

UNIVERSITI TEKNOLOGI MARA

**SHORT-TERM FORECAST OF NEW
CONFIRMED CASES COVID-19 IN
JAPAN USING TIME SERIES
ANALYSIS MODELS**

NOR ALYANI BINTI AZIZ

**Report submitted in fulfilment of requirement for
Bachelor of Science (Hons.) Management Mathematics
Faculty of Computer and Mathematical Science**

January 2021

STUDENT'S DECLARATION

I certify that this report and the research to which it refers are the product of my own work and that any ideas or quotation from the work of other people, published or otherwise are fully acknowledged in accordance with the standard referring practices of the discipline.



.....

NOR ALYANI BINTI AZIZ

2019361611

JANUARY 27, 2021

ABSTRACT

The new strain of coronavirus (COVID-19) was found to have started in Wuhan, China in late December 2019. The virus has spread to countries all over the world including Japan. The World Health Organization (WHO) declared COVID-19 as a pandemic on 11 March 2020 due to the increasing number of new confirmed cases and deaths each day. The COVID-19 outbreak has impacted the nation of Japan adversely and the number of confirmed cases in Japan continues to increase day by day. On 7 April 2020, Japan declared a state of emergency to prevent the pandemic from worsening. This study is conducted to forecast new daily confirmed cases of COVID-19 in Japan over a short-term period. Four univariate time series models, namely, the Naïve Model, Mean Model, Autoregressive Integrated Moving Average (ARIMA) Model and Exponential State Space Model were applied. This study analyzes daily data from 22 January to 10 April 2020 collected from the Our World in Data website. The prediction involves five phases of data analysis as well as five different partitions of estimation and evaluation parts in every model to ensure the accuracy of forecast values. R and RStudio software were used in this study to analyze the data. The results reveal that Naïve model with 99 percent of estimation part and 1 percent evaluation part produces the lowest value of error measures for Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scaled Error (MASE).

Keywords: COVID-19, Naïve, Mean, ARIMA, Exponential State Space Model, R programming

TABLE OF CONTENTS

CONTENTS	PAGE
SUPERVISOR’S APPROVAL	ii
STUDENT’S DECLARATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
LIST OF ABBREVIATIONS	x
CHAPTER ONE: INTRODUCTION	
1.1 Background of the Study	1
1.2 Problem Statement	2
1.3 Objective of the Study	4
1.4 Scope of the Study	4
1.5 Significance of the Study	5
CHAPTER TWO: LITERATURE REVIEW	
2.1 Coronavirus Disease (COVID-19)	7
2.2 Time Series Analysis	9
2.3 Summary	14
CHAPTER THREE: RESEARCH METHODOLOGY	
3.1 Method of Data Collection	15

3.2	Method of Data Analysis	15
3.3	Univariate Time Series Modelling	17
3.3.1	Naïve Model	
3.3.2	Mean Model	
3.3.3	Autoregressive Integrated Moving Average (ARIMA)	
3.3.4	Exponential State Space Model	
3.3.5	Model Selection Criteria	
3.4	Statistical Software R Programming	23

CHAPTER FOUR: RESULTS AND DISCUSSIONS

4.1	Data Cleaning	26
4.2	Descriptive Analysis	27
4.3	Univariate Time Series Modelling	32
4.4	Evaluation Procedures	35
4.5	Forecast Future Values	37
4.6	Discussion	38

CHAPTER FIVE: CONCLUSIONS AND RECOMMENDATIONS

5.1	Conclusions	40
5.2	Recommendations	41

REFERENCES	42
-------------------	----

APPENDICES

APPENDIX A: Dataset	47
APPENDIX B: R Syntax for Check on Stationary Condition for ARIMA Model	49
APPENDIX C: Results of Error Measures for All Models	50