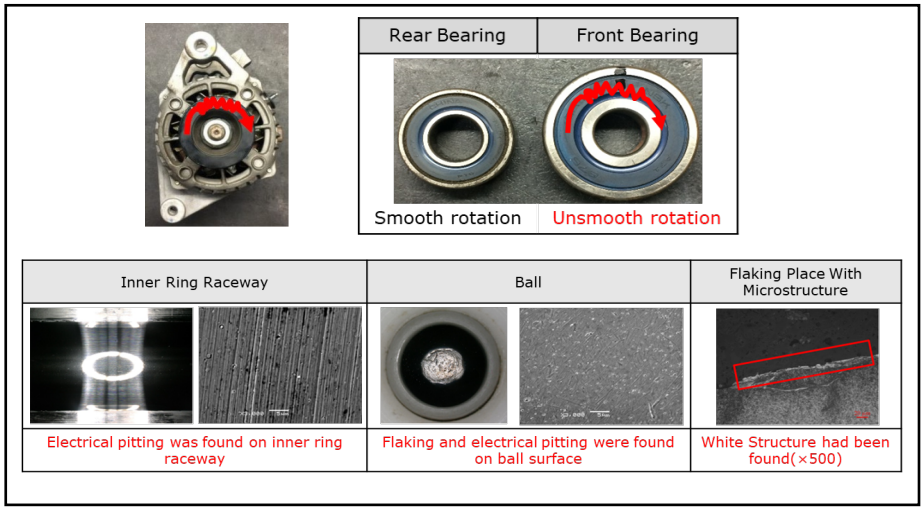


Figure 3.17 Investigation Results Bearing Failure

The investigation employed the Alternator Failure Mode and Effects Analysis (FMEA) and the Fault Tree Analysis (FTA) to determine the root cause of the failure. The FMEA table outlined several potential failure modes, such as abnormal bearing side conditions, lubrication irregularities, dimensional discrepancies, and over-stress conditions. The analysis revealed that issues such as non-conforming outer ring hardness, improper grease amount, and radial clearance beyond specifications contributed to the bearing failure as Figure 3.18. These findings necessitate further on-site verification to confirm the root cause under actual vehicle conditions.

•• PROTECTED

PROCESS FAILURE MODE AND EFFECTS ANALYSIS (P-FMEA)

SHEET _1_ OF _22_

ISSUE	DETAIL	DATE	AUTHOR

PART NAME : ALTERNATOR ASSY,W/REGULATOR
PART NUMBER : 27060-F2170 (TG104211-500*)

PROCESS FUNCTION REQUIREMENTS	POTENTIAL FAILURE MODE	POTENTIAL EFFECTS OF FAILURE	S	CLASS	POTENTIAL CAUSES MECHANISMS OF FAILURE	O	CURRENT PROCESS CONTROL	D	RPN	RECOMMENDED ACTION	RESPONSIBILITY & TARGET COMPLETION DATE	ACTION RESULTS				
												ACTIONS TAKEN	S	O	D	RPN
DR.Bearing press Fitting - Assemly DR Bearing into Frame drive end.by pressing	Dr.Frame mix model	-Cannot intall to Engine	3		-Miss model DR Frame suply by outline (Human Error)	1	- Spl.Jig final detection 100%	2	6							
-Assembly Dr. Bearing with Dr.Frame End correct model	Lathe chips attached to Dr.Frame	-Alt.Noise	1		- Operator does not appearance check part.	1	-Performance test bench check 100% (Hearing noise with master)	2	2							
and press with Dr.Frame End correct model and press to finished position	Set Dr.Frame on Pallet miss position	-Cannot intall to Engine	3		- Camera program error	1	- Spl.Jig final detection 100%	2	6							
	Bearing fitting Not finish position	- Useful life is decreasing. and Bearing noise	3		-ØID of DR Frame too small	1	- Checking bearing height of 1st part before working. - Confirm End position press by Linear	1	3							

SYMBOLS S = SEVERITY CLASS : Pc, S, E, R
O = OCCURRENCE D = DETECTION RPN = RISK PRIORITY NUMBER
O and S : 1 - 5 / Low - High D : 1 -5 / 5 Remote (Low) Possibility of Detection

ROUTING : SUPPLIER ⟶ (COPY) TMT QC

APPROVE	CHECK	ISSUE
Chutintorn 22/07/22 SMTH	Bang-on 22/07/22 SMTH	Napat 22/07/22 SMTH
MANAGER.	ASST. MGR.	QC Engineer
QUALITY CONTROL		

Figure 3.18 Process Failure Mode and Effects Analysis (FMEA) - Bearing

The FTA highlighted the cascading effect of these failures, indicating that high load conditions, vibration, and fluctuating revolutions could exacerbate the bearing failure. It underscored the importance of addressing these root causes to prevent recurrence. The identification of these failure modes and their potential causes is critical for developing effective countermeasures to enhance the durability and reliability of the automotive components as Figure 3.19.

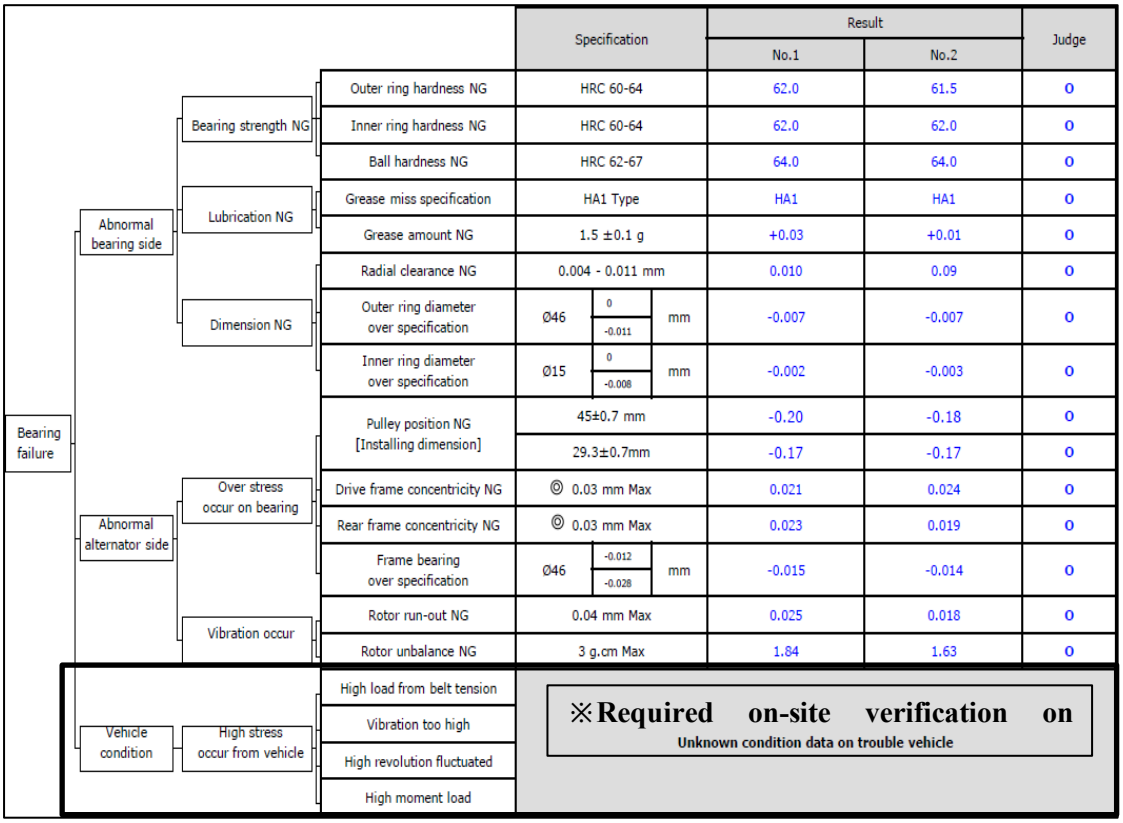


Figure 3.19 Fault Tree Analysis [FTA] – Bearing Failure [75]

In the context of automotive suppliers, companies like Company MNC have implemented stringent quality control measures to minimize warranty claims. By adopting advanced analytical tools and continuous improvement practices, Company MNC has significantly reduced the incidence of such failures. This proactive approach in quality management underscores the importance of integrating thorough analytical frameworks like FMEA and FTA in automotive part analysis to ensure customer satisfaction and reduce warranty costs [146] [147].

In response to persistent vehicle alternator issues, Company MNC initiated a joint verification activity, known as the Go & Study activity. This initiative was conducted in collaboration with the Regional Headquarters, Quality Assurance (QA) plants, and the Company A Warranty Department at the Company A Service Center. The purpose of this collaborative effort is to gather comprehensive information on the affected vehicles and compare them with properly functioning counterparts to identify the phenomenon and root cause of the issues.

During the Go & Study activity, Company MNC received two vehicle complaints from end-users. These complaints were meticulously compared against a good vehicle to ascertain the differences and pinpoint the underlying problems. It is

noteworthy that both vehicles under investigation were still within the warranty period, being less than three years old or having a mileage of under 100,000 kilometers. This adherence to the warranty period is crucial for ensuring the legitimacy of the complaints and the subsequent investigations.

The vehicle owners reported that the vehicle alternators of their Bezza 1.3L models produced abnormal noises after the engine started. This issue warranted a detailed examination to understand the discrepancies between the affected and unaffected vehicles. The Go & Study activity provided a structured approach to systematically collect and analyze vehicle data, facilitating an in-depth understanding of the underlying issues. The structured approach included not only the collection of data but also a comprehensive comparison with a baseline vehicle, which was essential for identifying the specific factors contributing to the abnormal alternator noise.

The results of the vehicle information collection from the vehicle owners are presented in Table 3.13. This table details the specifics of the affected vehicles and facilitates a thorough comparison with the baseline (good vehicle) to aid in identifying the root cause of the alternator noise issue. The data collected included various parameters that were critical in understanding the extent and nature of the alternator problems. By comparing these parameters with those of a properly functioning vehicle, the investigation aimed to isolate the anomalies and understand their implications on the alternator's performance.

Table 3.13
Summary of joint investigation for Alternator abnormal noise cases

Item	Vehicle No.1 (Complained)	Vehicle No.2 (Complained)	Vehicle No.3 (Reference)
Investigation Date	20 January 2020	21 January 2020	22 January 2020
Location	IG Enterprise	Asjita Sdn Bhd	Company MNC
Vehicle Model / Engine	Bezza 1.3L (1NR)	Bezza 1.3L (1NR)	Bezza 1.3L (1NR)
Mileage / MIS	57,926 km / 18.4 MIS	64,199 km / 23.3 MIS	25,362 km / Not available
Vehicle Usage Profile	Sales vehicle; predominantly highway and urban traffic use; operated 7 days/week at speeds of 80–100 km/h; average distance of 200 km/day; single occupant	Ride-hailing vehicle (Grab); predominantly highway and urban traffic use; operated 6 days/week at speeds of 80–110 km/h; average distance of 110 km/day; driver with up to two passengers	Company vehicle; detailed usage information not available
Observed Complaint	Intermittent abnormal noise originating from alternator	Intermittent abnormal noise originating from alternator	No abnormal noise observed
Vehicle Modification Status	No modification reported	Engine belt replaced on 22 July 2019 (3,499 km usage at inspection)	No modification reported

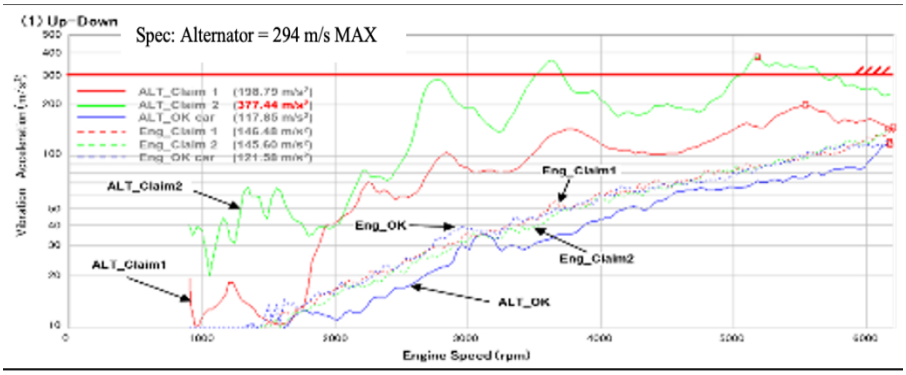
In the course of investigating vehicle alternator issues, Company MNC conducted a thorough collection of vehicle data and information. The team used a belt tension gauge to measure the belt tension on all the collected vehicles, revealing differing tension readings among the tested vehicles. These discrepancies were attributed to the vehicle specifications that utilized an auto belt tension mechanism, resulting in inconsistent readings at various points. Subsequent data collection involved measuring electrical load and vibration using a FFT analyzer. The results indicated that the electrical load and vibration levels in both complaint vehicles were higher than those in a properly functioning vehicle. This data provided crucial insights into the underlying issues affecting the vehicles.

Further validation was performed by downloading data from the FFT analyzer's data logger, confirming that high vibration levels were causing abnormal alternator noise. Detailed data collection and measurements are presented in Table 3.14, while Company MNC also estimated failure analysis using Fault Tree Analysis (FTA) as depicted in Figure 3.20. The Go & Study activity data suggested that the abnormality stemmed from the condition of the auto belt tensioner. To substantiate these findings, additional checks on belt tension in good vehicles in the market and new vehicles from Company A were deemed necessary.

The comprehensive belt tension confirmation involved forty-nine vehicles, comprising twenty-one units from the market and twenty-eight new vehicles from Company A Manufacturing, all of which had passed Quality Control (QC) inspection. The data collection was carried out by Company MNC Quality Assurance (QA) Department. The results, shown in Table 3.15, Figures 3.20, 3.21 and 3.22, indicated that both market and new vehicles exhibited belt tensions exceeding specifications.

Based on data collected from actual vehicles during joint investigations, the primary cause of bearing failure was identified as inconsistent or excessive belt tension, which resulted in abnormal noise. To address end-user complaints and prevent further increases in warranty costs, Company MNC conducted additional studies to improve the front bearing design and implemented field fix actions through component repair. These initiatives were aimed at reducing warranty expenditure while improving product reliability and overall business profitability.

Table 3.14
Comparison of Belt Tension, Electrical Load, and Vibration Characteristics

Item	Vehicle No.1 (Complained)	Vehicle No.2 (Complained)	Vehicle No.3 (Reference)
Belt Tension (Drawing: 238 ± 34 N) Specification : 204–272 N	253–257 N	179–190 N	315–327 N
	Difference in tension range observed despite automatic belt tension system		
Electrical Load Idle 700 rpm 3000 rpm	**No Load:** 14.06 V, 14.00 A 14.16 V, 14.16A	**No Load:** 14.14 V, 11.20 A 14.24 V, 11.50A	**No Load:** 14.09 V, 9.80 A 14.19 V, 10.40 A
	Load: 13.60 V, 67.8A 13.67 V, 69.5 A	**Load:** 13.69 V, 70.30 A 13.73 V, 71.30 A	**Load:** 13.68 V, 60.80 A 13.75 V, 62.00 A
	Electrical load of complained vehicles is higher than reference vehicle		
Vibration Level Specification Limit: 294 m/s²	ALT Max: 201.54 m/s² E/G Max: 184.35 m/s²	ALT Max: 377.44 m/s² E/G Max: 208.29 m/s²	ALT Max: 117.98 m/s²E/G Max: 147.70 m/s²
Pulley Load – Parking Condition: Maximum 1850 N	Hot start, load 1043 N (at tension 262–272 N)	Hot start, load 757 N (at tension 185–199 N)	Hot start, load 1231 N (at tension 276–298 N)
Results from Data Logger			
	High vibration levels correlated with alternator abnormal noise occurrence		
Vehicle Data Collection			

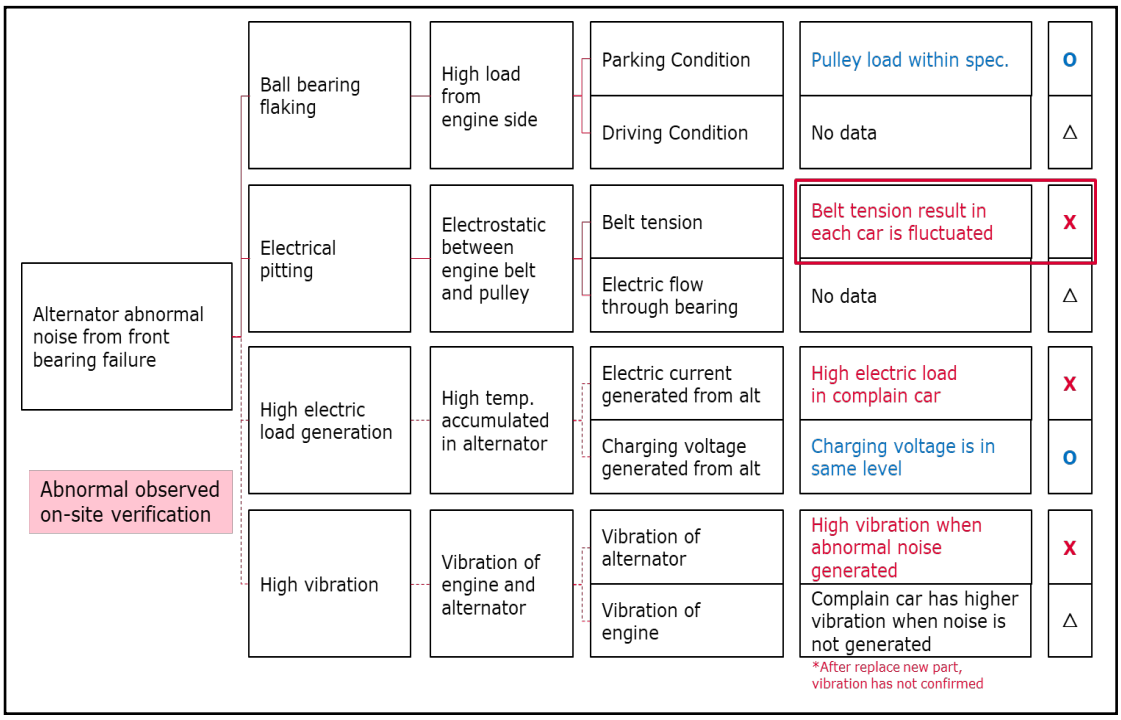
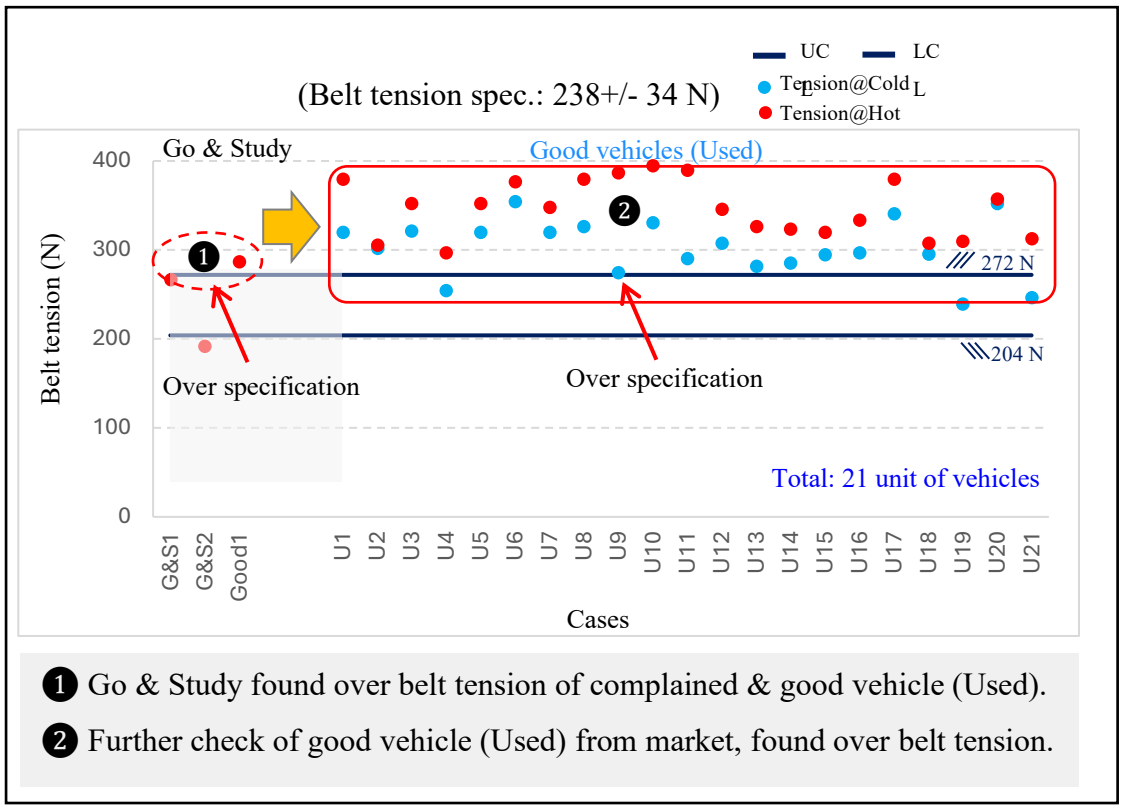


Figure 3.20 Estimate Failure Analysis by Fault Tree Analysis [75]

Table 3.15
Belt Tension Confirmation from Good Vehicle and New Vehicles

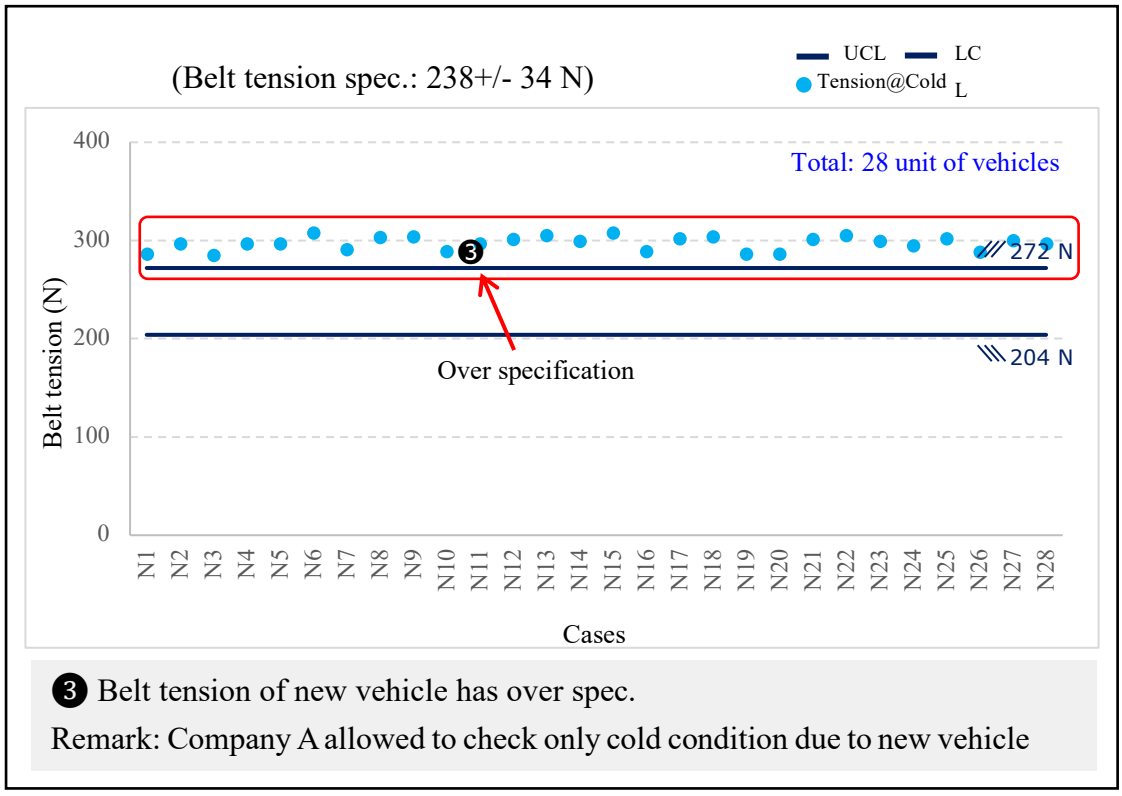
Target	**Good Vehicle (Used)** **(Using in market from End User)**	**New Vehicle** **(After QC passed in Company A)**
Allowed Quantity	Twenty-one (21) Units of vehicle	Twenty-eight (28) units of vehicles
Analyzer	Company MNC Quality Assurance staff	Company A staff (Not allow outside person)
Tools	CLAVIS Belt Tension Meter *Training by parent company with manual	CLAVIS Belt Tension Meter *Training by Company MNC staff with manual



❶ Go & Study found over belt tension of complained & good vehicle (Used).

❷ Further check of good vehicle (Used) from market, found over belt tension.

Figure 3.21 Complained & Good Vehicles (Used) [75]



❸ Belt tension of new vehicle has over spec.

Remark: Company A allowed to check only cold condition due to new vehicle

Figure 3.22 New vehicles after Quality Control check by Company A Manufacturing [75]

3.7.1 Feasibility Study To Conduct Component Repair

3.7.1.1 SWOT analysis

In evaluating the feasibility of component repair for Company MNC a SWOT analysis was conducted to identify the strengths, weaknesses, opportunities, and threats associated with this initiative as shown in Table 3.16.

Company MNC boasts a nationwide network of service dealers that support repair and return activities. This extensive network ensures that customers can access repair services conveniently across various locations. Moreover, Company MNC possesses the expertise to conduct training programs for Original Equipment Manufacturer (OEM) networks, aimed at enhancing the skills of mechanics. This capability is crucial for maintaining high repair standards and ensuring the reliability of the repaired components.

Despite these strengths, there are notable weaknesses in the component repair initiative. The Company MNC service dealer network is required to maintain a stock of component parts to avoid prolonged waiting times for end-users. This requirement can be challenging to manage and may lead to logistical complexities. Additionally, the introduction of component repair might reduce Company MNC aftermarket sales, as the demand for complete assemblies could decline. This potential decrease in sales must be carefully considered when implementing the repair strategy.

The component repair initiative presents significant opportunities for Company MNC including the potential reduction in warranty claim costs for both Company MNC and its OEM customers. By effectively repairing components, Company MNC can minimize the number of claims, thereby enhancing customer satisfaction and reducing financial liabilities. Furthermore, engaging in cost-sharing negotiations with OEM customers can result in mutually beneficial arrangements, fostering stronger partnerships and creating win-win situations.

However, the initiative also faces certain threats. There is a risk of secondary defects arising from repairs and replacements carried out by mechanics. Ensuring that the repairs are conducted with high precision is essential to avoid recurring issues. Additionally, the potential for high repeating cases due to inadequate repair skills poses a significant threat. It is imperative that the training programs for mechanics are comprehensive and effective to mitigate this risk.

Company MNC extensive service network and commitment to training programs for OEM mechanics play a crucial role in minimizing warranty claims and enhancing component reliability. These initiatives have significantly contributed to the company's ability to reduce costs and improve customer satisfaction. The findings from this SWOT analysis underscore the importance of leveraging Company MNC strengths and opportunities while addressing its weaknesses and threats to implement a successful component repair strategy.

Table 3.16
SWOT analysis

STRENGTHS	WEAKNESSES
• Company MNC have nationwide Service Dealer for support repair and return activity • Company MNC have expertise to conduct Training to OEM network for skill up Mechanic	• Company MNC Service Dealer network required to keep and maintain component part in stock to avoid vehicle owner longer time waiting • Company MNC Aftermarket Sales will reduce after introduce component repair due to low order for complete assembly
OPPORTUNITIES	**THREATS**
• Reduce warranty claim cost for Company MNC and OEM customers • Lowest cost sharing negotiation with OEM customers as win-win situation	• Secondary defect after repair and replacement done by Company A Service Center • A high recurrence of cases may result in reduced vehicle owner retention and loss of customer trust

3.7.1.2 Business Model Canvas (BMC) analysis

Company MNC has developed a comprehensive business model to ensure its strategic initiatives are effectively commercialized and that its business remains viable in the competitive market. This model is detailed through the Business Model Canvas, which encompasses key partnerships, activities, resources, value propositions, customer relationships, channels, customer segments, cost structure, and revenue streams as shown in Table 3.17.

Company MNC collaborates with several strategic partners to support its business operations. These partnerships include manufacturing plants and parent company. These alliances are vital for fostering innovation, ensuring high production quality, and facilitating market expansion. The collaboration with R&D centers enables

Company MNC to stay at the forefront of technological advancements and maintain a competitive edge.

The company engages in several critical activities to deliver maximum value to its customers. These activities include conducting Original Equipment Manufacturer (OEM) feedback, part recovery, and part investigation. Through OEM feedback, Company MNC gathers valuable insights into customer needs and market trends. Part recovery and investigation allow the Company Bo to identify and address issues promptly, ensuring the reliability and performance of its products. These activities are integral to Company MNC commitment to quality and customer satisfaction.

Company MNC operations rely on several key resources essential for executing its business activities. These resources include warranty data, engineering drawings, budget allocations, and skilled sales engineers. Warranty data provides a historical perspective on product performance, helping predict future trends and prevent potential issues. Engineering drawings and budgets ensure that product designs meet high standards and that projects are adequately funded. Skilled sales engineers bridge the gap between Company MNC and its customers, ensuring that customer needs are understood and met effectively.

Company MNC offers several compelling value propositions to its customers. These include the analysis of warranty data via Power BI, which enables data-driven decision-making and proactive problem resolution. The company also provides component repair and returns services, backed by a robust support network. Additionally, Company MNC focuses on offering low-cost solutions that save time while maintaining high quality, adhering to OEM standards. These value propositions are designed to enhance customer satisfaction and loyalty by delivering reliable and cost-effective solutions.

Maintaining strong customer relationships is a priority for Company MNC. The company engages in various initiatives to achieve this, such as organizing OEM conferences and systems, issuing technical bulletins, and maintaining a comprehensive dealer system. These efforts ensure that customers are well-informed, supported, and connected to the latest developments and services offered by Company MNC. By fostering strong relationships, Company MNC builds trust and long-term loyalty among its customer bases.

Company MNC utilizes multiple channels to reach its customers and deliver its value propositions. These channels include OEM part centers, service dealers, and part

stockists. By leveraging these channels, Company MNC ensures that its products and services are accessible to customers, enhancing convenience and reliability. The strategic use of these channels allows Company MNC to effectively distribute its offerings and support its customers across various locations.

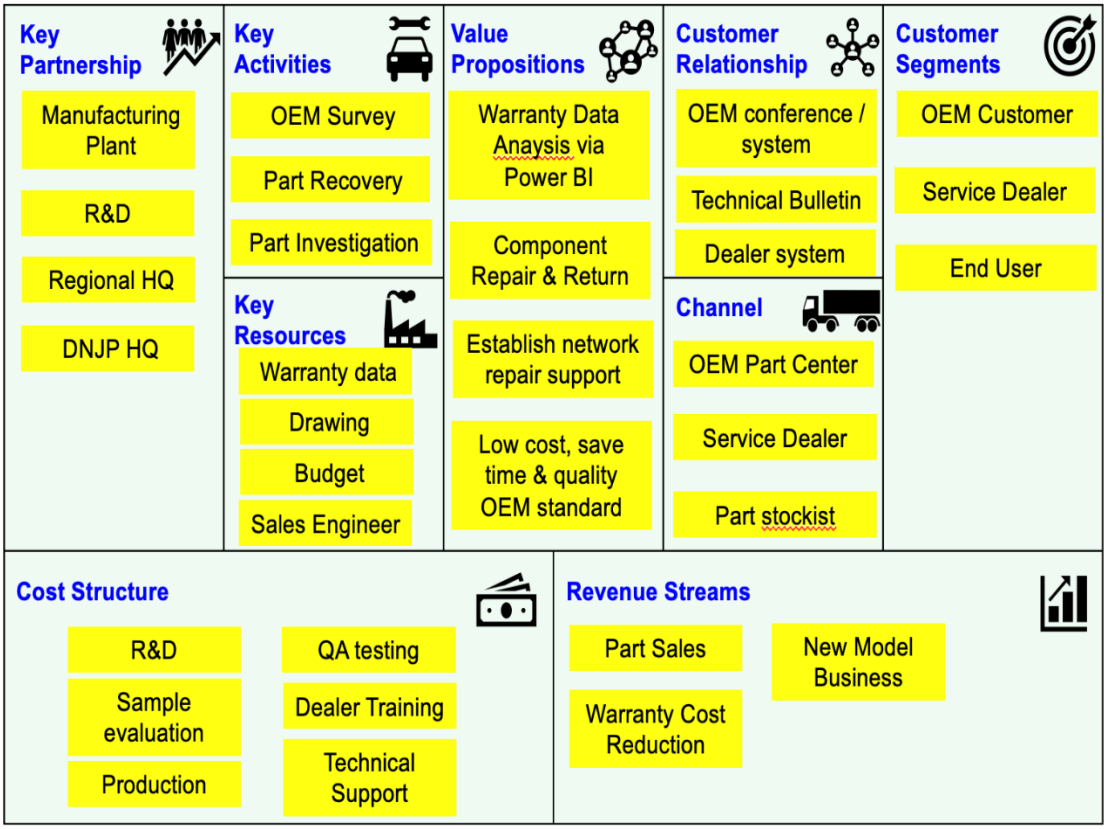
The primary customer segments for Company MNC include OEM customers, service dealers, and end-users. Each segment has distinct needs and expectations, and Company MNC tailors its offerings to meet these requirements effectively. By understanding and addressing the specific needs of each customer segment, Company MNC can deliver targeted solutions that enhance customer satisfaction and loyalty.

The cost structure of Company MNC encompasses various essential activities and resources. These include R&D, sample evaluation, production, quality assurance testing, dealer training, and technical support. By investing in these areas, Company MNC ensures that it maintains high standards of quality and innovation. The allocation of resources to these critical activities enables Company MNC to deliver reliable and advanced products to its customers.

Company MNC generates revenue through several streams. These include part sales, new model business, and warranty cost reduction. By diversifying its revenue streams, Company MNC ensures financial stability and growth potential. Part sales provide a steady income, while new model business drives innovation and market expansion. Warranty cost reduction initiatives help improve profitability by minimizing the costs associated with warranty claims.

In conclusion, the Business Model Canvas for Company MNC outlines a comprehensive and strategic approach to maintaining business viability and achieving market success. By leveraging strong partnerships, engaging in key activities, utilizing essential resources, offering valuable propositions, maintaining robust customer relationships, utilizing effective channels, addressing specific customer segments, and managing costs and revenue streams efficiently, Company MNC can continue to meet customer needs, drive innovation, and achieve sustainable growth.

Table 3.17
Business Model Canvas Analysis



3.8 Process of Component Repair: Comparison of Repair Concepts

In evaluating the feasibility of different repair concepts for component repair, Company MNC conducted a comparative analysis of three options: utilizing the OEM network, establishing a centralized repair shop, and leveraging the Company MNC service dealer network.

The first option involves utilizing the OEM network for immediate repair and replacement of components. This approach emphasizes prompt action to address component failures, ensuring minimal downtime for the end-user. The primary responsibility lies with the OEM network, which can swiftly replace defective parts with new ones. This option is deemed feasible and beneficial due to its ability to provide quick solutions and maintain customer satisfaction. Therefore, it has been recommended for a detailed feasibility study.

The second option proposes the establishment of a centralized repair shop by Company MNC. In this model, cores are collected from the OEM network, followed by a full component repair, after which the repaired components are returned to the

customers. However, this approach has notable drawbacks. The profit margins are relatively low, and there is a high potential for secondary defects during the repair process. Additionally, the lead time for repairs is considerably long, which can negatively impact customer satisfaction. Given these challenges, this option has been evaluated as not suitable for proceeding with further feasibility studies.

The third option involves leveraging the Company MNC service dealer network, where service dealers send the cores for repair and exchange them with repair units. Despite its potential, this option faces several significant hurdles. The service dealer network lacks experience in component repair, and there is insufficient capital to maintain a stock of component parts. Moreover, similar to the centralized repair shop, the lead time for repairs remains long, further diminishing its viability. Consequently, this option has also been deemed unsuitable for further feasibility studies. By considering these options and their respective advantages and drawbacks, Company MNC aims to implement the most effective and efficient repair strategy.

CHAPTER 4
RESULT AND DISCUSSION

4.1 Introduction

This chapter reports the analytical findings and model evaluation outcomes for the development of a predictive warranty management approach at Company MNC, with emphasis on alternator assembly claims for Company A. Building on the earlier methodological framework, the analysis demonstrates how structured warranty data can be transformed into actionable evidence to support lower-cost repair strategies, particularly through component-level (component parts) replacement rather than full assembly replacement. This direction is consistent with Company MNC's Alternator component repair implementation, which aims to restore product function while reducing claim cost exposure and material wastage. Since the introduction of Alternator component repair activities in FY2024, warranty claim costs have shown a declining trend, suggesting early effectiveness of the repair-focused strategy. Moving forward, the same data-driven approach will be extended to other high-impact products, supported by accelerated root cause investigation and scalable big-data analytics to improve forecasting accuracy and decision speed.

4.2 Development of Novel Big Data Analysis

Company MNC plans to deploy the Automotive Warranty Analysis System (AWAS) in April 2025 as a centralized platform for advanced warranty analytics and decision support. The system leverages Microsoft Power BI to transform large-scale warranty datasets into structured, interactive, and actionable insights. AWAS consolidates warranty information such as claim period, customer details, part name, part number, and vehicle model, enabling efficient filtering, traceability, and multi-dimensional analysis. This approach reflects modern manufacturing intelligence principles, where operational data is continuously analysed to support proactive quality and cost management.

4.2.1 Evolution of Warranty Management

Warranty management within the automotive sector has undergone a substantial methodological transformation, progressing from basic record-keeping functions to advanced analytical systems that support forward-looking decision-making. As illustrated in Figure 4.1, this evolution reflects the increasing integration of data analytics, algorithmic modelling, and digital diagnostic technologies into warranty operations.

Historically, warranty activities were dominated by claim reporting and cost recording after component failure had already occurred. Such practices limited manufacturers to corrective actions rather than preventive measures, often leading to prolonged service cycles, escalating warranty expenditures, and reduced customer confidence. The absence of analytical foresight constrained effective planning and weakened the ability to respond efficiently to emerging quality issues.

The current generation of warranty management frameworks adopts a fundamentally different approach by embedding continuous data monitoring and intelligent analytics into the decision process. Real-time diagnostic inputs, combined with machine learning techniques, enable the early recognition of abnormal failure behaviour and emerging defect trends. This capability allows manufacturers to initiate timely interventions, adjust service strategies, and mitigate warranty exposure before failures propagate widely. AI-based fault identification methods have been shown to support predictive maintenance planning and improve reliability outcomes by shifting warranty management from reaction-driven processes to anticipation-driven systems [148].

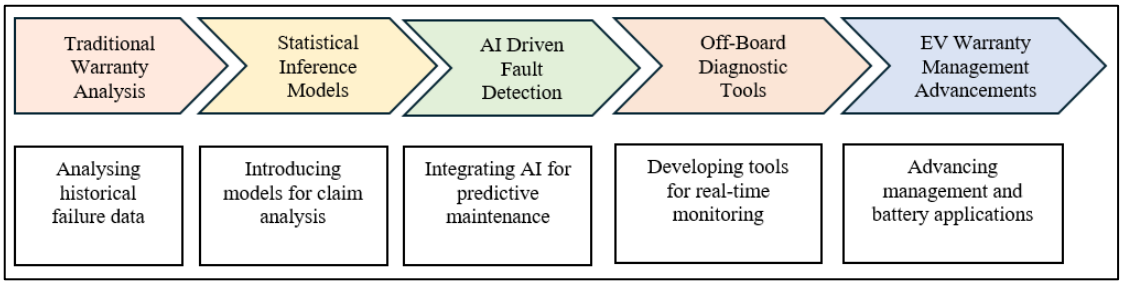


Figure 4.1 Evolution of Warranty and Management

The increasing adoption of electric vehicles and electronically complex architectures has further amplified the need for adaptive warranty analytics. Failure mechanisms related to battery ageing, thermal loads, and extended duty cycles introduce new sources of uncertainty that are not adequately addressed by traditional warranty models. Consequently, the deployment of AI-enabled warranty intelligence has become essential for managing evolving risk profiles and ensuring sustainable warranty strategies that align with rapid technological change and long-term industry transformation.

4.2.2 Role of AI and Business Intelligence in Warranty Analysis

The increasing availability of digital warranty data has expanded the role of Artificial Intelligence (AI) and Business Intelligence (BI) in supporting warranty analysis within the automotive industry. AI based analytical techniques enable continuous evaluation of warranty claim behaviour by identifying abnormal claim trends, repeated failure patterns, and unusual cost distributions. Through machine learning and statistical analysis, the system is capable of detecting potential reliability issues earlier, allowing corrective actions to be implemented before the issues develop into large scale field failures. This approach improves warranty monitoring effectiveness and supports more structured quality improvement activities.

Predictive analytics further strengthens warranty management by linking historical warranty records with operational variables such as production volume, supplier quality performance, product modifications, and service claim trends. The integration of these variables enables the forecasting of future warranty risks and supports proactive planning for service operations, spare parts preparation, and warranty cost control. Compared to conventional reporting methods, predictive models provide better visibility of future claim behaviour and improve decision making capability under changing operational conditions.

The analytical outputs generated through AI models are further enhanced through BI based visualisation platforms such as Microsoft Power BI. Interactive dashboards enable engineers and management personnel to monitor warranty performance, claim trends, root cause distribution, and operational indicators in real time. The integration of AI and BI allows complex analytical outputs to be translated

into practical and understandable information, improving response speed and supporting evidence based decision making.

In addition, the implementation of AI and BI within the AWAS framework was developed in accordance with structured industrial processes and data preparation procedures to ensure analytical consistency and reliability. The preparation stage involved warranty data extraction, data cleansing, duplicate removal, standardisation of failure coding, and validation of historical claim records before analytical processing was conducted. These preparation activities were necessary to reduce data inconsistency and improve the reliability of forecasting and visualisation results.

The implementation process also considered standard industrial quality practices such as failure trend monitoring, claim verification procedures, root cause analysis, and continuous improvement activities. The integration of AI and BI technologies within these operational processes enables more systematic warranty governance and improves coordination between engineering, quality assurance, and management functions. Beyond forecasting capability, the proposed system also contributes to operational efficiency through faster data processing, improved reporting transparency, reduction of manual analysis activities, and enhanced traceability of warranty information. These contributions strengthen the practical value of the AWAS framework and support long term quality improvement and cost optimisation within the automotive warranty management process.

4.2.3 Overview of Dashboard Development

This project develops a data-driven warranty analytics dashboard using Microsoft Power BI as the primary visualization platform. The dashboard provides an intuitive and comprehensive interface for analysing warranty performance, component reliability, and claim patterns. the system integrates Company MNC's warranty claim data and presents high-level Key Performance Indicators (KPIs), including total claims, claim rate, average claim cost, and claim duration. Interactive filtering is enabled across multiple dimensions such as claim year, vehicle model, dealer code, and component category, allowing users to drill down into specific trends and failure mechanisms.

Advanced data modelling is implemented through relational table design using primary and foreign keys, enabling granular comparison across vehicles, parts, and time periods. Calculated measures and DAX expressions ensure consistent and automated

KPI computation. The dashboard is optimized for desktop and mobile access, ensuring usability for managers, engineers, and analysts. Row-Level Security (RLS) is implemented to protect sensitive information and enforce role-based access control. Additionally, anomaly detection features highlight unusual claim behaviour, enabling early investigation before issues escalate into systemic failures.

Overall, the dashboard fulfils the project objectives by:

a) Enabling rapid analysis of large warranty datasets,
b) Establishing a unified Power BI-based warranty analytics platform, and
c) Delivering actionable, user-friendly visual insights for improved decision-making.

4.2.4 Data Selection and Preparation

An extensive data cleansing and preprocessing phase was conducted to ensure data accuracy, consistency, and reliability. This included removal of duplicate records, resolution of formatting inconsistencies, handling of missing values, and standardization of categorical fields.

Key data fields selected for analysis include:

a) Claim Year and Month
b) Dealer Code and Country Code
c) Vehicle Model and VIN
d) Registration, L/O, and Repair Dates
e) Mileage or Operating Hours
f) Failed Part and Cause Description
g) Claim Quantity and Total Claim Amount

These fields are critical for tracking warranty incidents, identifying failure causes, and assessing cost impact. Relationships between datasets were established using identifiers such as VIN and Claim Year, enabling cross-referencing and real-time filtering as shown in Figure 4.2, 4.3 and 4.4.

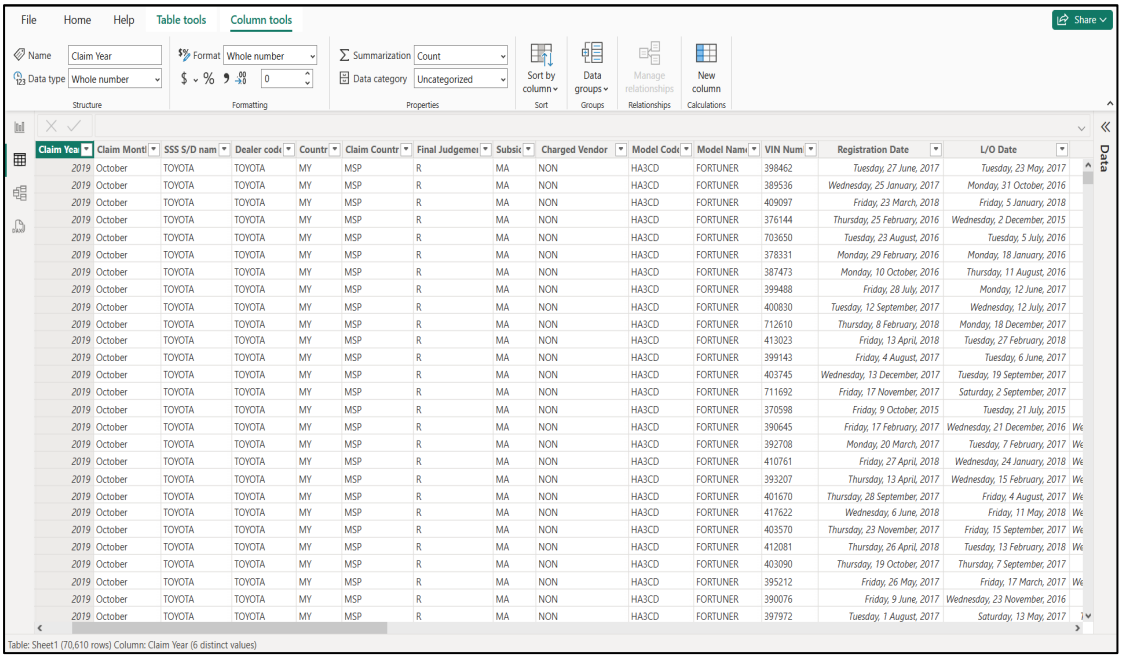


Figure 4.2 Warranty Data Selection For Dealer Code

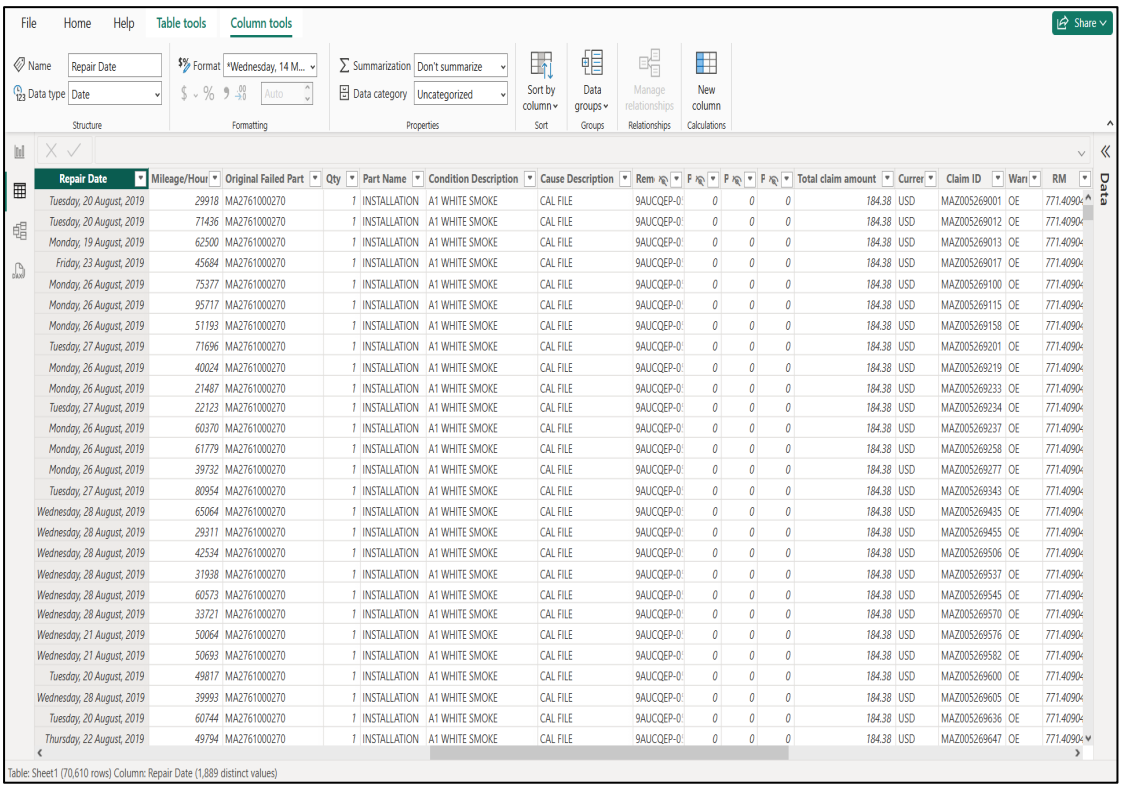


Figure 4.3 Warranty Data Selection for Cause Description

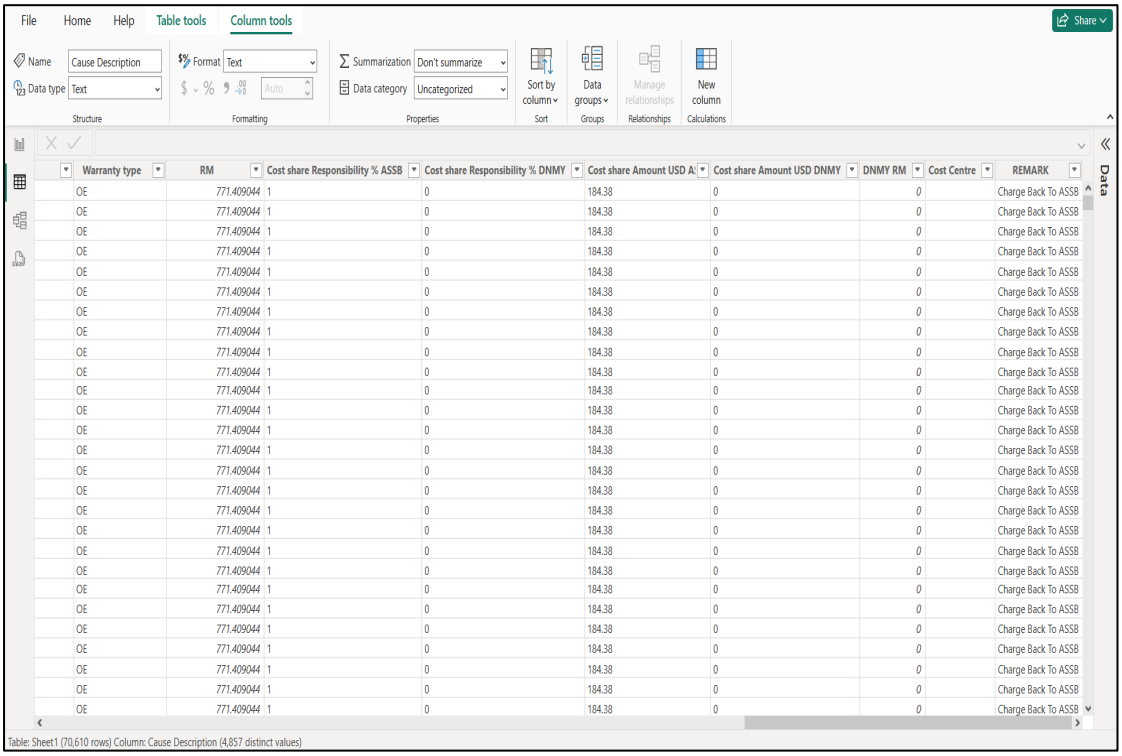


Figure 4.4 Warranty Data Selection for Claimed (Year/Month)

4.2.4.1 Time-Based Warranty Trend Analysis

A time-based warranty trend analysis was conducted using a combined bar and line chart, where L/O Date counts are shown as bars and Repair Date counts as a line series as shown in Figure 4.5. Data spans from Q2 2019 to Q1 2024, enabling identification of post-delivery quality trends.

A pronounced peak is observed in Q3–Q4 2019, indicating a concentration of warranty repairs that may be linked to production or component issues. Claim volumes decline after early 2020, suggesting the effectiveness of corrective actions. The chart also highlights the lag between manufacturing and repair, supporting optimization of warranty coverage planning and preventive actions. Although not a classical Pareto chart, it reflects Pareto principles, where a limited number of periods account for a large share of warranty issues.

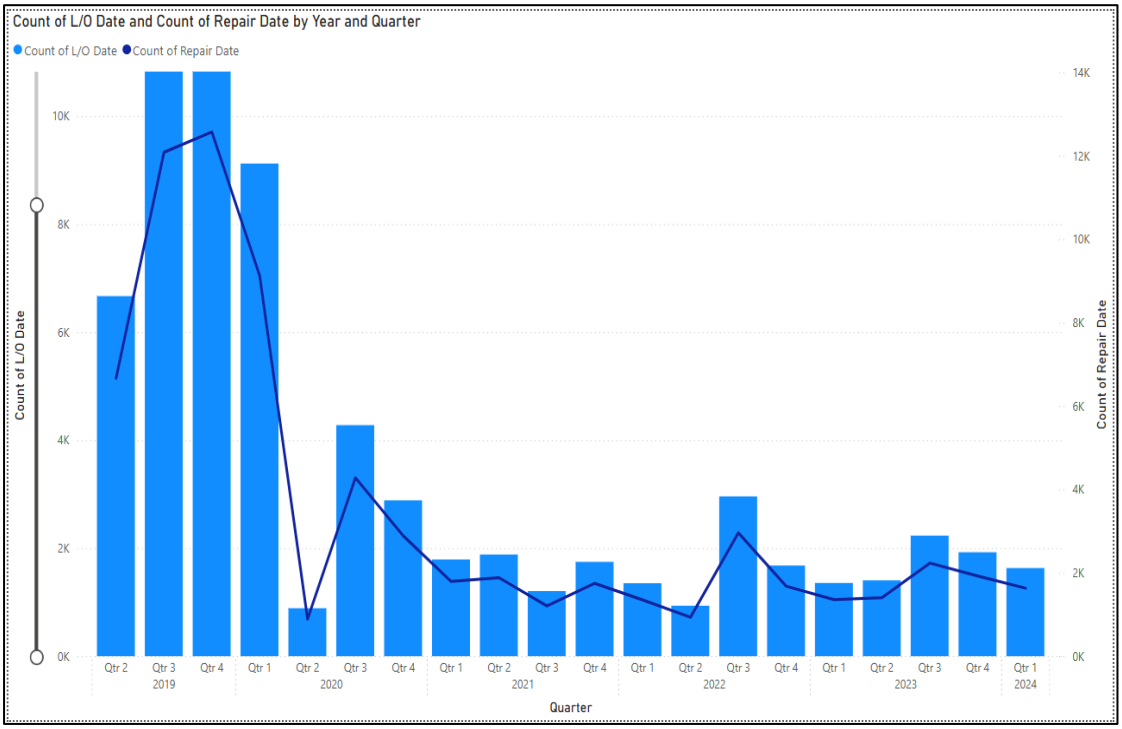


Figure 4.5 Count of L/O Date and Repair Date

4.2.4.2 Model Distribution of Warranty Claims

Figure 4.6 illustrates warranty claim distribution by vehicle model using a donut chart. HILUX (35.65%), FORTUNER (20.42%), and NEW MYVI (14.05%) contribute the highest claim proportions. This insight enables targeted quality investigations and prioritized resource allocation for high-risk models. The donut chart effectively communicates proportional contributions while maintaining clarity, supporting data-driven decision-making aligned with continuous improvement objectives.

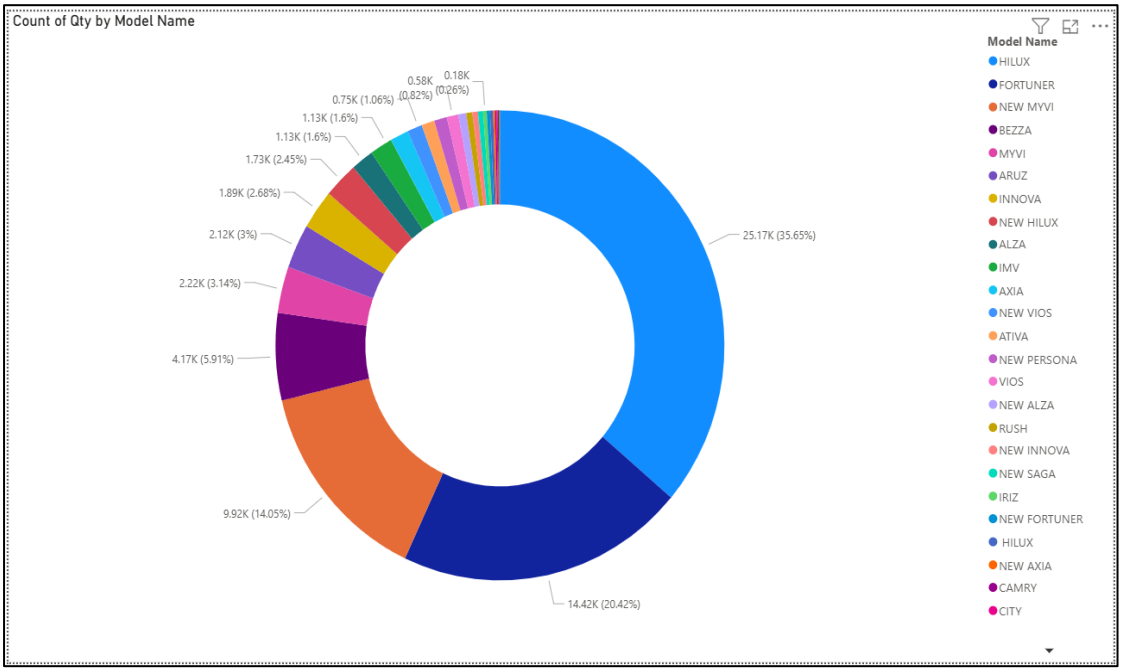


Figure 4.6 Count of Quantity by Model Name

4.2.4.3 Part Name Failure Analysis

Part-level failure analysis is critical for identifying components that most significantly affect system reliability. Figure 4.7 presents a bar chart of claim quantities by part name, enabling direct comparison across multiple categories. Bar charts are particularly suitable for ranking failure-prone components in complex systems and identifying outliers for focused investigation [133], [149]. This analysis supports quality control, predictive maintenance planning, inventory optimization, and supplier management, ensuring improvement efforts are concentrated on high-impact components.

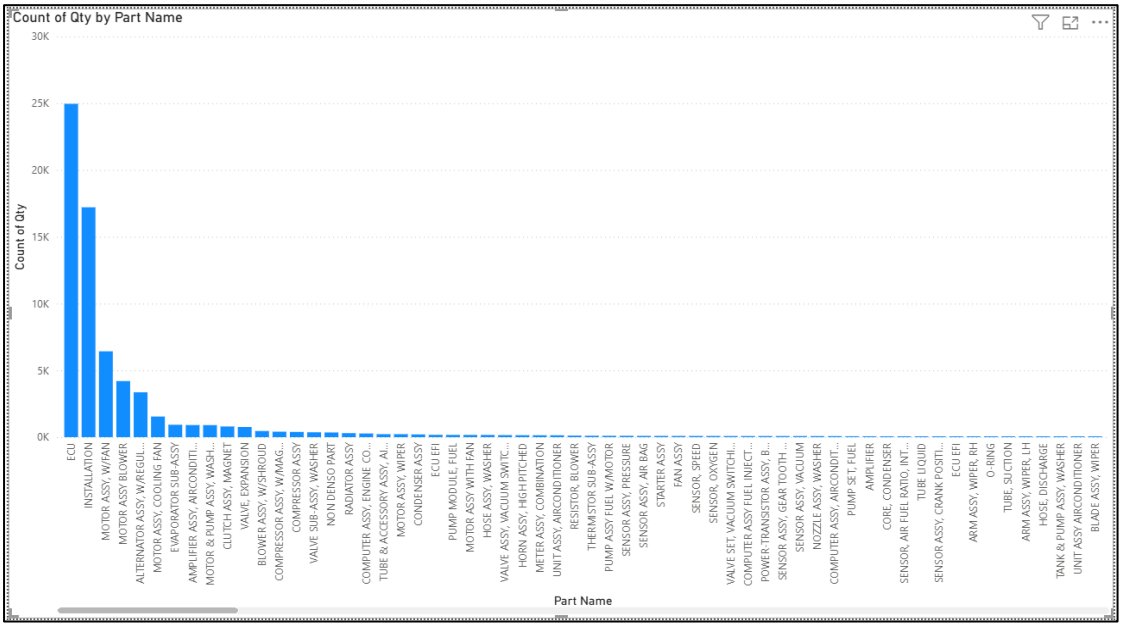


Figure 4.7 Count of Quantity by Part Name

4.2.4.4 Cause Description Categorization

Cause Description analysis aggregates failure reasons to identify dominant root causes. Figure 4.8 presents this information using a pie chart, clearly illustrating the proportional contribution of each cause. This visualization supports prioritization of corrective actions and process improvements by highlighting the most influential failure drivers.

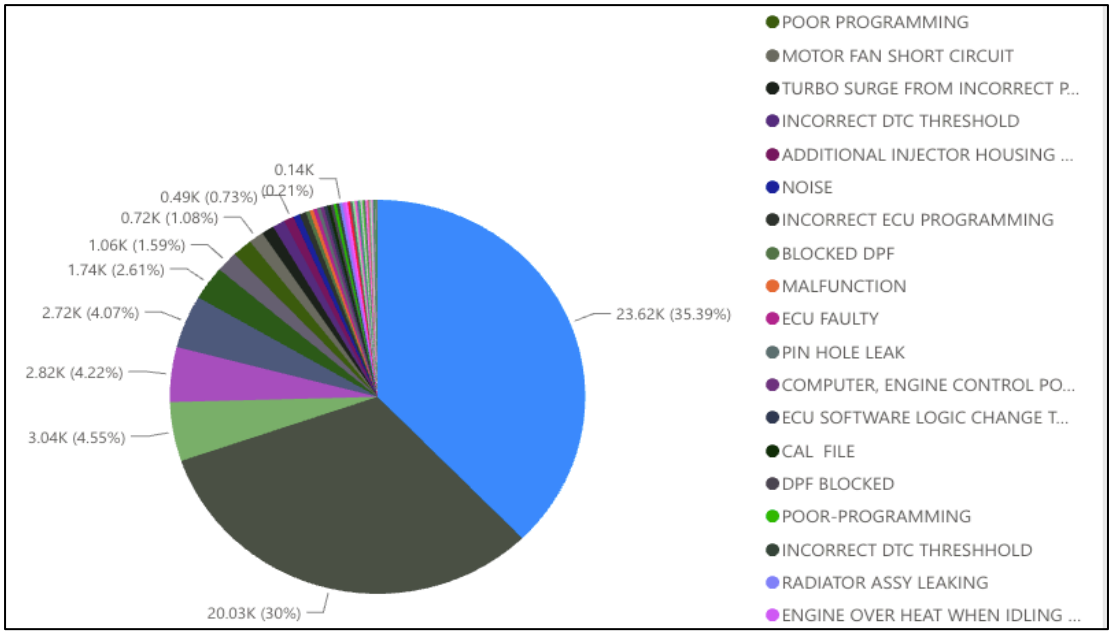


Figure 4.8 Count of Quantity by Cause Description

4.2.4.5 Country-Based Claim Analysis

Figure 4.9 displays claim volumes by country or region using a bar chart. The overseas claim records the highest claim volume, followed by Domestic markets, while other regions exhibit significantly lower activity. This geographical insight assists in identifying region-specific issues related to quality, logistics, or service practices.

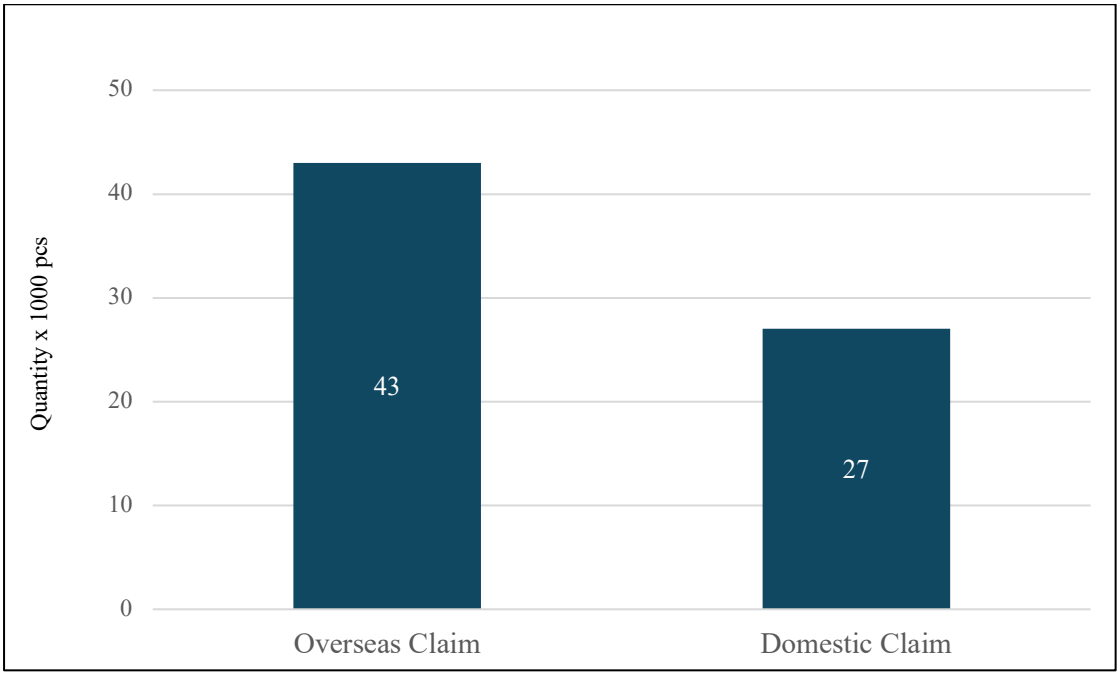


Figure 4.9 Sum of Quantity by Claim Country

4.2.4.6 Warranty Claims by Failed Part

Figure 4.10 shows claim distribution by failed part, where ECU (53.35%), Installation issues (24.38%), and Motor Assembly with Fan (9.1%) dominate. This information is essential for prioritizing engineering countermeasures and guiding cost reduction strategies. The pie chart format effectively highlights proportional impact across components.

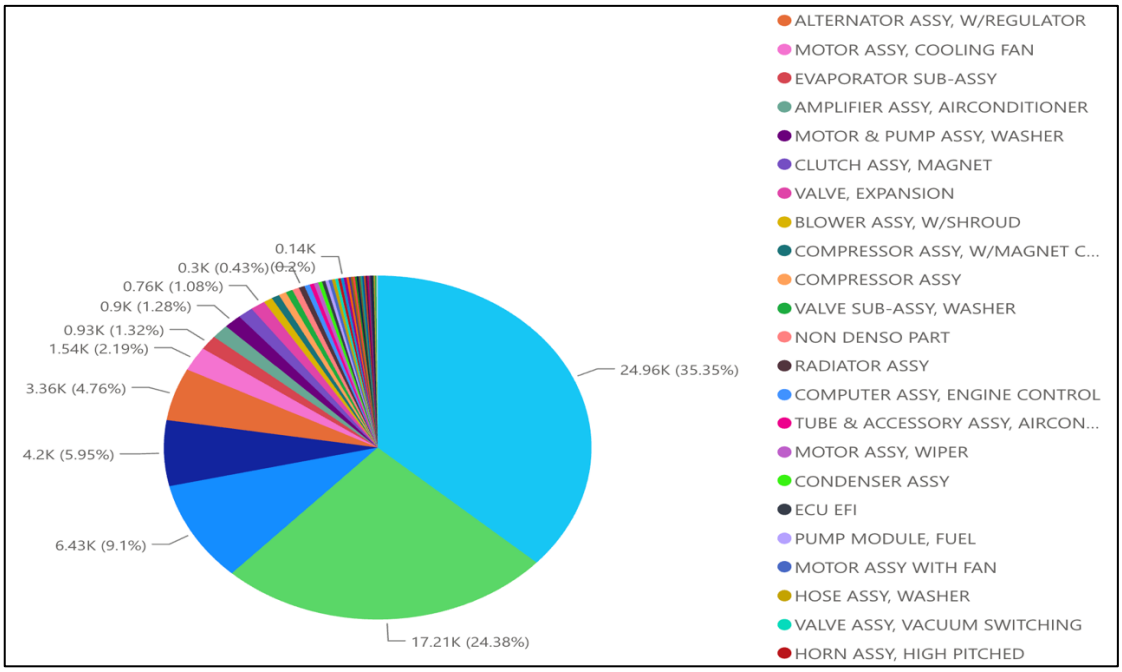


Figure 4.10 Count of Quantity by Part Name

4.2.4.7 Data Preparation and Visualization for Warranty Optimization

Reliable warranty forecasting is highly dependent on rigorous data preparation prior to analytical modelling. In this study, raw warranty records were subjected to systematic preprocessing procedures, including data filtering, categorical standardisation, and consistency validation, to ensure accurate representation of repair timelines, failure classifications, and associated cost information. These preparatory steps were essential to establish a dependable analytical foundation and to prevent distortion of subsequent predictive outcomes. Microsoft Power BI was employed as the primary visualisation platform to support structured exploration of seasonal behaviour, failure distributions, and cost evolution across the study period [150] [151].

Figures 4.11 to 4.14 present the progression from raw transactional data to aggregated trend and cost visualisations. Over the five-year observation window from FY2019 to FY2023, the mean annual warranty expenditure was RM 13.96 K, with a standard deviation of RM 9.72 K, indicating considerable inter-annual variability. Statistical inference using a 95% confidence interval (RM 7.82 K–RM 20.10 K) confirms that the observed cost reductions are statistically significant and unlikely to be attributable to random variation.

The downward trend in warranty costs observed after FY2020 is strongly associated with improvements in diagnostic accuracy, targeted technical training

initiatives, and increasing product maturity. In particular, corrective actions addressing excessive “No Trouble Found” (NTF) claims reduced unnecessary replacements caused by misdiagnosis at service centres. However, a moderate increase in warranty activity in FY2023 suggests the influence of new model introductions, evolving technologies, and supply chain disturbances, underscoring the need for forward-looking analytical tools.

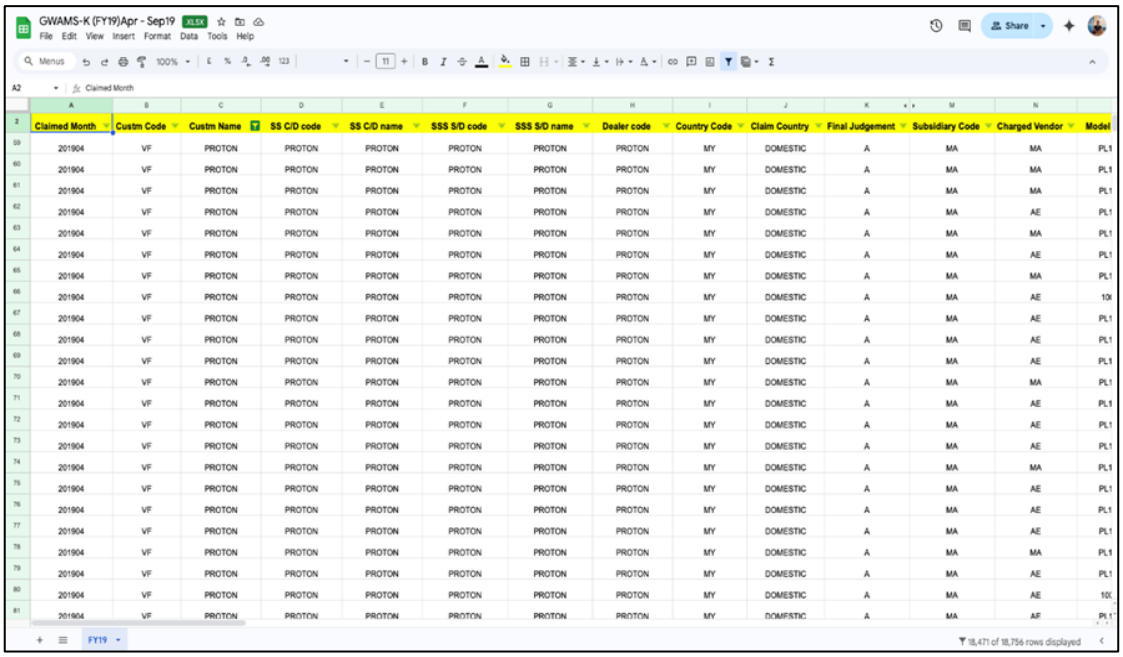


Figure 4.11 Raw Warranty Data in Excel

The declining warranty claim pattern between 2019 and 2022, illustrated in Figure 4.12, reflects enhanced manufacturing stability and quality control maturity. Prior studies have shown that as products progress through their lifecycle, failure frequencies typically decrease due to iterative design refinement, process stabilisation, and improved defect detection [152]. The integration of big data analytics and predictive maintenance strategies further reinforces this trend by enabling early fault identification and optimised service interventions.

Parametric analysis was subsequently applied to evaluate key warranty performance indicators, including production volume effects, claim timing characteristics, and lifecycle-related failure behaviour. This analytical stage enabled structured scenario evaluation, predictive modelling, and continuous monitoring of field performance through an interactive Power BI dashboard, as shown in Figure 4.15 [153].

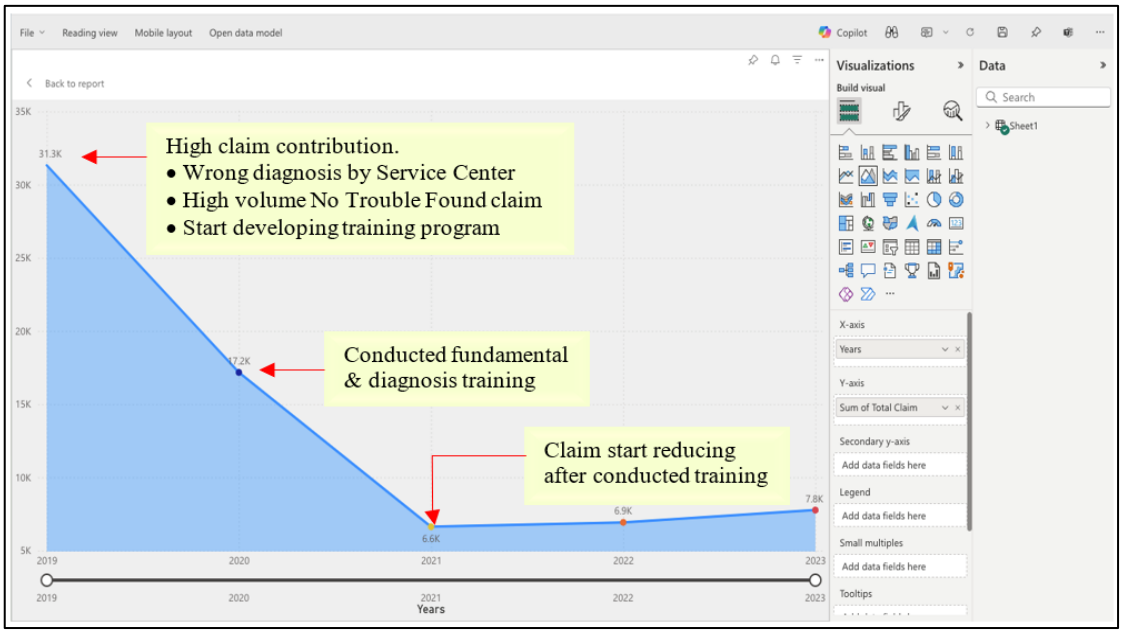


Figure 4.12 Five-Year Warranty Claims Trend

A detailed cost distribution analysis, presented in Figure 4.13, highlights the concentration of warranty expenditure in the early years of the study. FY2019 and FY2020 together accounted for approximately 93% of total five-year warranty costs, primarily due to widespread diagnostic inconsistencies and elevated NTF claim rates. In contrast, the substantial reduction observed from FY2021 onwards reflects improvements in fault confirmation procedures, proactive servicing, and data-driven quality interventions.

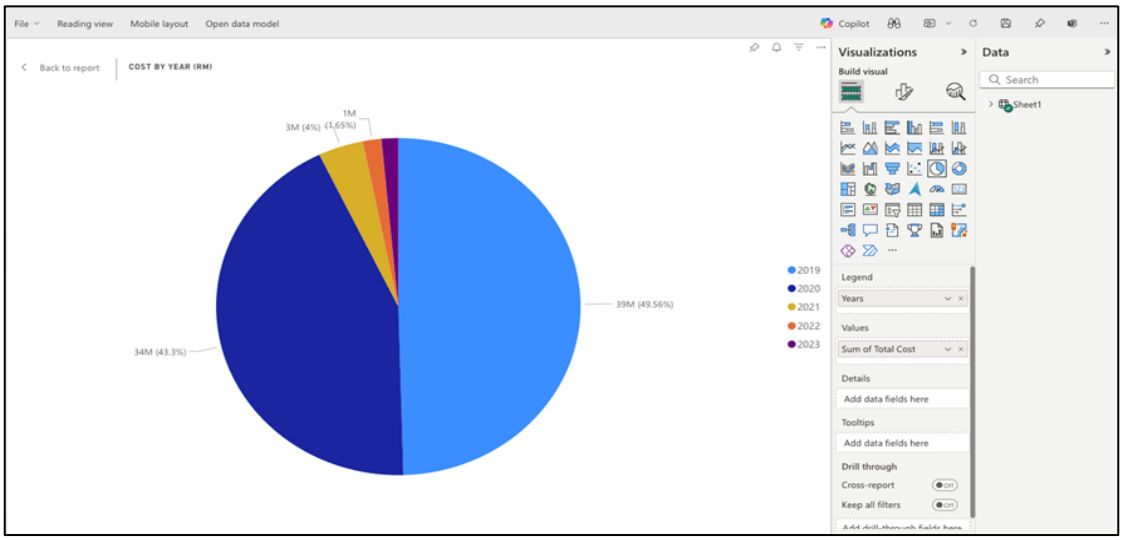


Figure 4.13 Warranty Cost Distribution by Year

Monthly cost fluctuations captured in Figure 4.14 provide further insight into product lifecycle dynamics and operational influences. Early-stage claim spikes are commonly associated with initial field exposure of new products, while later variations may arise from seasonal usage patterns, supply constraints, or changes in repair strategy. The application of AI-enabled predictive analytics, combined with BI-based visualisation, allows these fluctuations to be transformed into actionable intelligence, supporting more accurate warranty reserve planning and risk mitigation.

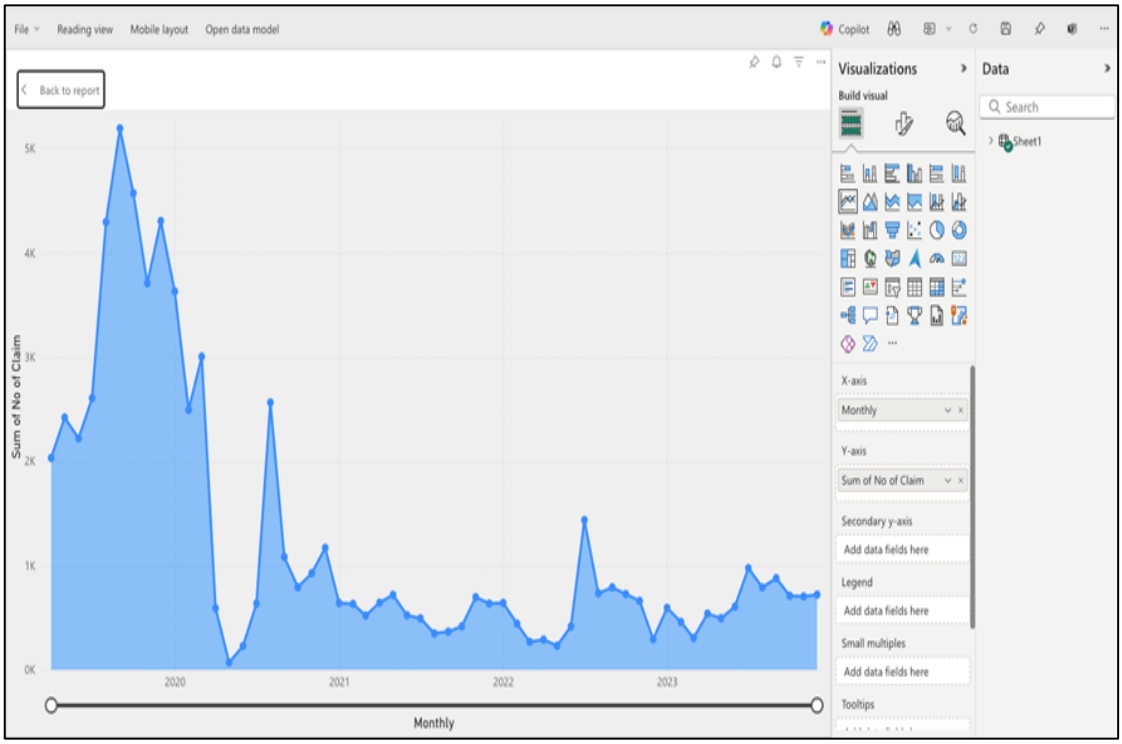


Figure 4.14 Warranty Cost Distribution by Year

Recent advances in machine learning, including time-series forecasting and anomaly detection, have demonstrated significant improvements in warranty prediction accuracy, with reported gains exceeding 30% in comparable industrial applications. Such capabilities are increasingly critical in the context of electric vehicles and advanced driver-assistance systems, where failure behaviour is more dynamic and data-intensive. The integration of Power BI with AI-driven diagnostic outputs enables automated cost forecasting, early risk detection, and optimised repair decision-making.

Internal benchmarking conducted as part of this study further validates the effectiveness of the proposed approach. The legacy Excel-based method achieved approximately 58% accuracy in anomaly identification, whereas the AI-integrated

system incorporating Python-based analytics improved detection accuracy to 86%, representing a 28% performance gain. This outcome demonstrates the practical benefits of transitioning from reactive warranty management to a predictive, data-driven framework, supporting improved operational efficiency, cost control, and long-term customer satisfaction.

4.3 AWAS Dashboard Design

The AWAS dashboard integrates key visualizations into a centralized decision-support platform as shown in Figure 4.15. Warranty trends from 2019–2023 exhibit a typical product lifecycle pattern, with early peaks, subsequent decline, and stabilization. Toyota accounts for 66.71% of total claims, reflecting market share and component-specific reliability challenges, particularly in ECU and air-conditioning systems.

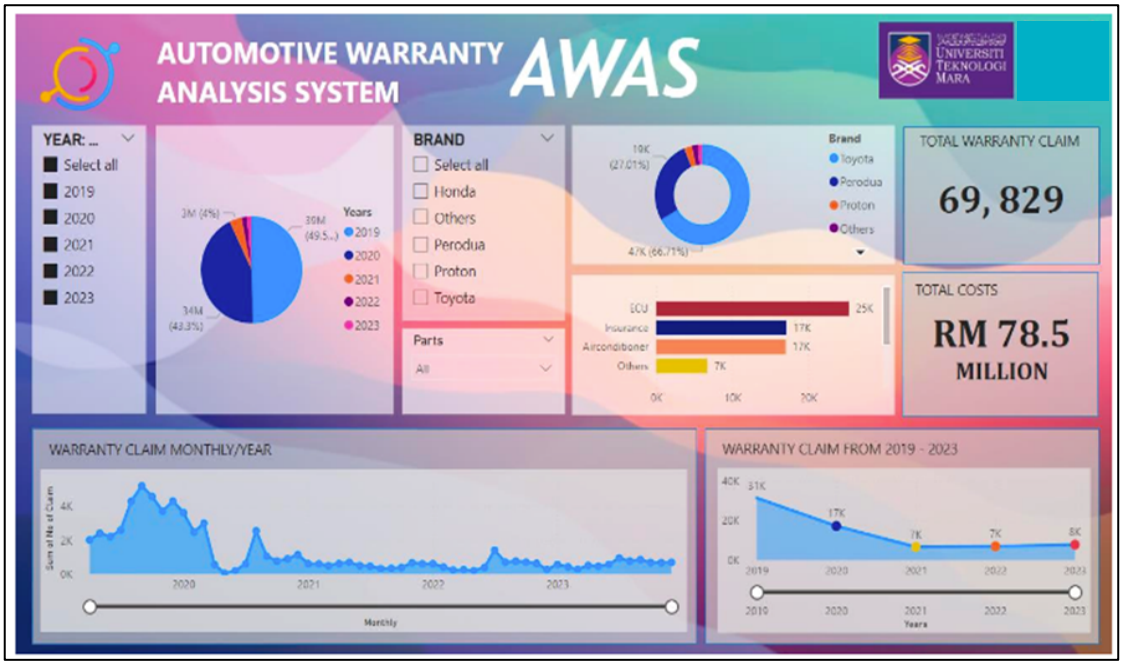


Figure 4.15 Automotive Warranty Analysis System (AWAS) dashboard.

AI-driven diagnostics and predictive analytics have contributed to claim reductions post-2021. However, emerging trends in 2023 indicate the need for continuous monitoring and adaptive analytics. Power BI supports this by enabling real-time visualization, automated forecasting, and informed policy refinement. [152].

4.4 Warranty Trends and KPI Insights

AWAS tracks critical KPIs such as claim frequency, resolution time, and average claim cost over a five-year period. These indicators reveal measurable progress in warranty performance while identifying persistent issues requiring targeted action. Continuous KPI monitoring enables proactive quality improvement, cost optimization, and enhanced customer satisfaction [153], [154].

Implementation challenges included inconsistent historical records, limited high-resolution failure patterns, and technical integration issues between Python analytics and Power BI. These challenges underscore the importance of robust data pipelines and system integration planning.

4.5 System Impact

The implementation of the Automotive Warranty Analysis System (AWAS) highlights the strategic value of big data analytics in modern warranty management by facilitating a transition from corrective, post-failure actions to anticipatory and evidence-based decision processes. Through its scalable architecture, the system is capable of accommodating increasing data volumes and adapting to changing operational and industry demands, ensuring long-term applicability and robustness [155], [156].

By integrating advanced analytical capabilities with day-to-day operational requirements, AWAS provides a structured and sustainable framework for enhancing warranty governance. This integration supports continuous improvement in product reliability, more effective warranty cost management, and strengthened long-term financial performance, aligning technical analytics with organisational quality and cost optimisation objectives [75], [157].

4.6 Interface Dashboard

The final AWAS interface in Figure 4.16 presents an integrated and interactive visualization of warranty performance from 2019–2024. Key indicators include Total Cases (71k), Total Cost (RM79.18M), and Total Claim Amount (RM21.68M). Core features include:

- Time-based analysis of L/O and Repair Dates
- Model-based filtering for high-risk vehicle identification
- Part and cause analysis for root cause investigation
- Geographic and dealer segmentation for regional performance tracking

The dashboard supports the project objectives of organizing large warranty datasets, enhancing visualization clarity, and enabling effective root cause analysis for data-driven decision-making. Furthermore, the effectiveness of the AWAS dashboard is strengthened through the integration of time based analytical parameters and contextual visual indicators within the figures and tables. In Figure 4.15 and Figure 4.16, warranty performance is not only represented through total claim values and frequency distributions, but also through temporal factors such as claim occurrence period, repair lead time, resolution duration, and lifecycle trend progression from 2019 to 2024. These time related indicators provide clearer justification of how warranty behaviour evolves throughout the product lifecycle, enabling more accurate interpretation of recurring failures, delayed corrective actions, and high risk operational periods [158] [159]. In addition, the integration of model based filtering, part specific failure analysis, geographic segmentation, and customer claim trends enhances the interpretability of the dashboard by linking statistical outputs with practical operational factors. This allows the figures and tables to reflect not only quantitative performance indicators, but also the underlying engineering and operational conditions contributing to warranty escalation. Consequently, the dashboard provides a more comprehensive and evidence based justification for decision making, supporting proactive intervention, continuous monitoring, and strategic warranty cost optimisation [160] [161][162].

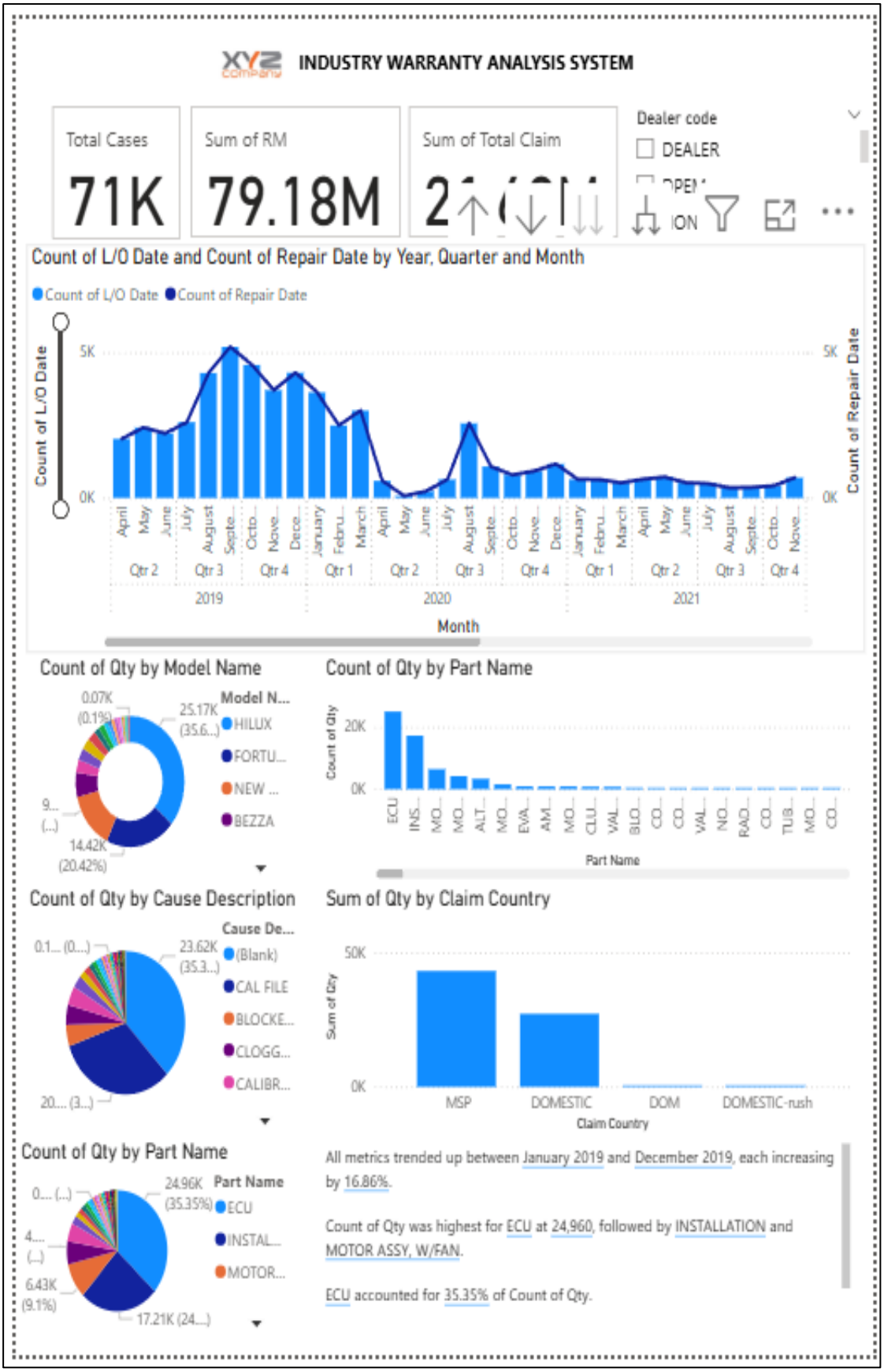


Figure 4.16 Graphical User Interface

4.7 Model Development Results

4.7.1 Training Performance Analysis

Six forecasting models were developed, covering both statistical and AI-based approaches with varying feature complexity. The Linear Basic model employed ten core predictors, while the Linear Enhanced and tree-based models utilised an expanded 23-feature set to capture richer temporal and business interactions. ARIMA was applied as a univariate time-series model, relying solely on historical warranty costs. An Ensemble model combined multiple model outputs to improve overall robustness and stability.

Table 4.1
Training Performance Comparison

Model	Features	Train R^2	RMSE	MAPE (%)
Linear Basic	10	0.208	4301	109.7
Linear Enhanced	23	0.659	2822	76.1
ARIMA	1	0.154	4444	99.8
Random Forest	23	0.755	2391	58.8
XGBoost	15	0.739	2471	69
Ensemble	17	0.567	2888	79.1

Table 4.1 compares the training performance of the forecasting models using R^2, RMSE, and MAPE. The Train R^2 values represent the fitting capability during model training, while RMSE and MAPE evaluate prediction error performance. The results show that higher training accuracy did not consistently translate into reliable generalisation performance. Random Forest achieved the highest training fit (Train R^2 = 0.755). XGBoost also demonstrated strong training performance (Train R^2 = 0.739), suggesting comparatively better generalisation control but still limited forecasting reliability. Instead, RMSE, MAPE, forecasting stability, and practical generalisation capability must also be considered to determine the most suitable forecasting model for warranty prediction applications.

4.7.2 Bias-Variance Trade-off in Warranty Datasets

The results reflect a typical bias–variance trade-off. Simpler models (e.g., Linear Basic) tend to be more stable across different samples (lower variance) but may underfit and miss nonlinear structure (higher bias). Conversely, flexible models (Random Forest, XGBoost) can capture complex patterns (lower bias) but become highly sensitive to small dataset fluctuations (higher variance), particularly when the training dataset contains only 48 monthly observations. In this study, model complexity exceeded the information content available in the historical record, leading to inflated training performance but weak deployment reliability. Overall, XGBoost provided the most balanced individual model behaviour among the AI methods because regularisation partially constrained variance while retaining pattern-learning capacity.

4.7.3 Overfitting Detection and Practical Implication

Overfitting was most evident in Random Forest, where the very negative expected out-of-sample R^2 suggests that learned patterns were specific to the training period and did not represent transferable warranty behaviour. This risk is amplified in warranty datasets because failure trends can shift due to quality improvement actions, process changes, supplier countermeasures, and reporting practices. In practical deployment, overfitted models create high budgeting risk because they generate confident but inaccurate forecasts. The findings confirm that model selection must prioritise *stability and business realism* over training fit when historical data is limited.

4.7.4 Multi-Model Forecast Generation

The forecast comparison indicates strong divergence across models when extrapolating beyond Month 48 (training endpoint). This divergence reflects model sensitivity to feature design, seasonal representation, and how each method handles extrapolation as shown in Figure 4.17 and Figure 4.18. Forecasts diverged markedly across models beyond the training horizon (Month 48), reflecting differences in feature representation, treatment of seasonality, and extrapolation behaviour. As illustrated in Figures 4.17 and 4.18, models exhibited varying sensitivity to temporal dynamics,

highlighting the impact of modelling assumptions on out-of-sample warranty cost projections.

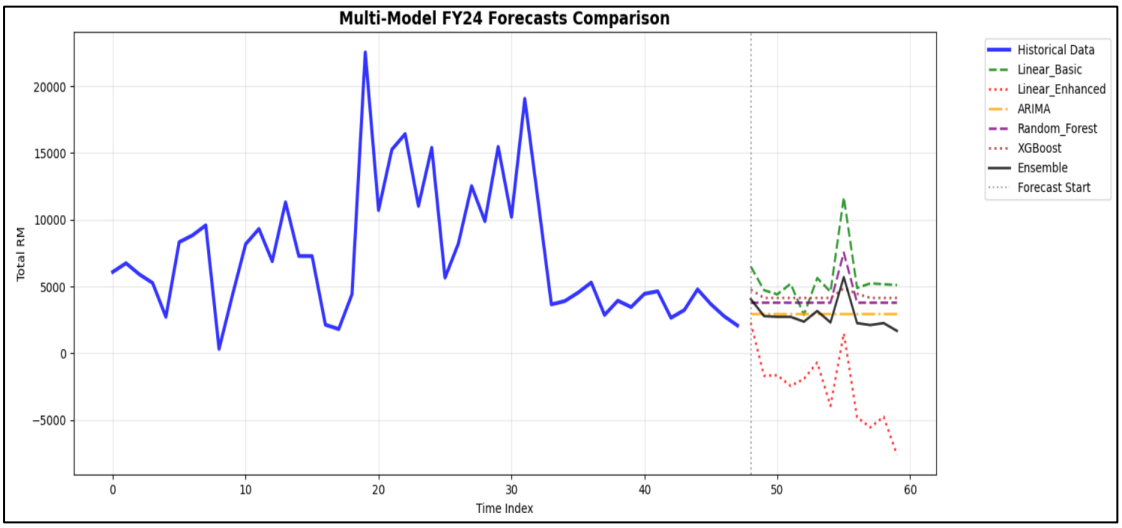


Figure 4.17 Historical Warranty Costs Versus FY2024 Forecasts (Multiple Models)

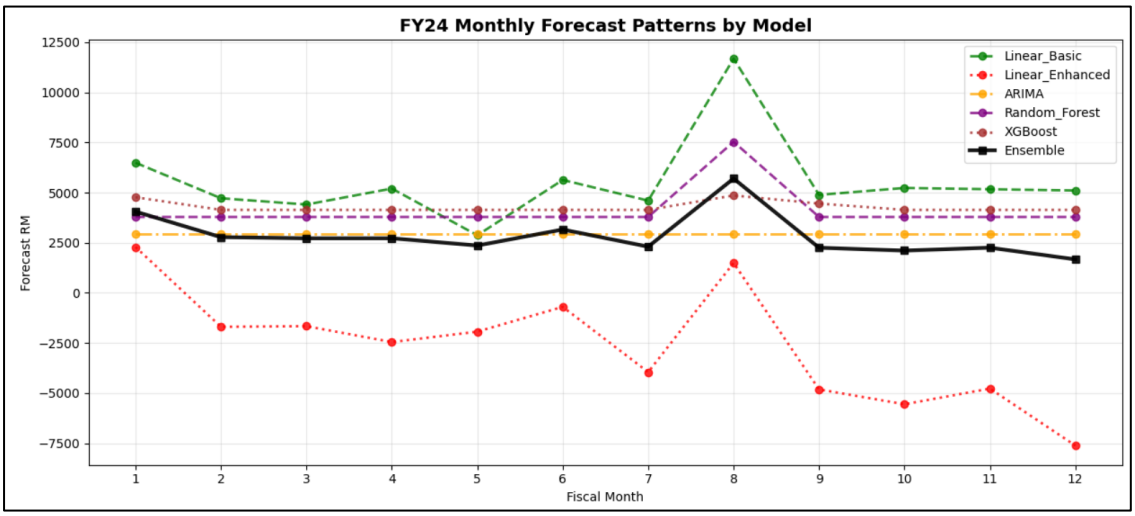


Figure 4.18 FY2024 Monthly Forecast Comparison by Model

a) Annual Total Predictions by Model

Annual totals ranged widely. A major failure was observed in the Linear Enhanced model, which produced negative monthly costs and a negative annual total. Since negative warranty cost is not physically meaningful, this output indicates violation of business constraints and invalidates the model for deployment. This issue commonly occurs when linear models extrapolate beyond the training range in the presence of multicollinearity, unstable coefficients, or excessive feature interactions as

shown in Table 4.2. For real deployment, such models must include non-negativity constraints or post-processing rules.

Table 4.2
FY24 Annual Total Comparison by Model

Model	FY24 Forecast Total (RM)	FY24 Monthly Average (RM)	FY24 Variance Range (RM)
Linear Basic	66,003	5,500	8,807
Linear Enhanced	-31,354	-2,613	9,901
ARIMA	35,311	2,943	0
Random Forest	49,183	4,099	3,749
XGBoost	51,351	4,279	718
Ensemble	34,099	2,842	4,030

b) Monthly Pattern Observations

XGBoost produced the most stable pattern (variance range RM718), while Linear Basic showed high volatility and an extreme November peak (RM11,680). ARIMA generated a flat forecast, indicating that the configured ARIMA structure did not capture seasonal dynamics under the chosen parameters, or that the differencing/identification procedure suppressed variation. The Ensemble forecast displayed moderate variation and produced an annual total close to the actual FY2024 warranty cost, demonstrating practical benefit from averaging out extreme behaviours among individual models as shown in Table 4.3.

Table 4.3
FY24 Forecasted Monthly Warranty Costs (Apr 2024 – Mar 2025)

Month	Linear Basic	Linear Enhanced	ARIMA	RF	XGBoost	Ensemble
Apr-24	6,490	2,286	2,943	3,786	4,769	4,055
May-24	4,719	-1,692	2,943	3,786	4,141	2,779
Jun-24	4,407	-1,660	2,943	3,786	4,141	2,723
Jul-24	5,204	-2,451	2,943	3,786	4,141	2,724
Aug-24	2,873	-1,933	2,943	3,786	4,141	2,362
Sep-24	5,632	-696	2,943	3,786	4,141	3,161
Oct-24	4,601	-3,950	2,943	3,786	4,141	2,304
Nov-24	11,680	1,493	2,943	7,535	4,859	5,702

Dec-24	4,888	-4,812	2,943	3,786	4,455	2,252
Jan-25	5,234	-5,554	2,943	3,786	4,141	2,110
Feb-25	5,170	-4,771	2,943	3,786	4,141	2,254
Mar-25	5,106	-7,615	2,943	3,786	4,141	1,672
Total	66,004	-31,354	35,316	49,183	51,352	34,098

4.8 Validation Against Actual Data

4.8.1 FY24 Actual Warranty Performance

FY2024 actual accepted claims totalled RM32,743.77 across 3,537 claims (Table 4.4). The monthly pattern shows a peak in May (RM5,266.22) and a trough in March (RM1,159.13), reflecting substantial seasonal variability that all models struggled to reproduce precisely.

Table 4.4
FY24 Actual Warranty Data Summary

Month	Claims Count	Total RM	Monthly Pattern
Apr-24	443	3,436.52	Above average
May-24	575	5,266.22	Peak month
Jun-24	302	2,634.33	Below average
Jul-24	363	3,177.10	Average
Aug-24	404	4,624.42	Above average
Sep-24	271	1,805.23	Low
Oct-24	322	2,803.89	Average
Nov-24	222	2,496.96	Below average
Dec-24	211	1,487.35	Low
Jan-25	192	2,318.48	Below average
Feb-25	131	1,504.14	Low
Mar-25	101	1,159.13	Trough month
Total FY24	3,537	32,743.77	

4.8.2 Actual vs. Predicted Performance and Ranking

Compares the actual monthly warranty claim cost in FY2024 with the forecasted results generated by the best performing model (Ensemble) and the weakest performing model (Linear Enhanced) as shown in Figure 4.19. The actual warranty trend shows fluctuating claim costs throughout the year, with the highest value recorded in May at approximately RM5,300. These fluctuations indicate that warranty behaviour is influenced by factors such as failure occurrence timing, customer usage patterns, claim reporting variation, and operational conditions.

The Ensemble model showed forecasting results that were closer to the actual trend and maintained more stable prediction performance throughout the year. The model was able to follow the general movement of the warranty claim pattern with lower forecasting deviation, which is consistent with its lower MAPE and total forecasting error presented in Table 4.5. This indicates that the Ensemble model provided better forecasting reliability for practical warranty prediction applications. In comparison, the Linear Enhanced model showed large forecasting deviations and produced unrealistic negative warranty cost predictions, particularly from October onward. The largest error occurred in March, indicating poor model stability and weak generalisation performance. This condition suggests that the model was overfitted to the training dataset and was unable to adapt to unseen warranty trends.

The forecasting errors observed in Figure 4.19 were also influenced by several industrial factors, including inconsistent historical claim records, non-trouble found (NTF) cases, fluctuating warranty patterns, and limited dataset size. In addition, factors such as maintenance practice variation, environmental conditions, and delayed claim reporting contributed to the instability of the forecasting trend. To improve forecasting performance, future analysis should include larger datasets, better data standardisation, removal of inconsistent records, and integration of additional operational parameters such as repair lead time, vehicle mileage, and service history.

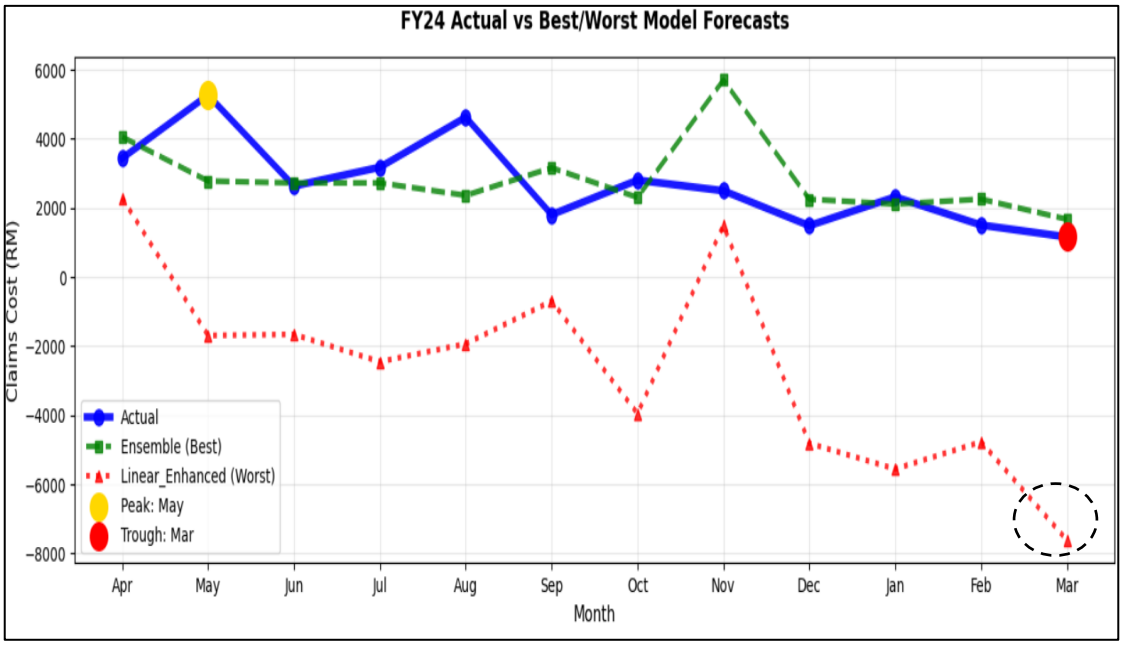


Figure 4.19 Actual versus Predicted Warranty Claim Cost Performance for FY2024

Table 4.5
Model Ranking (by MAPE – Mean Absolute Percentage Error)

Rank	Model	MAPE (%)	RMSE	Total Error (%)	Directional Accuracy (%)
1	Ensemble	42.3	1463.8	4.2	36.4
2	ARIMA	47.8	1221	8.0	0
3	Random Forest	88.5	2131.6	50.3	9.1
4	XGBoost	95.0	1984.7	57.0	9.1
5	Linear Basic	154.2	3777.3	101.8	54.5
6	Linear Enhanced	250.4	5871.9	-195.8	27.3

Table 4.5 shows that the Ensemble model achieved the best overall forecasting performance, with the lowest MAPE (42.3%) and annual total error (4.2%), indicating better forecasting stability and reliability compared to the other models. ARIMA ranked second, with lower RMSE and moderate MAPE values; however, its Directional Accuracy was 0%, meaning the model failed to correctly predict the month to month movement trend of the warranty cost. Directional Accuracy refers to the ability of the model to correctly identify whether the data trend increases or decreases between consecutive periods. For example, a Directional Accuracy of 36.4% means that the model correctly predicted the trend direction in approximately 4 out of 11 forecasting intervals.

Random Forest and XGBoost showed weak forecasting performance, with high MAPE values and low Directional Accuracy, indicating that higher model complexity did not improve prediction performance under limited dataset conditions. Although the Linear Basic model achieved the highest Directional Accuracy (54.5%), its MAPE remained very high at 154.2%, showing that the model could identify the trend direction better but failed to provide accurate numerical forecasting as shown in Figure 4.20. Therefore, the Ensemble model was selected as the best forecasting model because it provided more balanced and reliable performance in terms of both forecasting accuracy and overall prediction stability.

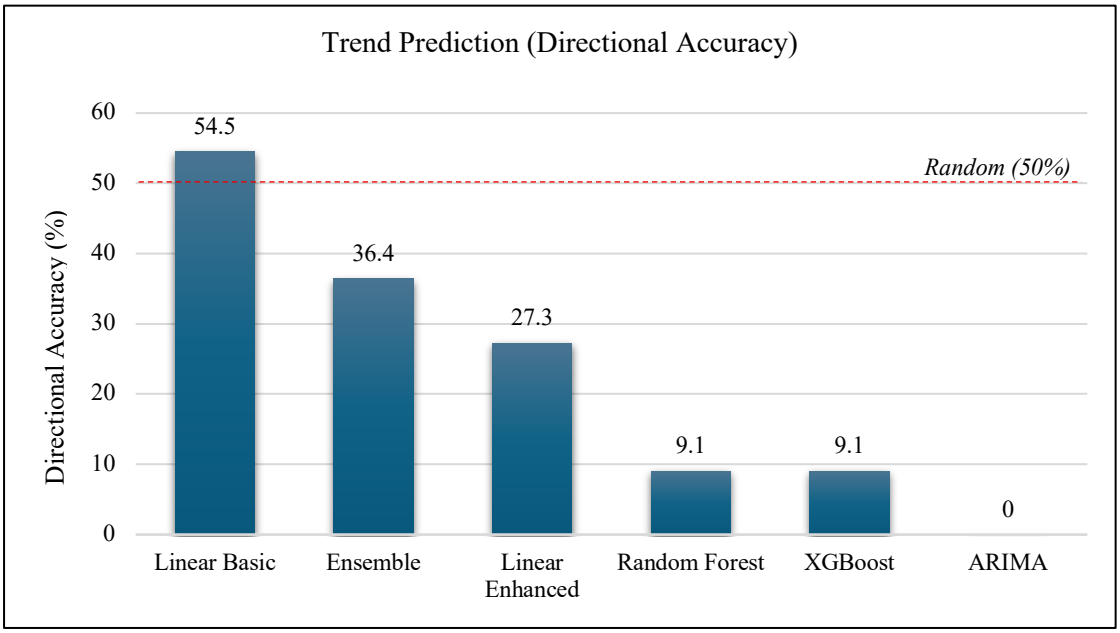


Figure 4.20 Directional Accuracy of Forecasting Models in Predicting Warranty Cost Trends

Monthly error patterns demonstrated in Table 4.6 and that all models struggled with seasonal extremes, particularly underestimating the May peak (actual: RM 5,266 vs. best prediction: RM 4,769 by XGBoost) and overestimating the March through (actual: RM 1,159 vs. lowest prediction: RM 1,672 by Ensemble). This systematic pattern indicated fundamental limitations in seasonal pattern recognition with available historical data.

Table 4.6
Monthly Warranty Cost Actual vs Forecast Comparison by Model

Month	Actual	Lin_Basic	Lin_Enh	ARIMA	RF	XGB	Ensemble
Apr	3,437	6,490	2,286	2,943	3,786	4,769	4,055
May	5,266	4,719	-1,692	2,943	3,786	4,141	2,779
Jun	2,634	4,407	-1,660	2,943	3,786	4,141	2,723
Jul	3,177	5,204	-2,451	2,943	3,786	4,141	2,724
Aug	4,624	2,873	-1,933	2,943	3,786	4,141	2,362
Sep	1,805	5,632	-696	2,943	3,786	4,141	3,161
Oct	2,804	4,601	-3,950	2,943	3,786	4,141	2,304
Nov	2,497	11,680	1,493	2,943	7,535	4,859	5,702
Dec	1,487	4,888	-4,812	2,943	3,786	4,455	2,252
Jan	2,318	5,234	-5,554	2,943	3,786	4,141	2,110
Feb	1,504	5,170	-4,771	2,943	3,786	4,141	2,254
Mar	1,159	5,106	-7,615	2,943	3,786	4,141	1,672

4.8.3 Model Selection Recommendations

The validation results identified the Ensemble model as the most suitable forecasting approach for warranty prediction analysis. The model achieved the lowest forecasting error, with a MAPE of 42.3% and annual total error of 4.2%, indicating more consistent prediction performance compared to the other models. The averaging mechanism within the Ensemble approach reduced extreme forecasting deviations and improved overall prediction stability during FY2024 evaluation. In contrast, Random Forest and XGBoost demonstrated unstable forecasting behaviour despite recording high fitting performance during training. The models produced large forecasting errors when tested on unseen data, indicating limited generalisation capability. This condition was mainly caused by overfitting, where the models adapted too closely to historical training patterns that were not fully representative of future warranty conditions.

Although ARIMA achieved relatively lower RMSE and moderate MAPE values, the model was not selected as the best forecasting approach because it failed to accurately capture the actual month to month warranty trend movement. The model recorded 0% Directional Accuracy, indicating that it was unable to correctly predict whether the warranty claim trend increased or decreased between consecutive periods. The forecasting output also appeared relatively flat and less responsive to sudden

warranty fluctuations caused by recurring alternator noise issues and changing operational conditions. As a result, the model showed limited practical capability for dynamic warranty forecasting applications.

The forecasting instability was further influenced by irregular warranty claim behaviour throughout the fiscal period. Significant fluctuation in monthly warranty cost was observed due to recurring alternator noise complaints associated with design related component weaknesses. Variations in failure occurrence timing, repair activities, and claim submission patterns contributed to inconsistent trend movement, increasing the complexity of warranty forecasting. Several operational conditions also affected forecasting performance, including inconsistent claim records, non-trouble found (NTF) cases, delayed reporting, and varying customer operating environments. These factors introduced uncertainty into the dataset and reduced prediction consistency, particularly for complex machine learning models that required larger and more stable datasets for effective learning.

Overall, the findings indicate that forecasting reliability in industrial warranty applications depends not only on model complexity, but also on data quality, trend consistency, and model stability during unseen data evaluation. The Ensemble model demonstrated better practical suitability for warranty planning and monitoring due to its balanced forecasting behaviour and lower prediction deviation under fluctuating industrial conditions.

4.9 Warranty Data Visualization

4.9.1 Warranty Claims Trend: Quantity Versus Paid Amount

The long-term warranty trend (FY2018–FY2025) highlights a consistent relationship between claim quantity and paid amount, where reductions in claim volume are generally reflected in lower warranty payout. In FY2018, Company MNC recorded 4,690 warranty claims with a total paid amount of RM616,000. Claim volume gradually declined over subsequent years, with notable reductions in FY2020 (3,452 claims) and FY2021 (2,977 claims). The paid amounts followed the same downward direction, indicating improved field reliability and/or effectiveness of quality and containment actions. The FY2023–FY2026 outlook shows continued reduction, with warranty claim paid amounts of RM388K (FY2023), RM342,000 (FY2024) and RM291,000

(FY2025), supporting the quality improvement and repair-oriented initiatives able achieved significant mitigate warranty cost results in total twenty five percent (25%) within 3 years in Figure 4.21.

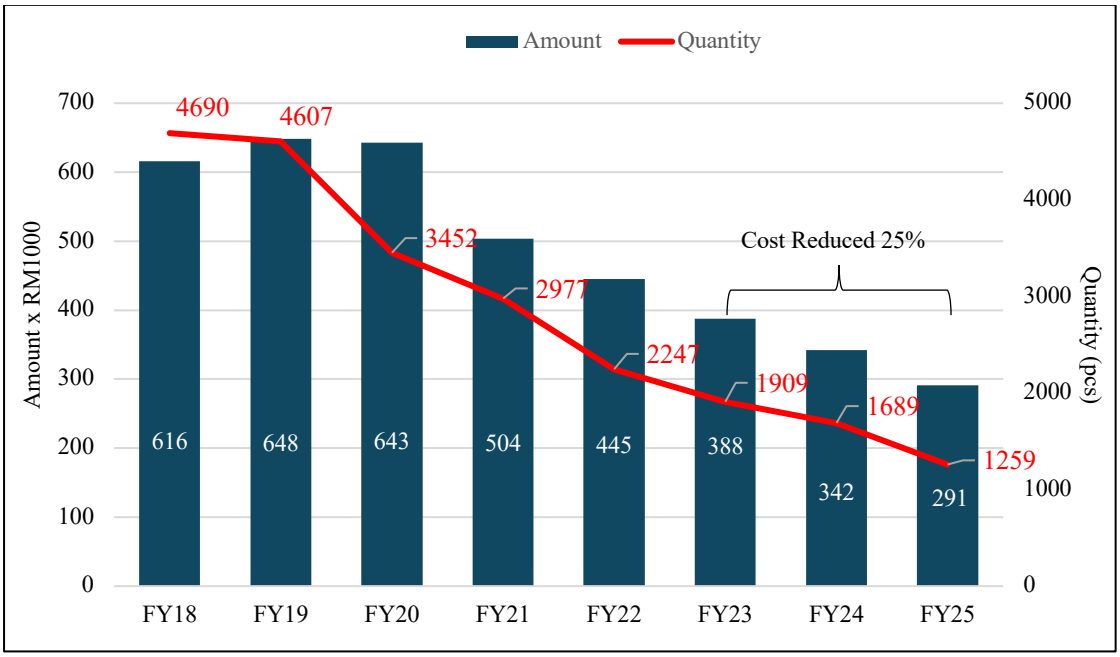


Figure 4.21 Total Warranty Claim Based on Quantity vs Paid Amount

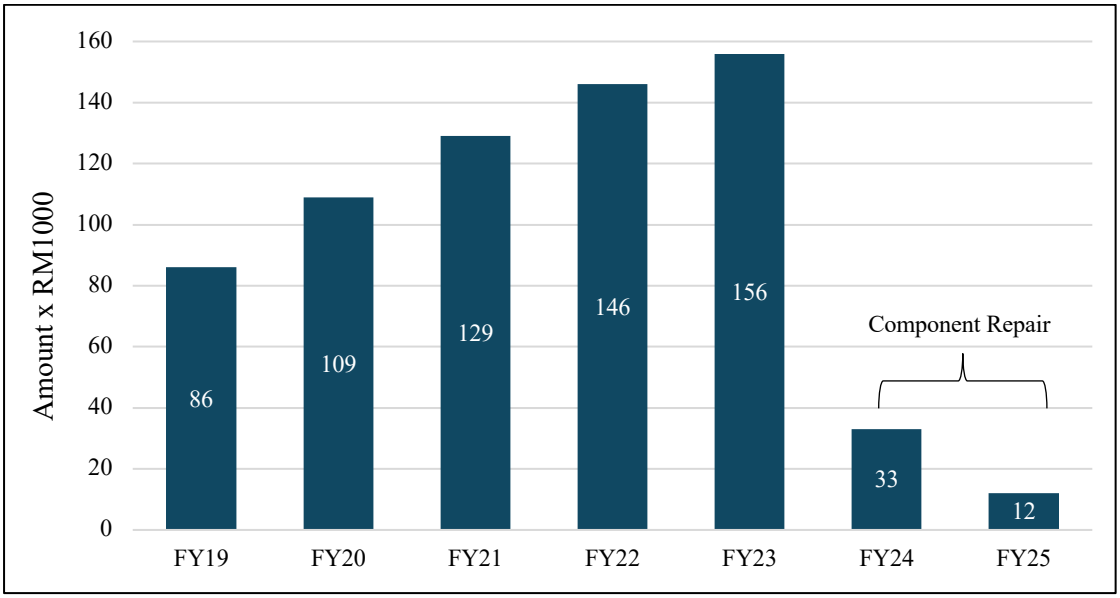


Figure 4.22 Total Company A Alternator Assy Warranty Claim Trend

Figure 4.22 shows the warranty claim trend for Company A alternator assemblies from FY2019 to FY2025. The warranty claim amount increased steadily from RM86K in FY2019 to RM156K in FY2023, indicating a recurring reliability issue

associated with the alternator assembly. The increasing trend was mainly contributed by abnormal noise complaints reported during vehicle operation. Detailed investigation identified that the noise issue was primarily caused by design related weaknesses within the alternator system, particularly involving bearing degradation and internal component instability during prolonged operating conditions.

The continuous increase in warranty claims between FY2019 and FY2023 indicates that the earlier corrective actions were insufficient to effectively control the recurring failure trend. The repeated occurrence of the same failure mode suggests that the issue was systematic and required a more sustainable warranty mitigation approach beyond conventional part replacement activities. A significant reduction in warranty claim amount was observed after the implementation of component repair activities beginning in FY2024. The warranty claim amount decreased from RM156K in FY2023 to RM33K in FY2024, representing a reduction of RM123K or approximately 78.8%. The claim amount further reduced to RM12K in FY2025, contributing to a total reduction of RM144K or approximately 92.3% compared to FY2023. This improvement was mainly achieved through the introduction of component repair activities as part of the warranty cost reduction strategy. Instead of complete alternator replacement, selected failed components were repaired and reused following technical verification and failure assessment. This approach significantly reduced warranty replacement cost, improved component utilisation efficiency, and minimised recurring warranty expenditure associated with the noise issue. The results demonstrate that the implementation of component repair strategy combined with warranty trend analysis can effectively support warranty cost optimisation and improve operational efficiency. The findings also indicate that integrating root cause analysis with targeted repair activities provides a more practical and sustainable solution for managing recurring warranty issues within the automotive industry.

4.9.2 Sales Growth Context

Company MNC sales increased from RM594 million (FY2019) to RM648 million (FY2020) and RM674 million (FY2021). A step-change growth was observed in FY2022 and FY2023, where sales reached RM981 million and RM1,002 million, respectively. The positive results continue towards RM1,062 million by FY2025 as shown in Figure 4.23. The combination of sales growth and mitigation warranty cost

indicates that reliability improvements are occurring alongside business expansion, reinforcing Company MNC's competitive position and supporting more confident warranty provisioning decisions. The achievement results shown that the Company MNC achieved the Key Performance Index (KPI) companywide in six percent of total sales within 3 years (FY2023-FY2025).

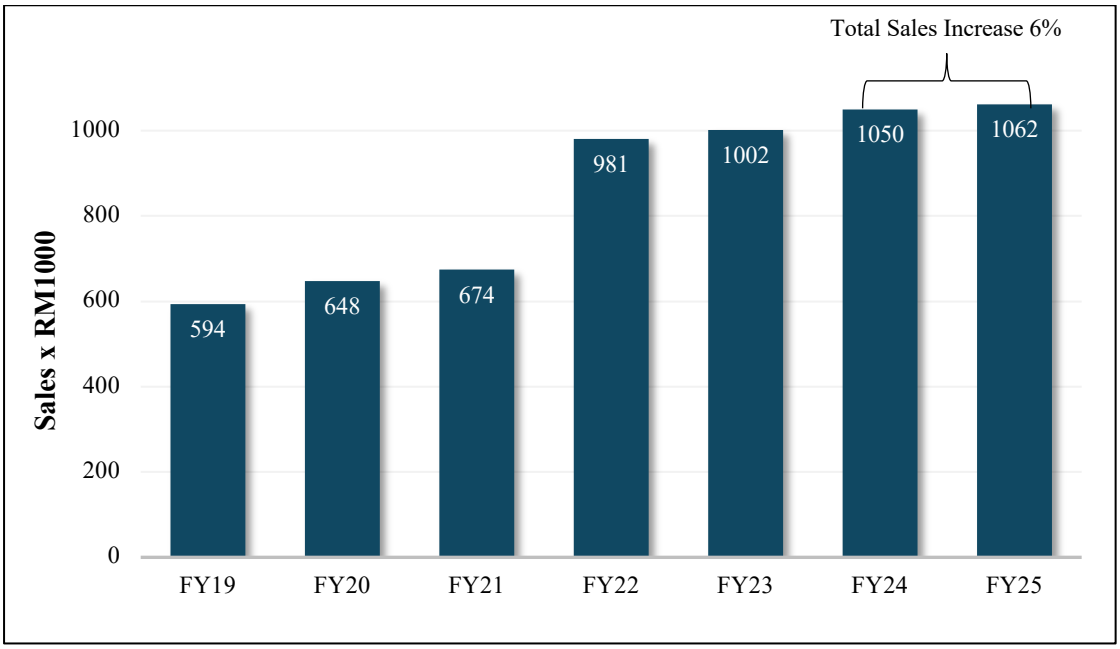


Figure 4.23 Total Company MNC Total Sales Trend

CHAPTER 5
CONCLUSION AND RECOMMENDATION

5.1 Conclusion

This This research successfully achieved the objectives of developing and evaluating a data driven warranty management system model for the Malaysian automotive industry through the integration of big data analytics, predictive analysis, and interactive visualisation techniques. The study focused on warranty claim data associated with alternator assemblies supplied by Company MNC to Company A within the domestic automotive market. By analysing nine years of historical warranty data, the research identified significant patterns related to failure trends, warranty cost behaviour, recurring component issues, and operational inefficiencies.

The first objective, which focused on analysing warranty claim trends and identifying major contributing factors, was successfully achieved through structured data cleansing, classification, and visual analytics. The findings showed that warranty cost demonstrated a decreasing trend from RM648K in FY2019 to RM388K in FY2023, representing an overall reduction of approximately 40.1%. Despite this improvement, the study identified persistent issues related to recurring failures, non-trouble found (NTF) cases, inconsistent failure coding, and delayed claim resolution processes. The analysis also confirmed that alternator related failures were strongly associated with issues such as bearing degradation, voltage regulator malfunction, thermal exposure, and inconsistent maintenance practices.

The second objective, which involved the application of predictive and statistical analytical methods for warranty management improvement, was achieved through the implementation of machine learning and data driven analytical approaches. The predictive model demonstrated acceptable performance with $R^2 \geq 0.60$, indicating satisfactory predictive capability for real industrial warranty datasets characterised by operational variability and data uncertainty. The integration of predictive analytics enabled improved identification of high risk components, failure trends, and potential warranty escalation patterns. This demonstrates that analytical approaches can support

earlier intervention and more effective decision making compared to conventional reactive practices.

The third objective, which focused on developing an integrated visualisation and decision support platform, was achieved through the development of the Automotive Warranty Analysis System (AWAS) using Microsoft Power BI. The AWAS dashboard successfully integrated warranty claim trends, claim cost monitoring, repair lead time analysis, geographic segmentation, model based filtering, and root cause analysis into a single platform. The inclusion of time based parameters and operational indicators improved the interpretability of warranty behaviour throughout the product lifecycle, enabling more structured monitoring and evaluation. The system also improved transparency in warranty data analysis and supported evidence based decision making for quality improvement and warranty cost optimisation.

The findings of this research demonstrate that the integration of predictive analytics, engineering analysis, and business intelligence tools can significantly enhance warranty management effectiveness within the automotive industry. In addition to improving data visibility and analytical capability, the proposed framework contributes to proactive quality management by enabling earlier detection of recurring failures and supporting targeted corrective actions. The study further highlights the importance of data integrity, benchmarking, and performance threshold monitoring in ensuring reliable and meaningful warranty analysis.

Overall, this research provides both theoretical and practical contributions to the field of automotive warranty management. From an academic perspective, the study expands current understanding of data driven warranty analysis within real industrial environments. From an industrial perspective, the proposed AWAS framework offers a scalable and practical solution for improving warranty governance, product reliability, operational efficiency, and customer satisfaction within the Malaysian automotive industry.

5.2 Recommendations for Future Work

Although this research successfully developed and validated a data driven warranty management framework, several opportunities remain for further enhancement and industrial expansion. Future research should extend the analytical scope beyond alternator assemblies to include additional high risk automotive

components such as compressors, electronic control units, radiators, and steering systems. This would improve the generalisability and robustness of the proposed framework across broader vehicle systems and failure conditions.

Future studies should also explore the integration of real time data acquisition technologies such as Internet of Things (IoT) sensors, connected vehicle systems, and cloud based monitoring platforms. The incorporation of real time operational data would enable continuous condition monitoring and strengthen predictive maintenance capability by supporting earlier detection of abnormal component behaviour before failure occurrence. In addition, future work may incorporate advanced artificial intelligence techniques such as deep learning, ensemble learning, and adaptive forecasting models to improve prediction accuracy and anomaly detection performance. These approaches may further enhance the capability of the system to manage complex and highly variable warranty datasets in large scale automotive environments.

Another important area for future improvement involves strengthening data governance and standardisation across service centres, suppliers, and manufacturing operations. Improved standardisation of failure coding, diagnostic reporting, and repair documentation would reduce inconsistencies in warranty datasets and improve analytical reliability. This would also support more effective benchmarking and cross functional collaboration throughout the warranty management process.

Future research should additionally consider integrating financial and sustainability related indicators into the AWAS framework. This includes evaluating warranty cost impact relative to sales volume, carbon footprint associated with part replacement activities, and cost savings achieved through targeted repair strategies. Such integration would provide additional value for strategic planning and support organisational sustainability initiatives.

Finally, future studies may evaluate the implementation of AWAS within broader industrial ecosystems involving multiple manufacturers, suppliers, and service networks. Expanding the framework beyond a single organisation would strengthen its scalability, improve benchmarking capability, and establish a more comprehensive automotive warranty intelligence platform for the Malaysian automotive industry.

REFERENCES

[1] M Hamsani, "DNMY Warranty Claim Trend (5 Years)," Company XYZ Bi-Annually Report, Shah Alam, 2023.

[2] Road Transport Department, "Malaysia's Official Open Data Portal," 16 January 2024. [Online]. Available: https://data.gov.my/dashboard/car-popularity. [Accessed 08 Jun 2024].

[3] Alshurideh, Hikmat, Ra'ed , Haitham, Said, The Effect of Information Technology on Business and Marketing Intelligence Systems, Jordan: Springer International Publishing, 2023/2/9.

[4] Zainil Hanim Saidin, Wan Afezah Wan Abdul Rahman, Beni Widarman Yus Kelana, Rawiyah Abd Hamid, "Does Relationship Quality Affect Customer Loyalty of Malaysian National Carmakers: Alternative Attractiveness as Moderator," *International Journal of Supply Chain Management,* Vols. Vol 7, No 3, no. DOI: https://doi.org/10.59160/ijscm.v7i3.2264, p. 160–173, 2018.

[5] Jamaludin Akbar, Mohd Rizaimy Shaharudin, Siti Fairuza Hassam, Nurdiyana Nazihah Zainal, Amir Imran Zainoddin, "Quality of Service Practices within Business Market: An Automotive Industry Experience," *International Journal of Supply Chain Management,* Vols. Vol 7, No 6, no. https://doi.org/10.59160/ijscm.v7i6.2634, p. 204–220, 2018.

[6] Akhirahmad, M.; Nawi, M.N.M.; Taib, C.A.; Nor, A., "Design failure mode effect analysis and customer approach matrix: A case study in an automotive company," in *Proceedings of the 2nd Conference on Technology & Operations Management (2ndCTOM)*, Sintok, Kedah, Malaysia, 2018.

[7] M. F. Ahmad, Yeong Ting, Nik Hisyamudin Muhd Nor, Shiau Wei Chan, Mohd Fahrul Hassan, A.H Nor Aziati, "Comparative study of lean practices between Japanese and Malaysia automotive service centres," in *AIP Conference Proceedings*, Batu Pahat, Johor, Malaysia, October 2017.

[8] Nick Mikulec, Tim Felke and Shaunak Bangale, "Analysis of Warranty Data to Identify Improvements to Vehicle Reliability and Service Information," *SAE International Journal of Passenger Cars - Electronic and Electrical Systems,* Vols. Vol. 10, Issue 2, no. doi:10.4271/2017-01-1687, pp. 405-413, Aug. 2017.

[9] Kumar, Anil Luthra, Sunil Khandelwal, Dinesh Kumar Mehta, Rajneesh Chaudhary, Nityanand Bhatia, Sukhdev, "Measuring and improving customer retention at authorised automobile workshops after free services," *Journal of Retailing and Consumer Services,* vol. 39(C), no. DOI: 10.1016/j.jretconser.2017.07.007, pp. 93-102, 2017.

[10] Malaysian Automotive Association, "Media Release: Malaysian Automotive Industry Achieves Record Sales in 2023," Jan 2024.

[11] Corporation, DENSO, "DENSO focuses on quality to reduce warranty claims," [Online]. Available: https://www.denso.com/global/en/news/news-releases/2019/20191203-01/. [Accessed Jun 2024].

[12] DENSO Corporation, "DENSO's advanced technologies,," [Online]. Available: https://www.denso.com/global/en/innovation/advanced-technologies/. [Accessed June 2024].

[13] Shahanaghi Kamran, Noorossana Rassoul, Jalali Naini Seyed Gholamreza, Heydari Majeed, "Failure modeling and optimizing preventive maintenance strategy during two-dimensional extended warranty contracts," *Engineering Failure Analysis,,* vol. 28, no. https://doi.org/10.1016/j.engfailanal.2012.09.006, pp. 90-102, 2013-03-01.

[14] Yeu-Shiang Huang, Chih-Chiang Fang, Chia-Yu Kuo, "A study on two-dimensional warranty with a preventive maintenance policy under burn-in or run-in," *Applied Stochastic Models in Business and IndustryEarly View,* no. https://doi.org/10.1002/asmb.2859, 01 April 2024.

[15] Sunith Bandaru, Kalyan Deb, Vineet R. Khare, Rahul Chougule, "Quantitative modeling of customer perception from service data using evolutionary optimization," *GECCO '11: Proceedings of the 13th annual conference on Genetic and evolutionary computation,* no. https://doi.org/10.1145/2001576.2001813, pp. 1763 - 1770, July 2011.

[16] Pisal, S.M.M., Aris, I.B., Sahari, M.M., Seng, C.C., Bukhari, S.N.B., Mansor, N.F.B., Munir, M.F.B.M., "Moving towards sustainable human capital development for Malaysian automotive industry," in *2015 10th Asian Control Conference: Emerging Control Techniques for a Sustainable World, ASCC*

2015, Sutera Harbour ResortKota Kinabalu; Malaysia;, 8 September 2015, Article number 7244874.

[17] Anantha Raj A. Arokiasamy, Huam Hon Tat, "Assessing the relationship between service quality and customer satisfaction in the malaysian automotive insurance industry," *Middle East Journal of Scientific Research,* vol. 20(9), no. 1023-1030, p. DOI:10.5829/idosi.mejsr.2014.20.09.12029, January 2014.

[18] MG Maarof, S Sorooshian, SA Hamid, "What Drives A Successful Kaizen Implementation? An Insight from Malaysian Automotive Manufacturers," in *International Conference on Industrial Engineering and Operations Management*, Rabat, Morocco, April 11-13, 2017.

[19] N.M. Yusop, D.A. Wahab, N. Saibani, "Realising the automotive remanufacturing roadmap in Malaysia: challenges and the way forward," *Journal of Cleaner Production,* Vols. 112, Part 3, no. https://doi.org/10.1016/j.jclepro.2015.03.072, pp. 1910-1919, 20 January 2016.

[20] Peng Chan Chang, Chun-Ling Chuang, Wang Ching Chuang, Wei-Ching Lin, "An examination of the effects of quality on customer loyalty: The automobile industry in Taiwan," *Total Quality Management & Business Excellence,* Vols. 26(1-2):1-15, no. DOI:10.1080/14783363.2012.704268, pp. 203-217, August 2012.

[21] Malaysia Automotive Association, "Another Record Sales; Tiv 2023 Surpassed Previous Year High!," http//www.maa.org.my, Petaling Jaya, Selangor, 16 January 2024.

[22] R. Godwin-Jones, "Big data and language learning: Opportunities and challenges," *Emerging Technologies,* Vols. 25, Issue 1 , no. http://hdl.handle.net/10125/44747, p. 4–19, February 2021.

[23] S. Wu, "Warranty data analysis: A review," *Quality and Reliability Engineering International,,* vol. 28(8), no. DOI: https://doi.org/10.1002/qre.1282, pp. 795-805., 2012.

[24] K Teplická, S Khouri, T Mudarri, M Freňáková, "Improving the Quality of Automotive Components through the Effective Management of Complaints in Industry 4.0," *Applied Sciences,* vol. 13, no. 14, p. https://doi.org/10.3390/app13148402, 2023.

[25] Fernanda Campos Bueno, José Carlos de Toledo, Sérgio Luis da Silva, "The role of DFMEA in the integration between product development and quality engineering: cases in large companies," *Product Management & Development,* vol. 17, no. http://dx.doi.org/10.4322/pmd.2019.009, pp. 85-95, 2019.

[26] N Hissham AH, N Fazli AM, M Fauzi I, M Ilham K, A Malek AW,, "Development of a Visualization Framework for Automotive Warranty Data Management”," *Journal of Computing Research and Innovation,* Vols. vol. 11, no. 1,, no. doi:10.24191/jcrinn.v11i1.624, pp. pp. 157–172,, Mar. 2026.

[27] Fikri Dweiri, Sameer Kumar, Sharfuddin Ahmed Khan, Vipul Jain, "Designing an integrated AHP based decision support system for supplier selection in automotive industry," *Expert Systems with Applications,* vol. 62, no. DOI:10.1016/j.eswa.2016.06.030, pp. 273-283, 15 November 2016.

[28] V. Mani, Rajat Agrawal, Vinay Sharma, "Supplier selection using social sustainability: AHP based approach in India," *International Strategic Management Review,* Vols. Volume 2, Issue 2,, no. https://doi.org/10.1016/j.ism.2014.10.003, pp. 98-112, 2014.

[29] Muhammad Jawad Akhtar, Afshan Naseem, Faaiz Ahsan, "A novel hybrid approach to explore the interaction among faults in production process with extended FMEA model using DEMATEL and cloud model theory," *Engineering Failure Analysis,* vol. 157, p. https://doi.org/10.1016/j.engfailanal.2023.107876, 2024.

[30] Ionica Oncioiu, Ovidiu Constantin Bunget, Mirela Cătălina Türkeș, Sorinel Căpușneanu, Dan Ioan Topor, Attila Szora Tamaș, Ileana-Sorina Rakoș and Mihaela Ștefan Hint, "The Impact of Big Data Analytics on Company Performance in Supply Chain Management," *Sustainability,* vol. 11(18):4864, no. https://doi.org/10.3390/su11184864, 2019.

[31] Mohammad Babakmehr, Sascha Baumanns, Abdallah Chehade, Thomas Hochkirchen, Mahdokht Kalantari, Vasiliy Krivtsov, David Schindler, "Data-driven framework for warranty claims forecasting with an application for automotive components," *Engineering Reports,* Vols. 6,5, no. https://doi.org/10.1002/eng2.12764, 2023.

[32] Muhammad Anshari, Mohammad Nabil Almunawar, Syamimi Ariff Lim, Abdullah Al-Mudimigh, "Customer relationship management and big data enabled: Personalization & customization of services," *Applied Computing and Informatics,* vol. 15, no. https://doi.org/10.1016/j.aci.2018.05.004, pp. 94-101, 2019..

[33] F Scherpen, A Draghici, J Niemann, "Customer experience management to leverage customer loyalty in the automotive industry," *Procedia - Social and Behavioral Sciences,* vol. 238, no. https://doi.org/10.1016/j.sbspro.2018.04.014, pp. 374-380, 2018.

[34] Aysegul Ucar, Mehmet Karakose, Necim Kırımça, "Artificial Intelligence for Predictive Maintenance Applications: Key Components, Trustworthiness, and Future Trends," *Applied Sciences,* vol. 14(2), no. https://doi.org/10.3390/app14020898, p. 898, 2024.

[35] Andreas Theissler, Judith Pérez-Velázquez, Marcel Kettelgerdes, Gordon Elger, "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry," *Reliability Engineering & System Safety,* vol. 215, no. https://doi.org/10.1016/j.ress.2021.107864, 2021.

[36] Chia-Fen Chi, Davin Sigmund, Martin Octavianus Astardi, "Classification Scheme for Root Cause and Failure Modes and Effects Analysis (FMEA) of Passenger Vehicle Recalls," *Reliability Engineering & System Safety,* vol. 200, no. https://doi.org/10.1016/j.ress.2020.106929, 2020.

[37] Karina Cecilia Arredondo-Soto, Alejandro Jiménez-Zaragoza, Marco Augusto Miranda-Ackerman,Julio Blanco-Fernández, Alejandra García-Lechuga, Guadalupe Hernández-Escobedo,Jorge Luis García-Alcaraz, "Design and Repair Strategies Based on Product–Service System and Remanufacturing for Value Preservation," *Sustainability,* vol. 14 (14), no. https://doi.org/10.3390/su14148560, 2022.

[38] Amel Ben Mabrouk, Anis Chelbi, "Joint preventive maintenance and extended warranty strategy for leased unreliable equipment submitted to imperfect repair at failure," *IFAC-PapersOnLine,* Vols. 55, 10, no. https://doi.org/10.1016/j.ifacol.2022.09.553, pp. 1201-1206, 2022.

[39] Suzaituladwini Hashim, Nurul Fadly Habidin, Juriah Conding, Nurzatul Ain , Seri Lanang Jwaya, Anis Fazdlin Mohd Zubir, "Total Productive Maintenance and Innovation Performance in Malaysian Automotive Industry," *International Journal of Engineering Research and Development,* vol. 3(11), no. e-ISSN: 2278-067X, p-ISSN: 2278-800X,, pp. 62-67, 2012.

[40] Krivtsov, Vasiliy Vasilievich, "A Monte Carlo approach to modeling and estimation of the generalized renewal process in repairable system reliability analysis," *Systems Science and Systems Engineering,* vol. 14, no. 2, 2023.

[41] N Haris, A K Junoh,W Z A Wan Muhamad, "Warranty Claim Quantity Forecasting via Multiple Linear Regressions," in *1st International Conference on Engineering and Technology, ICoEngTech 2021*, Perlis, 2021.

[42] Shaomin Wu, Artur Akbarov, "Support vector regression for warranty claim forecasting," *European Journal of Operational Research,* Vols. Volume 213, Issue 1 ,, no. https://doi.org/10.1016/j.ejor.2011.03.009, pp. 196-204, 2011.

[43] Ammar Y. Alqahtani, Surendra M. Gupta, "Warranty and Preventive Maintenance Analysis for Sustainable Reverse Supply Chains," *Journal of Management Science and Engineering,* Vols. 2, 1,, no. https://doi.org/10.3724/SP.J.1383.201004, pp. 69-94, 2017.

[44] Izzet Sahin, Hakan Polatoglu, Quality, warranty and preventive maintenance, Vols. vol. 4, no. 2, pp. 78-89, New York: Springer Science & Business Media, 2012, p. DOI: 10.2345/jme.2019.0002.

[45] Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., & Elger, G, "Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry," *Reliability engineering & system safety,,* vol. 215, no. DOI: https://doi.org/10.1016/j.ress.2021.107864, p. 107864., 2021.

[46] Nacchia, M., Fruggiero, F., Lambiase, A., & Bruton, K, "A Systematic Mapping of the Advancing Use of Machine Learning Techniques for Predictive Maintenance in the Manufacturing Sector," *Applied Sciences,* vol. 11(6), no. DOI: https://doi.org/10.3390/app11062546, p. 2546, 2021.

[47] Paula Fraga-Lamas, Tiago M. Fernández-Caramés, "A Review on Blockchain Technologies for an Advanced and Cyber-Resilient Automotive Industry," *ieeexplore.ieee.org,* vol. 7, 2019.

[48] Liao, Albert, Warranty Chain Management, Digitalization and Sustainability, Vols. vol. 5, no. 4, pp. 132-146,, Springer, 2022 , p. DOI: 10.4567/wmr.2021.0056.

[49] Salizanah, Muhamad Lutfi,, "2023 Cost Sharing Activities," in *Supplier Briefing 2023 Cost Sharing Activity*, Rawang, 2023.

[50] Yihai He, Changchao Gu, Zhenzhen He, Jiaming Cui, "Reliability-oriented quality control approach for production process based on RQR chain," *Total Quality Management & Business Excellence,* Vols. 29, 5-6, no. https://doi.org/10.1080/14783363.2016.1224086, pp. 652-672, 2018.

[51] Jelena Kurilova-Palisaitiene, Erik Sundin, Bonnie Poksinska, "Remanufacturing challenges and possible lean improvements," *Journal of Cleaner Production,* vol. 172, no. https://doi.org/10.1016/j.jclepro.2017.11.023, pp. 3225-3236, 2018.

[52] Jalal Ashayeri, Ning Ma , Renata Sotirov, "The redesign of a warranty distribution network with recovery processes," *Transportation Research Part E: Logistics and Transportation Review,* vol. 77, no. https://doi.org/10.1016/j.tre.2015.02.017, pp. 184-197, 2015.

[53] Hairi, Mohd Hamsani Mohd, "DNMY GWAMS data analysis," Glenmarie, Shah Alam, 2023.

[54] Alessandro Manello, Giuseppe Calabrese, "The influence of reputation on supplier selection: An empirical study of the European automotive industry," *Journal of Purchasing and Supply Management,* Vols. 25,1, no. https://doi.org/10.1016/j.pursup.2018.03.001, pp. 69-77, 2019.

[55] JD Wisner, KC Tan, K Leong, Principles of supply chain management: A balanced approach, Vols. vol. 8, no. 3, pp. 102-117,, United States of America: SOUTH-WESTERN CENGAGE Learning, 2021, p. DOI: 10.3456/ijscm.2020.0045..

[56] Kamran Rashid, M.M. Haris Aslam, "Business excellence through total supply chain quality management," *Asian Journal on Quality,* vol. 13, no. DOI: https://doi.org/10.1108/15982681211287829, pp. 309-324, 2012.

[57] Luz María Valdez-de la Rosa, Luis Alberto Villarreal-Villarreal, Gustavo Alarcón-Martínez, "Quality and innovation as drivers for manufacturing

competitiveness of automotive parts suppliers," *The TQM Journal,* Vols. 33,5, no. DOI: https://doi.org/10.1108/TQM-12-2018-0218, pp. 966-986, 2021.

[58] Yuan Wang, Sachin B. Modi, Tobias Schoenherr, "Leveraging sustainable design practices through supplier involvement in new product development: The role of the suppliers' environmental management capability," *International Journal of Production,* vol. 232, no. https://doi.org/10.1016/j.ijpe.2020.107919, 2021.

[59] Huan, P. A. N. G., Dongyang, S. H. I., Daocheng, W. A. N. G., & Jingyi, L. I. U., "Automobile reliability evaluation and automobile-enterprise operational decision system based on warranty data," *Journal of Automotive Safety and Energy,,* vol. 13(2), no. DOI: 10.3969/j.issn.1674-8484.2022.02.004,, p. 250., 2022.

[60] Zhonghui, S., Yanying, G., Zhonghong, S., Shouchen, Y., & Baoyu, H., "Maintenance cost prediction for the vehicle based on maintenance data," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering,,* vol. 238(6), no. DOI: https://doi.org/10.1177/09544070221147080, pp. 1550-1556, 2024.

[61] E Fadda, G Perboli, M Rosano, JE Mascolo, D Masera, "A Decision Support System for Supporting Strategic Production Allocation in the Automotive Industry," *Sustainability,* vol. 14 (4), no. DOI: https://doi.org/10.3390/su14042408, 2022.

[62] Fabio Arena, Mario Collotta, Liliana Luca, Marianna Ruggieri,Francesco Gaetano Termine, "Predictive Maintenance in the Automotive Sector: A Literature Review," *Mathematical and Computational Applications,* vol. 27, no. https://doi.org/10.3390/ mca27010002, p. 2, 2022.

[63] Dnyanesh Rajpathak, Yiming Xu, Ian Gibbs, "An integrated framework for automatic ontology learning from unstructured repair text data for effective fault detection and isolation in automotive domain," *Computers in Industry,* vol. 123, no. https://doi.org/10.1016/j.compind.2020.103338, p. 103338, 2020.

[64] I. Institute for Business Value, [Online]. Available: https://www.ibm.com/thought-leadership/institute-business-value/en-us/report/data-powered-warranty. [Accessed 04 Jan 2026].

[65] Deloitte, “deloitte-uk-automotive-analytics,”, 2015, 2015. [Online]. Available: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/manufacturing/deloitte-uk-automotive-analytics.pdf. [Accessed 2026 Jan 4].

[66] Automotive Survey Report Partner organizations, "Global Automotive Warranty Survey Report Commercial Services," Bearing Point, 2008. [Online]. Available: https://www.bearingpoint.com/files/autowarrantyreport_final_web.pdf. [Accessed 4 Jan 2026].

[67] Wahab, D. A., Manan, N. F. A., Hannan, M. A., Abdullah, S., & Hussain, A. (2008)., "Designing for comfort and reliability in an intelligent car seat.," *American Journal of Applied Sciences,,* Vols. 5(12),, p. 1787–1792., 2008.

[68] N. Haris, A. K. Junoh, and W. Z. A. Wan Muhamad,, "Warranty Claim Quantity Forecasting via Multiple Linear Regressions," *Journal of Physics, Conference Series, IOP Publishing Ltd,,* no. doi: 10.1088/1742-6596/1962/1/012043, Jul. 2021.

[69] C. López, M. F. Morales-Contreras, I. M. Langella, and J. Alonso-Monge,, "Modeling supply chain disruptions due to geopolitical Reasons: A systematic literature review," *Transp Res E Logist Transp Rev,* Vols. 202,, no. doi: 10.1016/j.tre.2025.104290., Oct. 2025.

[70] A. A. Wilson, A. Nehme, A. Dhyani, and K. Mahbub,, "A Comparison of Generalised Linear Modelling with Machine Learning Approaches for Predicting Loss Cost in Motor Insurance," *Risks,,* Vols. vol. 12, no. 4,, no. doi: 10.3390/risks12040062., Apr. 2024,.

[71] Samsudin, S. N., Abdullah, B., Yusoff, N., "Reliability Simulation Mechanism Model of Big Data Mileage Prediction for Automotive Warranty,," *International Journal of Intergrated Engineering,,* Vols. Vo. 17, No. 2, no. DOI: https://doi.org/10.30880/ijie.2025.17.02.017, pp. 208-221,, 2025.

[72] A. Mukhtar, B. Hofer, D. Jannach, and F. Wotawa,, "Spreadsheet debugging: The perils of tool over-reliance," *Journal of Systems and Software,,* Vols. vol. 184,, no. doi: 10.1016/J.JSS.2021.111119., pp. 111119,, Feb. 2022,.

[73] Y. C. Y. Y. H. C. J. H. a. S. W. W. Bao, "Data-driven stock forecasting models based on neural networks: A review," *Information Fusion,* vol. 113, no. doi: 10.1016/J.INFFUS.2024.102616, p. 102616, Jan. 2025.

[74] A. H. R. P. S. a. A. K. S. M. Javaid, "Digital economy to improve the culture of industry 4.0: A study on features, implementation and challenges," ," *Green Technologies and Sustainability,,* Vols. vol. 2, no. 2,, no. doi: 10.1016/J.GRETS.2024.100083., p. p. 100083, May 2024,.

[75] N. Hissham AH, M. Fauzi I, M. Ilham K, A. Malek AW, N. Fazli AM,, "Improvement Product Design Quality to Reduce Warranty Cost in Aftermarket Sales for OEM's Customer,," *Journal Mechanical Engineering,* Vols. Vol. 22, Issue 3,, no. doi.org/10.24191/jmeche.v22i3.5494, pp. page 1-15,, 15 September 2025.

[76] U. R. S. Yalavarthy, N. B. Kumar, A. R. V. Babu, R. P. Narasipuram, and S. Padmanaban, "Digital twin technology in electric and self-navigating vehicles: Readiness, convergence, and future directions,"," *Energy Conversion and Management,* Vols. vol. 26,, no. doi: 10.1016/J.ECMX.2025.100949, p. 100949, April 2025.

[77] N. Hissham AH, M. Hakim R, M. Fauzi I, F. Jika, N. Fazli AM,, "Development of a Visualization Framework for Automotive Warranty Data Management,," in *CSSR2025_12th International Conference on Science & Social Research 2025*, Shah Alam, 05 November 2025.

[78] J. F. Dollinger and V. Vauchey, "Towards efficient program execution on edge-cloud computing platforms," *Journal of Parallel and Distributed Computing,,* Vols. 205,, no. doi: 10.1016/J.JPDC.2025.105135., pp. 105135,, Nov. 2025,.

[79] J. López Espejel, M. S. Yahaya Alassan, M. Bouhandi, W. Dahhane, and E. H. Ettifouri,, "Low-cost language models: Survey and performance evaluation on Python code generation," *Eng Appl Artif Intell,,* Vols. 140,, no. doi: 10.1016/J.ENGAPPAI.2024.109490, pp. 109490,, Jan. 2025,.

[80] A. Malviya, A. Udhani, and S. Soni,, "R-Tool: Data analytic framework for big data," *2016 Symposium on Colossal Data Analysis and Networking, CDAN,* no. doi: 10.1109/CDAN.2016.7570960, Sep. 2016,.

[81] Heyn, H. M., Habibullah, K. M., Knauss, E., Horkoff, J., Borg, M., Knauss, A., & Li, P. J., "Automotive Perception Software Development: An Empirical Investigation into Data, Annotation, and Ecosystem Challenges,," *Proceedings - 2023 IEEE/ACM 2nd International Conference on AI Engineering - Software Engineering for AI, CAIN 2023,* Vols. 13-24, no. DOI: https://doi.org/10.48550/arXiv.2303.05947, 2023.

[82] J. Brodny and M. Tutak, "Stakeholder interactions and ethical imperatives in big data and AI development,," *Journal of Open Innovation: Technology, Market, and Complexity,,* Vols. vol. 11, no. 1,, no. doi: 10.1016/J.JOITMC.2025.100491., p. p. 100491, Mar. 2025,.

[83] Theodore T. Allena, Zhenhuan Sui, Kaveh Akbari, "Exploratory text data analysis for quality hypothesis generation," *Quality Engineering,* vol. 30, no. DOI 10.1080/08982112.2018.1481216, pp. 701-712, 2018.

[84] V. Kotu and B. Deshpande, "Regression Methods," *Data Science,,* no. doi: 10.1016/B978-0-12-814761-0.00005-8., p. pp. 165–197, Jan. 2019.

[85] R. Swarnkar, R. Sawant, R. Harikrishnan, and P. Srideviponmalar,, "Multiple Linear Regression Algorithm-based Car Price Prediction," *Proceedings of the 3rd International Conference on Artificial Intelligence and Smart Energy, ICAIS 2023,,* no. doi: 10.1109/ICAIS56108.2023.10073882., pp. pp. 675–681,, 2023,.

[86] Zhao, J., Jiang, Z. Z., & Sun, M., "A review on the gray market business using a data-driven approach.," *Data Science and Management,,* Vols. 1(1),, no. https://doi.org/10.1016/j.dsm.2021.02.003, p. 38–47., 2021.

[87] J. A. Fernandes de Souza, M. M. Silva, S. G. Rodrigues, and S. Machado Santos,, "A forecasting model based on ARIMA and artificial neural networks for end–OF–life vehicles,"," *J Environ Manage,,* Vols. vol. 318,, no. doi: 10.1016/J.JENVMAN.2022.1156, pp. p. 115616,, Sep. 2022,.

[88] J. D. Cummins and G. L. Griepentrog, "Forecasting automobile insurance paid claim costs using econometric and ARIMA models,"," *Int J Forecast,,* Vols. vol. 1, no. 3,, no. doi: 10.1016/0169-2070(85)90003-2., pp. pp. 203–215,, 1985,.

[89] N. Haris, A. K. Junoh, S. Ibrahim, A. Ahmed, W. Z. Azman, and W. Muhamad,, "Warranty Claim Cost Forecasting based on ARIMA Box-Jenkins Approach,"

2021. [Online]. Available: Available: http://myjms.mohe.gov.my/index.php/ijbtm. [Accessed 04 Jan 2026].

[90] L. Breiman, "Random forests,," *Mach Learn,,* Vols. vol. 45, no. 1,, no. doi: 10.1023/A:1010933404324/METRICS., pp. pp. 5–32,, Oct. 2001,.

[91] Hasan Ahmed Salman, Ali Kalakech, Amani Steiti,, "Random Forest Algorithm Overview,," *Babylonian Journal of Machine Learning,,* Vols. Vol. 2024,, no. DOI: https://doi.org/10.58496/BJML/2024/007, pp. 69-79,, 08 June 2024,.

[92] Becker, Thijs; Rousseau, Axel-Jan; Geubbelmans, Melvin; Burzykowski, Tomasz; Valkenborg, Dirk, "Decision trees and random forests," *American Journal of Orthodontics and Dentofacial Orthopedic,* vol. 164, no. DOI: 10.1016/j.ajodo.2023.09.011, pp. 894-897, 2023.

[93] Martin Becker, Simon Schölzel, "Warranty Provisions: Machine-Learning Versus Human Estimates," *European Accounting Review,* vol. 34 (5), no. https://doi.org/10.1080/09638180.2024.2444521, pp. 1945-1973, 01 Jan 2025.

[94] Nirmawana Simarmata, Ketut Wikantika, Trika Agnestasia Tarigan, Muhammad Aldyansyah , Rizki Kurnia Tohir, Adam Irwansyah Fauzi, Anggita Rahma Fauzia, "Comparison of random forest, gradient tree boosting, and classification and regression trees for mangrove cover change monitoring using Landsat imagery," The Egyptian," *Journal of Remote Sensing and Space Sciences,,* vol. 28, no. DOI: https://doi.org/10.1016/j.ejrs.2025.02.002, pp. 138-150, March 2025,.

[95] Azal Ahmad Khan, Omkar Chaudhari, Rohitash Chandra, "A review of ensemble learning and data augmentation models for class imbalanced problems: Combination, implementation and evaluation," *Expert Systems with Applications,* vol. 244, no. https://doi.org/10.1016/j.eswa.2023.122778, p. 122778, 15 June 2024.

[96] Tianqi Chen, Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," *ACM,* no. DOI: http://dx.doi.org/10.1145/2939672.2939785, pp. 785-794, 2016,.

[97] Nedaa Thamer Qassem, Ibrahim Ahmed Saleh, "Prediction of Software Cost Estimation Using Stacking Ensemble Learning method," *Procedia Computer*

Science, vol. 258, no. https://doi.org/10.1016/j.procs.2025.04.305, pp. 727-736, 2025.

[98] David Opeoluwa Oyewola, Emmanuel Gbenga Dada, Juliana Ngozi Ndunagu, "A novel hybrid walk-forward ensemble optimization for time series cryptocurrency prediction," *A Cell Press Journal,* vol. 8 (11), no. DOI: 10.1016/j.heliyon.2022.e11862, November 2022.

[99] Wael Hassanieh, Abdallah Chehade, Vasiliy Krivtsov,, "Deep auto-encoded Conditional Gaussian Mixture Model for warranty claims forecasting," *Reliability Engineering & System Safety,* vol. 263, no. https://doi.org/10.1016/j.ress.2025.111230, p. 111230, November 2025.

[100] Chairote Yaiprasert, Achmad Nizar Hidayanto, "AI-powered ensemble machine learning to optimize cost strategies in logistics business," *International Journal of Information Management Data Insights,* vol. 4 (1), no. DOI: https://doi.org/10.1016/j.jjimei.2023.100209, p. 100209, April 2024,.

[101] Teköz, Begüm., "Forecasting Warranty Claims for Month in Services Groups in Automotive Sector," Middle East Technical University, August 2022.

[102] H. Dalianis, "Evaluation Metrics and Evaluation,"," *Clinical Text Mining, Cham: Springer International Publishing,,* no. doi: 10.1007/978-3-319-78503-5_6., p. 45–53., 2018,.

[103] Stijn Hawinkel, Willem Waegeman, Steven Maere, "Out-of-Sample R2: Estimation and Inference," *The American Statistician,,* vol. 78, no. https://doi.org/10.1080/00031305.2023.2216252, p. 1, 2024.

[104] Deng, Ai, "Time series cross validation: A theoretical result and finite sample performance," *Economics Letters,* Vols. Volume 233,, no. https://doi.org/10.1016/j.econlet.2023.111369, p. 111369, December 2023,.

[105] Pan, Xing; Zhang, Manli; Chen, Xi, "A Method of Quality Improvement Based on Big Quality Warranty Data Analysis," *2018 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C),* no. DOI: 10.1109/QRS-C.2018.00115, 16-20 July 2018.

[106] Wu, Shaomin, "Warranty claim analysis considering human factors," *Reliability Engineering & System Safety,* vol. 96(1), no. DOI: https://doi.org/10.1016/j.ress.2010.07.010, pp. 131-138., 2011.

[107] Dr. Rong Luo, Berenica Vejvoda, "Further Adventures In data Cleaning: Working With Data In Excel and R," *Research Data Management in the Canadian Context,* no. DOI: https://doi.org/10.5206/DBAS6486, 2023.

[108] Kalaivany Natarajan, Jiuyong Li, Andy Koronios, "Data mining techniques for data cleaning," *Engineering Asset Lifecycle Management,* vol. 796–804, no. https://doi.org/10.1007/978-0-85729-320-6_91, 2010.

[109] Harode, A.; Ensafi, M.; Thabet, W., "Linking BIM to Power BI and HoloLens 2 to Support Facility Management: A Case Study Approach," *Buildings,* vol. 12 (6), no. DOI: https://doi.org/10.3390/buildings12060852, p. 852, 2022.

[110] Nurdin, A. A., Salmi, G. N., Sentosa, K., Wijayanti, A. R., & Prasetya, A., "Utilization of Business Intelligence in Sales Information Systems," *Journal of Information System Exploration and Research,* vol. 1 (1), no. DOI: https://doi.org/10.52465/joiser.v1i1.101, pp. 39-48, January 2023.

[111] Mitra, A., & Resnick, S. I., "Modeling Total Expenditure on Warranty Claims.," *Stochastic Models,,* Vols. 27(3),, no. DOI: https://doi.org/10.1080/15326349.2011.593406, p. 431–473., 2011.

[112] Jiayi Yan, Qiuchen Lu, Nan Li, Long Chen, Michael Pitt,, "Common data environment for digital twins from building to city levels,," *Automation in Construction,,* Vols. 174,, no. DOI: https://doi.org/10.1016/j.autcon.2025.106131., p. 106131, 2025,.

[113] N Hissham AH, N Fazli AM, M Fauzi I, A Malek AW, M Ilham K, "Development of Warranty Management Visualisation for Automotive Aftermarket with Integration Of AI," *Journal of Mechanical Engineering,* vol. 23 (1), no. https://doi.org/10.24191/jmeche.v23i1.5527, pp. 19-38, 15 January 2026.

[114] Peças, P., Encarnação, J., Gambôa, M., Sampayo, M., & Jorge, D., "PDCA 4.0: A New Conceptual Approach for Continuous Improvement in the Industry 4.0 Paradigm.," *Applied Sciences,,* vol. 11(16), no. DOI: https://doi.org/10.3390/app11167671, p. 7671, 2021.

[115] Xiao Liu, Rong Pan, "Analysis of Large Heterogeneous Repairable System Reliability Data With Static System Attributes and Dynamic Sensor

Measurement in Big Data Environment," *Technometrics,,* vol. 62 (2), no. DOI: https://doi.org/10.1080/00401706.2019.1609584, 16 Jul 2019.

[116] Arnold R, Chukova S, Hayakawa Y, Marshall S., "Warranty cost analysis with an alternating geometric process," *Warranty cost analysis with an alternating geometric process,* vol. 233(4), no. DOI: https://doi.org/10.1177/1748006X18820379, pp. 698-715, 6 January 2019.

[117] Bernard, G., Budde, L., Hänggi, R., & Friedli, T., "Driving continuous improvement with industry 4.0 technologies: lessons from multiple use case analysis.," *Applied Sciences,,* Vols. 15(4),, no. DOI: https://doi.org/10.3390/app15042191, p. 2191., 2025.

[118] Picado Argüello, B., & González-Prida, V., "Integrating Change Management with a Knowledge Management Framework: A Methodological Proposal.," *Information,,* Vols. 15(7),, no. DOI: https://doi.org/10.3390/info15070406, p. 406., 2024.

[119] Wang, F. K., "Quality evaluation of a manufactured product with multiple characteristics," *Quality and reliability engineering international,,* vol. 22(2), no. DOI: https://doi.org/10.1002/qre.712, pp. 225-236, 2006.

[120] J Jeon, SY Sohn, "Product failure pattern analysis from warranty data using association rule and Weibull regression analysis: A case study," *Reliability Engineering & System Safety,* vol. 133, no. DOI: https://doi.org/10.1016/j.ress.2014.08.015, pp. 176-183, 2015.

[121] I. Makarova, P. Buyvol, L. Gabsalikhova, A. Pashkevich, E. Tsybunov, and A. Boyko,, "Improving the reliability of autonomous vehicles in a branded service system using big data," *21st Int. Conf. Res. Educ. Mechatronics, REM 2020,,* Vols. 19,, no. DOI: 10.1109/REM49740.2020.9313084., pp. 19–24,, 2020.

[122] Srinivasan, R., Manivannan, S., Ethiraj, N., Devi, S. P., Kiran, S. V., "Modelling an Optimized Warranty Analysis methodology for fleet industry using data mining clustering methodologies with Fraud detection mechanism using pattern recognition on hybrid analytic approach.," *Procedia Computer Science,,* vol. 87, no. DOI: https://doi.org/10.1016/j.procs.2016.06.001, pp. 322-327., 2016.

[123] Murthy, D. N. P., & Djamaludin, I., "New product warranty: A literature review.," *International Journal of production economics,,* vol. 79(3), no. DOI: https://doi.org/10.1016/S0925-5273(02)00153-6, pp. 231-260., 2002.

[124] Wu, Shaomin, "Warranty data analysis: a review," *Quality and Reliability Engineering International,,* vol. 28(8), no. https://doi.org/10.1002/qre.1282, pp. 795-805., 2012.

[125] Hazucha, P., Moon, S. T., Schrom, G., Paillet, F., Gardner, D., Rajapandian, S., & Karnik, T., "High voltage tolerant linear regulator with fast digital control for biasing of integrated DC-DC converters.," *IEEE Journal of Solid-State Circuits,* vol. 42(1), no. DOI: https://doi.org/10.1109/JSSC.2006.885060, pp. 66-73., 2006.

[126] Torgunov, D., Trundle, P., Campean, F., Neagu, D., & Sherratt, A., "Vehicle Warranty claim prediction from diagnostic data using classification," *UK Workshop on Computational Intelligence,* no. DOI: https://doi.org/10.1007/978-3-030-29933-0_40, pp. 483-492, August 2019.

[127] Ebrahimi, A., & Mojtahedi, S., "Pattern analysis of auto parts failures in the after-sales service network; an interconnected approach of association rules mining and Bayesian networks in the automotive industry.," *International Journal of Quality & Reliability Management,,* vol. 41(4), no. DOI: https://doi.org/10.1108/IJQRM-02-2023-0031, pp. 1185-1207., 2024.

[128] Adomah, R., Alwada'n, T., Al Masarweh, M, "Enhanced Graphical Representation of Data in Web Application (Case Study: Covid-19 in the UK).," *International Journal of Advanced Computer Science and Applications,,* vol. 12(9), no. DOI:10.14569/IJACSA.2021.0120968, 2021.

[129] Li, G., Zhu, X., Wang, J., Wu, D., & Li, J., "Using lda model to quantify and visualize textual financial stability report.," *Procedia computer science,,* vol. 122, no. DOI: https://doi.org/10.1016/j.procs.2017.11.382, pp. 370-376., 2017.

[130] Himthani, P., Dubey, G. P., Sharma, B. M., & Taneja, A., "Big data privacy and challenges for machine learning.," *In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC),* no. DOI: 10.1109/I-SMAC49090.2020.9243527, pp. 707-713, 2020.

[131] Rai, B., & Singh, N, "A modeling framework for assessing the impact of new time/mileage warranty limits on the number and cost of automotive warranty claims.," *Reliability Engineering & System Safety,* vol. 88(2), no. DOI: https://doi.org/10.1016/j.ress.2004.07.006, pp. 157-169., 2005.

[132] Kaisler, S., Armour, F., Espinosa, J. A., & Money, W., "Big data: Issues and challenges moving forward.," *46th Hawaii international conference on system sciences,* no. DOI: 10.1109/HICSS.2013.645, pp. 995-1004, January 2013.

[133] Ismail, M. F., Hamid, N. H. A., Basri, M. H. M., & Manan, N. F. A., "Improving Automotive Accessory Development: A Warranty-Based Analysis of Material Defects and Process Enhancement," *Journal of Applied Engineering Design and Simulation,* vol. 5(2), no. DOI: https://doi.org/10.24191/jaeds.v5i2.112, pp. 1-19., 2025.

[134] Wadhera, S., Kamra, D., Kumar, A., Jain, A., & Jain, V., "A systematic Review of Big data tools and application for developments.," *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM),* no. DOI: 10.1109/ICIEM51511.2021.9445326, pp. 561-566, 2021.

[135] Liu, P., Wang, D., Zhang, Y., Ren, Y., & Ren, J., "A method for integrating operation and maintenance data of power information systems based on artificial intelligence technology," *Third International Conference on Machine Vision, Automatic Identification, and Detection (MVAID 2024),* vol. 13230, no. DOI: https://doi.org/10.1117/12.3036521, pp. 765-770, August 2024.

[136] Kristiyanto, D. Y., Sarno, R., Sungkono, K. R., Handayani, R., & Shepard, A, "Comprehensive Framework for Transitioning Monolithic to Microservices in MVC Context.," *In 2024 3rd International Conference on Creative Communication and Innovative Technology (ICCIT),* no. DOI: 10.1109/ICCIT62134.2024.10701144, pp. 1-7, August 2024.

[137] Ahmed, M. K., Umar, G. A., & Dawaki, M, "Development of a new minimalistic PHP micro framework: NOAPHP.," *Bima Journal of Science and Technology,,* vol. 8(3A), no. DOI: 10.56892/bima. v8i3A.810, 2024, pp. 262-269., 2024.

[138] Bi C, Nichols T, Lee H, Yang Y, Ye Z, Pan Y, Hong E, Kochunov P, Chen S., "BNPower: a power calculation tool for data-driven network analysis for

whole-brain connectome data," *Imaging Neuroscience.,* vol. 2, no. DOI: https://doi.org/10.1162/imag_a_00099, pp. 1-13., February 2024.

[139] S. N. A. B. &. Y. N. Samsudin, "Customer satisfaction and service experience in big data analytics for automotive service advisor.," *In 2022 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS),* no. DOI: 10.1109/I2CACIS54679.2022.9815482, pp. 84-89, June 2022.

[140] Wu, Shaomin, "A review on coarse warranty data and analysis," *Reliability Engineering & System Safety,* vol. 114, no. DOI: https://doi.org/10.1016/j.ress.2012.12.021, pp. 1-11, 2013.

[141] R Karim, K Suzuki, "Analysis of warranty claim data: a literature review.," *International Journal of Quality & Reliability Management,,* vol. DOI: https://doi.org/10.1108/02656710510610820, no. 22(7), pp. 667-686., 2005.

[142] Mohd Hanafi Azman Ong, Zamalia Mahmud, "An Exploratory Factor Analysis of Market Survey Instruments for Automobile Industry: A study on Malaysian Motor Vehicle Industry," *Thailand Statistician, 2020; 18(4): 481-490,* 2019.

[143] Wikipedia The Free Encyclopedia, [Online]. Available: https://en.wikipedia.org/wiki/Cronbach%27s_alpha. [Accessed 23 July 2024].

[144] De Battisti, F., Nicolini, G., & Salini, S., "The Rasch Model," *Modern Analysis of Customer Surveys: with applications using R,,* no. DOI: https://doi.org/10.1002/9781119961154.ch14, pp. 259-281., 2011.

[145] Lee Kwee Yoon, Cindy, "DNMY Customer Satisfaction Survey 2021 ~ 2023," Sales Purchase Administration, Shah Alam, 2023.

[146] Funing Xu, Ning Ding, Nan Li, Long Liu, Nan Hou, Na Xu, Weimin Guo, Linan Tian, Huixia Xu, Chi-Man Lawrence Wu, Xiaofeng Wu, Xiangfeng Chen, "A review of bearing failure Modes, mechanisms and causes," *Engineering Failure Analysis,* vol. 152, no. https://doi.org/10.1016/j.engfailanal.2023.107518, 2023.

[147] Munehiko Itoh, Atsumi Kato, Yoshitaka Shimono, Yasuhiko Haraguchi & Park Taehoon, "Supplier Development of Japanese Automotive Parts Suppliers—Purchasing Strategy of Denso in ASEAN," *SpringerBriefs in Economics,* no. https://doi.org/10.1007/978-981-13-2360-7_5, pp. 63-80, 2018.

[148] Shokouhyar, Ahmadi, Ashrafzadeh,, "Promoting a novel method for warranty claim prediction based on social network data,," Reliability Engineering and System Safety, vol. vol. 216, no. DOI: 10.1016/j.ress.2021.108010,, 2021., "Promoting a novel method for warranty claim prediction based on social network data," *Reliability Engineering & System Safety,,* vol. 216, no. DOI: https://doi.org/10.1016/j.ress.2021.108010, p. 108010., 2021.

[149] Samsudin, S. N., Abdullah, B., & Yusoff, N, "Customers' choice towards aftersales serviceability of counterfeit spare parts: a case study.," *Malaysian Journal of Computing (MJoC),* vol. 9(2), no. DOI: 10.24191/mjoc.v9i2.26548, pp. 1789-1798., 2024.

[150] Sarmast Hasan Kiadeh, Z., Shokouhyar, S., Omarzadeh, A., Shokoohyar, S, "Warranty operation enhancement through social media knowledge: a deep-learning methods.," *INFOR: Information Systems and Operational Research,,* vol. 62(2), no. DOI: https://doi.org/10.1080/03155986.2024.2303907, pp. 273-311., 2024.

[151] Soltanali, H., Rohani, A., Tabasizadeh, M., Abbaspour-Fard, M. H., & Parida, A, "Operational reliability evaluation-based maintenance planning for automotive production line.," *Quality Technology & Quantitative Management,,* vol. 17(2), no. DOI: https://doi.org/10.1080/16843703.2019.1567664, pp. 186-202., 2020.

[152] Babakmehr, M., Baumanns, S., Chehade, A., Hochkirchen, T., Kalantari, M., Krivtsov, V., & Schindler, D., "ata-driven framework for warranty claims forecasting with an application for automotive components," *Engineering Reports,* vol. 6(5), no. DOI: https://doi.org/10.1002/eng2.12764, p. e12764, 2024.

[153] Samsudin, S. N., Abdullah, B., & Yusoff, N, "Big data model of customer satisfaction and service experience (CSSE) for Automotive Aftersales Services," *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE),* vol. 6, no. DOI: 10.1109/ICRAIE52900.2021.9703988, pp. 1-6, December 2021.

[154] Bohani, F. A., Mohamed Rashid, F. S., Mahmud, Y.,Yahya, S. R., "Analyzing the impact of feature selection using information gain for airlines' customer

satisfaction.," *Malaysian Journal of Computing (MJoC),* vol. 9(1), no. DOI: 10.24191/mjoc.v9i1.24163, pp. 1673-1689., 2024.

[155] Mueller, C., & Mezhuyev, V., "AI models and methods in automotive manufacturing: a systematic literature review.," *Recent innovations in artificial intelligence and smart applications,,* no. DOI: https://doi.org/10.1007/978-3-031-14748-7_1, pp. 1-25., 2022.

[156] Kleyner, A., & Sanborn, K., "Modelling automotive warranty claims with build-to-sale data uncertainty.," *International Journal of Reliability and Safety,,* Vols. 2(3),, no. DOI: https://doi.org/10.1504/IJRS.2008.021063, pp. 179-189, 2008.

[157] Wu, S., & Akbarov, A., "Forecasting warranty claims for recently launched products.," *Reliability Engineering & System Safety,,* vol. 106, no. DOI: https://doi.org/10.1016/j.ress.2012.06.008, pp. 160-164, 2012.

[158] Mordor Intelligence, "Automotive Warranty Management - Market Share Analysis, Industry Trends & Statistics, Growth Forecasts (2024 - 2029)," [Online]. Available: https://www.researchandmarkets.com/reports/5996815/automotive-warranty-management-market-share. [Accessed 04 Jan 2026].

[159] K. G. Hillman, "Shareholder value, stakeholder management, and social issues: what's the bottom line?," *Strategic Management Journal,* pp. 125-139, 2022.

[160] S. Abd-Karim, "Developing the Value Management Maturity Model (VM3©)," *Journal of Design and Built Environment,* p. 1–10, 2014.

[161] P. Sharma and R. Srikonda, "APPLICATION OF VALUE ENGINEERING IN AFFORDABLE HOUSING IN INDIA," *International Journal of Engineering Technologies and Management Research,* 2021.

[162] Nurdin, A. A., Salmi, G. N., Sentosa, K., Wijayanti, A. R., & Prasetya, A. ., "Utilization of Business Intelligence in Sales Information Systems.," *Journal of Information System Exploration and Research,,* vol. 1(1), no. DOI: https://doi.org/10.52465/joiser.v1i1.101, pp. 39 - 48, 2022.

APPENDICES

APPENDIX 1

The Tables here present is the full list of features engineering from phase 3 in feature in engineering section in section 3.8.4.1 Feature Engineering

Table 3.13

Feature Metadata Summary

No.	Feature Name	Type	Description	Used in Sets	Missing Values
1	Time_Linear	Trend	Linear time index (1–48)	Basic, Enhanced, Full	0 (0%)
2	Time_Squared	Trend	Quadratic time trend	Enhanced, Full	0 (0%)
3	Year	Calendar	Calendar year (2019–2024)	All	0 (0%)
4	Month	Calendar	Calendar month (1–12)	All	0 (0%)
5	FY	Fiscal	Fiscal year identifier	All	0 (0%)
6	FM	Fiscal	Fiscal month identifier	All	0 (0%)
7	COVID_Impact	Binary	COVID period indicator	Basic, Enhanced, Full	0 (0%)
8–19	FM_1 to FM_12	Seasonal	Fiscal month dummy variables (1–12)	Basic, Enhanced, Full	0 (0%)
20	FY_2019	Annual	Fiscal year 2019 dummy	Enhanced, Full	0 (0%)
21	FY_2021	Annual	Fiscal year 2021 dummy	Enhanced, Full	0 (0%)
22	FY_2022	Annual	Fiscal year 2022 dummy	Enhanced, Full	0 (0%)
23	FY_2023	Annual	Fiscal year 2023 dummy	Enhanced, Full	0 (0%)
24	Calendar_Year	Legacy	Alternative year format	Full	0 (0%)
25	Calendar_Month	Legacy	Alternative month format	Full	0 (0%)
26	Quarter	Derived	Calendar quarter	Full	0 (0%)
27	Time_Index	Index	Primary time index	Full	0 (0%)
28	Time_Index_Complete	Index	Complete time sequence	Full	0 (0%)
29	Rolling_Mean_6M	Legacy	Pre-existing 6-month rolling mean	Full	5 (10.4%)

Table 3.14

Lag Feature Metadata Summary

No.	Feature Name	Type	Description	Lag Period	Missing Values	Treatment
30	Total_RM_lag1	Autoregressive	1-month warranty cost lag	1 month	1 (2.1%)	Preserve NaN
31	Total_RM_lag2	Autoregressive	2-month warranty cost lag	2 months	2 (4.2%)	Preserve NaN
32	Total_RM_lag3	Autoregressive	3-month warranty cost lag	3 months	3 (6.2%)	Preserve NaN
33	Total_RM_lag4	Autoregressive	4-month warranty cost lag	4 months	4 (8.3%)	Preserve NaN
34	Total_RM_lag5	Autoregressive	5-month warranty cost lag	5 months	5 (10.4%)	Preserve NaN
35	Total_RM_lag6	Autoregressive	6-month warranty cost lag	6 months	6 (12.5%)	Preserve NaN
36	Claims_Count_lag1	Volume	1-month claim count lag	1 month	1 (2.1%)	Preserve NaN
37	Claims_Count_lag2	Volume	2-month claim count lag	2 months	2 (4.2%)	Preserve NaN
38	Claims_Count_lag3	Volume	3-month claim count lag	3 months	3 (6.2%)	Preserve NaN
39	Claims_Count_lag4	Volume	4-month claim count lag	4 months	4 (8.3%)	Preserve NaN
40	Claims_Count_lag5	Volume	5-month claim count lag	5 months	5 (10.4%)	Preserve NaN
41	Claims_Count_lag6	Volume	6-month claim count lag	6 months	6 (12.5%)	Preserve NaN

Table 3.15

Rolling Window Feature Metadata Summary

No.	Feature Name	Type	Description	Window Size	Missing Values	Treatment
42	Total_RM_rolling_mean_3m	Smoothing	3-month rolling average	3 months	2 (4.2%)	Forward/ backward fill
43	Total_RM_rolling_std_3m	Variability	3-month rolling std deviation	3 months	2 (4.2%)	Forward/ backward fill
44	Total_RM_rolling_min_3m	Extremes	3-month rolling minimum	3 months	2 (4.2%)	Forward/ backward fill
45	Total_RM_rolling_max_3m	Extremes	3-month rolling maximum	3 months	2 (4.2%)	Forward/ backward fill

No.	Feature Name	Type	Description	Window Size	Missing Values	Treatment
46	Total_RM_rolling_mean_6m	Smoothing	6-month rolling average	6 months	5 (10.4%)	Forward/ backward fill
47	Total_RM_rolling_std_6m	Variability	6-month rolling std deviation	6 months	5 (10.4%)	Forward/ backward fill
48	Total_RM_rolling_min_6m	Extremes	6-month rolling minimum	6 months	5 (10.4%)	Forward/ backward fill
49	Total_RM_rolling_max_6m	Extremes	6-month rolling maximum	6 months	5 (10.4%)	Forward/ backward fill

Table 3.16

Growth Rate Feature Metadata Summary

No.	Feature Name	Type	Description	Growth Period	Missing Values	Treatment
50	Total_RM_growth_1m	Momentum	Monthly warranty cost growth rate	1 month	1 (2.1%)	Preserve NaN
51	Total_RM_growth_3m	Momentum	Quarterly warranty cost growth rate	3 months	3 (6.2%)	Preserve NaN
52	Total_RM_growth_12m	Momentum	Annual warranty cost growth rate	12 months	12 (25%)	Preserve NaN
53	Claims_Count_ growth_1m	Volume	Monthly claim count growth rate	1 month	1 (2.1%)	Preserve NaN
54	Claims_Count_ growth_12m	Volume	Annual claim count growth rate	12 months	12 (25%)	Preserve NaN

Table 3.17

Vehicle and Part Feature Metadata Summary

No.	Feature Name	Type	Description	Category	Missing Values	Treatment
55	Count_AR	Vehicle	Monthly AR claim count	Model Count	0 (0%)	Zero fill
56	Prop_AR	Vehicle	AR proportion of total claims	Model Proportion	0 (0%)	Zero fill
57	Count_BZ	Vehicle	Monthly BZ claim count	Model Count	0 (0%)	Zero fill
58	Prop_BZ	Vehicle	BZ proportion of total claims	Model Proportion	0 (0%)	Zero fill
59	Count_MV	Vehicle	Monthly MV claim count	Model Count	0 (0%)	Zero fill
60	Prop_MV	Vehicle	MV proportion of total claims	Model Proportion	0 (0%)	Zero fill
61	Count_AX	Vehicle	Monthly AX claim count	Model Count	0 (0%)	Zero fill
62	Prop_AX	Vehicle	AX proportion of total claims	Model Proportion	0 (0%)	Zero fill
63	Count_E_E	Part	Electrical/Electronic claim count	Part Count	0 (0%)	Zero fill
64	Prop_E_E	Part	E/E proportion of total claims	Part Proportion	0 (0%)	Zero fill
65	Count_A_C	Part	Air conditioning claim count	Part Count	0 (0%)	Zero fill
66	Prop_A_C	Part	A/C proportion of total claims	Part Proportion	0 (0%)	Zero fill
67–74	Additional Features	Various	Extended model and part features	Mixed	0 (0%)	Zero fill

APPENDIX 2

Appendix 2 presents selected Python code snippets from each phase of the warranty cost forecasting project, implemented in Google Colab using libraries such as pandas, scikit-learn, xgboost, and statsmodels. Figure 8.1 shows the data preprocessing steps, including loading, cleaning, and fiscal year formatting. Figure 8.2 contains model training code for Linear Regression, Random Forest, ARIMA, and XGBoost. Figure 8.3 illustrates prediction generation and error evaluation using MAE and RMSE. Figure 8.4 outlines the iterative forecasting loop across fiscal years, and Figure 8.5 presents the code for compiling model comparison results and calculating directional accuracy. These snippets support the methodology detailed in Chapter 3.

```
# Cell 1: Environment Setup & Data Loading
# Company A Alternator Warranty Claims Monthly Forecasting

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
from datetime import datetime
import re
import os

# Statistical and ML libraries
from statsmodels.tsa.stattools import adfuller
from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.tsa.arima.model import ARIMA
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from sklearn.ensemble import RandomForestRegressor
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import TimeSeriesSplit
from       sklearn.metrics       import       mean_absolute_error,
mean_squared_error, mean_absolute_percentage_error
import xgboost as xgb
```

Figure 8.1 Phase 1 Code Snippet

```
# Cell 4: Time Series Foundation Analysis
# Create monthly aggregated time series and initial analysis

print("="*60)
print("🔍 PHASE 2 - CELL 4: TIME SERIES FOUNDATION ANALYSIS")
print("="*60)

# Step 1: Correct COVID Period Identification
print("\n🔧 STEP 1: CORRECTING COVID PERIOD IDENTIFICATION")
print("-" * 50)

print("🚨 CORRECTING ALGORITHM ERROR FROM CELL 3:")
print("   Previous algorithm incorrectly flagged FY2019 as COVID-affected")
print("   FY2019 low activity (Aug 2019, Dec 2019) are pre-COVID business anomalies")
print("   REAL COVID impact: Aug-Sep 2021 in FY2021")

# Manually correct COVID period identification
covid_months_correct = [202108, 202109]  # Aug-Sep 2021 - the real COVID impact
monthly_ts['COVID_Affected_Corrected']    =    monthly_ts['Claimed Month'].isin(covid_months_correct)

print(f"\n✅ CORRECTED COVID IDENTIFICATION:")
print(f"   COVID affected months: {covid_months_correct}")
print(f"   These fall in FY2021 (as expected)")

# Update fiscal year summary with corrected COVID flagging
fy_summary_corrected = monthly_ts.groupby('Fiscal_Year').agg({
    'Claims_Count': ['count', 'sum', 'mean'],
    'Total_RM': ['sum', 'mean'],
    'COVID_Affected_Corrected': 'any'
}).round(2)

fy_summary_corrected.columns         =         ['Months_Available', 'Total_Claims', 'Avg_Monthly_Claims',
                                'Total_RM',        'Avg_Monthly_RM', 'COVID_Affected']
fy_summary_corrected = fy_summary_corrected.reset_index()

print("\nCorrected Fiscal Year Summary:")
print(fy_summary_corrected.to_string(index=False))
```

Figure 8.2 Phase 2 Code Snippet

```
# Cell 7: Feature Engineering for Multi-Model Approach
# Create comprehensive feature set for regression and time series
models

print("="*60)
print("🔍 PHASE 3 - CELL 7: FEATURE ENGINEERING FOR MULTI-MODEL
APPROACH")
print("="*60)

# Step 1: Prepare Base Dataset for Feature Engineering
print("\n📊 STEP 1: BASE DATASET PREPARATION")
print("-" * 50)

# Start with the clean monthly time series
feature_df = ts_complete.copy()
feature_df                                                    =
feature_df.sort_values('Time_Index_Complete').reset_index(drop=Tru
e)

print(f"Base dataset for feature engineering:")
print(f"   Shape: {feature_df.shape}")
print(f"      Date  range:  {feature_df['Date_Label'].iloc[0]}  to
{feature_df['Date_Label'].iloc[-1]}")
print(f"   Target variable: Total_RM")

# Rename time index for clarity
feature_df['Time_Index'] = feature_df['Time_Index_Complete']

# Check available columns and create missing ones if needed
print(f"   Available columns: {list(feature_df.columns)}")

# Create calendar year and month from Claimed Month if not available
if 'Calendar_Year' not in feature_df.columns:
    feature_df['Calendar_Year']       =       feature_df['Claimed
Month'].apply(lambda x: int(str(x)[:4]))
if 'Calendar_Month' not in feature_df.columns:
    feature_df['Calendar_Month']       =       feature_df['Claimed
Month'].apply(lambda x: int(str(x)[4:6]))
```

Figure 8.3 Phase 3 Code Snippet

```
# Cell 12: Iterative Validation - FY19→FY21→FY22→FY23 Strategy
# Implement progressive temporal validation excluding COVID-affected FY20

print("="*70)
print("🔍 PHASE 4 - CELL 12: ITERATIVE VALIDATION STRATEGY")
print("="*70)

# Step 1: Prepare Iterative Validation Framework
print("\n🔄 STEP 1: ITERATIVE VALIDATION FRAMEWORK SETUP")
print("-" * 60)

# Use the complete feature dataset from previous cells
iteration_data = modeling_datasets['enhanced_regression'].copy()

# Create fiscal year mapping for easy filtering
iteration_data['FY'] = iteration_data['Claimed Month'].apply(
    lambda x: int(str(x)[:4]) if int(str(x)[4:6]) >= 4 else int(str(x)[:4]) - 1
)

print(f"Iterative validation dataset:")
print(f"   Total observations: {len(iteration_data)}")
print(f"    Date range: {iteration_data['Date_Label'].iloc[0]} to {iteration_data['Date_Label'].iloc[-1]}")

# Verify fiscal year distribution
fy_counts = iteration_data['FY'].value_counts().sort_index()
print(f"\nFiscal year distribution:")
for fy, count in fy_counts.items():
    print(f"   FY{fy}: {count} months")

# Exclude FY2020 (incomplete data) as planned
iteration_data_clean = iteration_data[iteration_data['FY'] != 2020].copy()
print(f"\nAfter excluding FY2020:")
print(f"   Clean observations: {len(iteration_data_clean)}")
```

Figure 8.4 Phase 4 Code Snippet

```
# Cell 14: Multi-Model Production Forecasting Comparison
# Generate FY24 forecasts with all validated models and
comprehensive comparison

print("="*70)
print("🔍 PHASE 4 - CELL 14: MULTI-MODEL PRODUCTION FORECASTING
COMPARISON")
print("="*70)

# Step 1: Prepare Unified Training Dataset for All Models
print("\n🔄 STEP 1: UNIFIED TRAINING DATASET PREPARATION")
print("-" * 60)

# Use the same base data as Cell 13 for consistency
unified_data = modeling_datasets['enhanced_regression'].copy()
unified_data['FY'] = unified_data['Claimed Month'].apply(
    lambda x: int(str(x)[:4]) if int(str(x)[4:6]) >= 4 else
int(str(x)[:4]) - 1
)

# Exclude incomplete FY2020
train_unified = unified_data[unified_data['FY'] != 2020].copy()

print(f"Unified training dataset:")
print(f"   Total observations: {len(train_unified)}")
print(f"   Fiscal years: {sorted(train_unified['FY'].unique())}")
print(f"    Date range: {train_unified['Date_Label'].iloc[0]} to
{train_unified['Date_Label'].iloc[-1]}")

# Prepare target variable
y_unified = train_unified['Total_RM'].copy()
```

Figure 8.5 Phase 5 Code Snippet

AUTHOR'S PROFILE

Ts. Nor Hissham bin Abdul Hamid received his Master of Engineering Management from Universiti Teknologi MARA (UiTM), Malaysia, in 2023. He is currently pursuing a PhD in Mechanical Engineering at the Faculty of Mechanical Engineering, UiTM. His research focuses on product design quality improvement for warranty cost reduction in automotive aftermarket sales, with emphasis on warranty analytics, predictive modelling, and AI-integrated business intelligence systems. He has over 20 years of experience in the automotive industry and currently serving as Assistant General Manager for OEM Sales in a multinational automotive company, where he is responsible for warranty analysis, root cause investigation, and quality improvement initiatives. He has been actively involved with the Malaysia Board of Technologists (MBOT) since 2023 as a Panel Assessor and Accreditation Panel member under the Automotive Technology and Technical Accreditation Scheme (TTAS).

LIST OF PUBLICATION:

N. Hissham A. H., M. Fauzi I., M. Ilham K., A. Malek A. W., and N. Fazli A. M., "Improvement product design quality to reduce warranty cost in aftermarket sales for OEM's customer," *Journal of Mechanical Engineering*, vol. 22, no. 3, pp. 1–15, Sep. 2025, doi: 10.24191/jmeche.v22i3.5494.

N. Hissham A. H., M. Fauzi I., A. Malek A. W., and N. Fazli A. M., "Development of warranty management visualisation for automotive aftermarket with integration of AI," *Journal of Mechanical Engineering*, vol. 23, no. 1, pp. 19–38, Jan. 2026, doi: 10.24191/jmeche.v23i1.5527.

Nor Hissham Abdul Hamid, Nor Fazli Adull Manan, Mohd Fauzi Ismail, Muhammad Ilham Khalit, and Abdul Malek Abdul Wahab, "Development of a Visualization Framework for Automotive Warranty Data Management", *JCRINN*, vol. 11, no. 1, pp. 157–172, Mar. 2026. doi:10.24191/jcrinn.v11i1.624.

M. Fauzi I., N. Hissham A. H., M. H. Mahamad Basri, and N. Fazli A. M., "Improving automotive accessory development: A warranty-based analysis of material defects and process enhancement," *Journal of Applied Engineering Design and Simulation*, vol. 5, no. 2, pp. 1–19, 2025, doi: 10.24191/jaeds.v5i2.112.

LIST OF CONFERENCES:

N. Hissham A. H., M. Hakim R., M. Fauzi I., F. Jika, and N. Fazli A. M., "Development of a visualization framework for automotive warranty data management," in *Proceedings of the 12th International Conference on Science & Social Research (CSSR 2025)*, Nov. 2025, Awarded Best Presenter.

LIST OF INTELLECTUAL PROPERTY (IP):

N. Hissham A. H., M. Hakim R., A. Malek A. W., and N. Fazli A. M., "Component repair strategy for warranty cost reduction in automotive alternator assemblies," Copyright Act 1987, Literary Work, Registration No. CRLY2025P05673, Jun. 26, 2025.