

UNIVERSITI TEKNOLOGI MARA

**SHORT-TERM FORECAST OF NEW
CONFIRM CASE OF COVID-19 IN
HUBEI, CHINA USING NAÏVE
METHOD, MEAN MODEL,
AUTOREGRESSIVE INTEGRATED
MOVING AVERAGE (ARIMA) AND
STATE SPACE MODEL**

RAJA AZIZUL FITRI BIN RAJA AZRAN

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

July 2020

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.



.....
RAJA AZIZUL FITRI BIN RAJA AZRAN
2017994703

AUGUST 5, 2020

ABSTRACT

The new virus and disease have been found (COVID-19), with the outbreak started in December 2019 in Wuhan, unknown China. Coronavirus disease 2019 is a global pandemic that affects many countries. The number of confirmed cases worldwide on 4 May 2020 is 3,435,894 cases, with 239,604 deaths. On 25 March 2020, China recorded the highest number of confirmed cases with 81,848 cases and 3,827 deaths. The new confirmed cases keep on increasing day by day. Therefore, this study is conducted to forecast the new confirmed cases of COVID-19 in Hubei, China for a short-term period by using the Naïve method, Mean model, Autoregressive Integrated Moving Average (ARIMA), and State space model. There are 9 different training sets and test sets for every method in this study. All the methods are also divided into 4 datasets in which each dataset will predict a 3-step ahead forecast by using the best model that produces the least error measure. The result shows that the Naïve method is the best model for all 4 datasets since it produces the lowest error measure. However, the prediction of new confirmed cases by the Naïve method is not accurate to the actual new confirmed cases from the Kaggle website.

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

TABLE OF CONTENTS

CONTENTS	PAGE
SUPERVISOR’S APPROVAL	ii
STUDENT’S DECLARATION	iii
ACKNOWLEDGEMENTS	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	2
1.4 Scope of the Study	3
1.5 Significance of the Study	3
CHAPTER TWO: LITERATURE REVIEW	
2.1 COVID-19	4
2.2 Time Series Analysis	5
2.3 Summary	6

CHAPTER THREE: RESEARCH METHODOLOGY

3.1	Method of Data Collection	7
3.2	Naïve method	7
3.3	Mean Model	8
3.4	Autoregressive Integrated Moving Average (ARIMA)	9
3.5	State Space Model	10

CHAPTER FOUR: RESULTS AND DISCUSSIONS

4.1	Introduction	12
4.2	Data Analysis	15
4.2.1	Result of Dataset 1	15
4.2.2	Result of Dataset 2	20
4.2.3	Result of Dataset 3	21
4.2.4	Result of Dataset 4	22

CHAPTER FIVE: CONCLUSION AND RECOMMENDATION

5.1	Conclusion	24
5.2	Recommendation	25

REFERENCES	26
-------------------	----

APPENDICES	28
-------------------	----