



**Forecasting the Group Health Insurance Claim Size for
Outpatient Benefit Using ARIMA Method**

UNDERGRADUATE THESIS

Submitted as one of the requirements to obtain

Sarjana Aktuaria

By

Salwa Fayza Alkatraz

021202100031

FACULTY OF BUSINESS

ACTUARIAL SCIENCE STUDY PROGRAM

CIKARANG

MARCH 2025

PLAGIARISM REPORT

Thesis Salwa Fayza Draft.docx

ORIGINALITY REPORT

15% SIMILARITY INDEX	16% INTERNET SOURCES	12% PUBLICATIONS	10% STUDENT PAPERS
--------------------------------	--------------------------------	----------------------------	------------------------------

PRIMARY SOURCES

1	www.ijisae.org Internet Source	3%
2	graphsearch.epfl.ch Internet Source	2%
3	ia801208.us.archive.org Internet Source	1%
4	Aduroja, Samuel. "Intelligent Predictive Maintenance for Fire Detection System with Improved Holt-Winters Method", Universidade do Porto (Portugal), 2024 Publication	1%
5	Submitted to Aston University Student Paper	1%
6	c3.ai Internet Source	1%
7	e-journal.president.ac.id Internet Source	1%
8	deepai.org Internet Source	1%
9	Rini Cahyandari, Sukono, Riaman, Nurnadiah Zamri. "Model of Discrete-Time Surplus Process for Scheme of Productive Waqf Integration with Sustainable Fishermen's Welfare Benefits Based on Several Threshold Levels: Systematic Literature Review", Sustainability, 2025 Publication	1%

GPTZERO REPORT

AI Report

We predict this text is

Human Generated

AI Probability

4%

This number is the probability that the document is AI generated, not a percentage of AI text in the document.

Plagiarism

4%

Percentage of this document that matches internet sources, including websites, documents, books, etc.

CHAPTER I - 4/22/2025

Salwa Fayza Alkatraz

CHAPTER I

INTRODUCTION

1.1 Background of The Problem

Insurance is a contract between an insurance company (insurer) and a policyholder (insured), where the insured pays a premium to protect the risk of damage or legal liability to a third party based on the death of the insured or on the death or survival of the insured, if the benefits have been determined and/or if the funds are made by the insured based on the results of management (Rejda & McNamara, 2021).

There are various types of insurance, each serving different purposes. Life insurance that provides financial benefits in the event of death, sudden illness, or permanent disability due to an accident or illness. Vehicle insurance that protects valuable assets from accidents or other damage and covers potential losses. Education insurance that designed to offer financial security and future planning for children's education. Bancassurance is a collaboration between banks and insurance companies to meet customer protection needs. Marine insurance that covers goods during transportation by sea, land, or air, whether for a single shipment, declaration basis, or under a binding contract. Lastly, health insurance provides coverage for medical expenses, including drug costs, surgeries, and other healthcare needs.

In insurance, there are individual insurance programs and corporate insurance (group). One product that has these two types of insurance is health insurance. Unlike individual health insurance that only covers 1 person, group health insurance covers many people such as the main insured, spouse and child. For groups of individuals, this product is more beneficial than individual health insurance products because risk selection is easier for each participant and the premium rates offered are lower. Unlike other types of insurance, group health insurance has a short coverage period, up to one year and can be renewed.

ABSTRACT

Group health insurance is a health insurance product that provides health protection to a group of individuals that covers their medical expenses and related cost if they experience illness or accident that offers benefits such as lower premiums and easier risk selection. As an insurer, an insurance company needs to prepare for the size of claims that will occur in the next period in order to minimize the misfortunes caused by uncertain or inflated claims. This study predicts the number of group health insurance claims using the Autoregressive Integrated Moving Average (ARIMA) method. The ARIMA method uses historical data to forecast future values by identifying patterns and trends. This research is based on the need for PT. ABC to forecast future claims based on claims data from September 2022 to August 2023, which shows random fluctuations. With the ARIMA (3,1,0) model, this study successfully predicted the number of group health insurance claims for outpatient care. The model was selected based on Akaike Information Criterion (AIC) and validated using Shapiro-Wilk and Ljung-Box tests, which confirmed the normality and absence of autocorrelation in the residuals. The ARIMA (3,1,0) model was selected because it was the only model that met the statistical test requirements. This study forecasts claim amounts for the next five weeks, from August 26th, 2023, to September 23rd, 2023. The results show an MAE and RMSE of 970,418,438, an MSE of 1.36×10^{18} , and a MAPE of 4.05%, indicating a high level of accuracy.

Keywords: *Time Series Analysis, Insurance Claims, Autoregressive Integrated Moving Average (ARIMA) Method, Group Health Insurance, Forecasting*

ABSTRAK

Asuransi kesehatan kumpulan adalah produk asuransi kesehatan yang memberikan perlindungan kesehatan kepada sekelompok individu yang menanggung biaya pengobatan dan biaya terkait jika mereka mengalami sakit atau kecelakaan. Perusahaan asuransi harus memahami dan merencanakan ukuran klaim yang akan terjadi di masa depan. Dengan cara ini, perusahaan asuransi dapat meminimalisir kerugian yang disebabkan oleh klaim yang klaim yang tidak pasti atau meningkat. Penelitian ini memprediksi jumlah klaim asuransi kesehatan kelompok menggunakan metode Autoregressive Integrated Moving Average (ARIMA) dengan data historis untuk meramalkan nilai masa depan dengan mengidentifikasi pola dan tren. Penelitian ini didasarkan pada kebutuhan PT. ABC untuk memperkirakan klaim masa depan berdasarkan data klaim dari September 2022 hingga Agustus 2023, yang menunjukkan fluktuasi acak. Dengan model ARIMA (3,1,0), penelitian ini berhasil memprediksi jumlah klaim asuransi kesehatan kelompok untuk rawat jalan. Model ini dipilih berdasarkan Akaike Information Criterion (AIC) dan divalidasi menggunakan uji Shapiro-Wilk dan Ljung-Box, yang mengonfirmasi kenormalan dan tidak adanya autokorelasi pada residual. Model ARIMA (3,1,0) dipilih karena merupakan satu-satunya model yang memenuhi persyaratan uji statistik. Penelitian ini memprakirakan besaran klaim untuk 5 minggu ke depan, mulai dari 26 Agustus 2023 hingga 23 September 2023. Hasil peramalan menunjukkan MAE dan RMSE sebesar 970.418.438, MSE sebesar $1,36 \times 10^{18}$, dan MAPE 4,05%, yang mengindikasikan tingkat akurasi yang tinggi.

Kata Kunci: *Peramalan, Klaim Asuransi Kesehatan, Metode Autoregressive Integrated Moving Average (ARIMA), Asuransi Kesehatan Kumpulan, analisis time series*

TABLE OF CONTENTS

PANEL EXAMINER APPROVAL	ii
STATEMENT OF ORIGINALITY	iii
SCIENTIFIC PUBLICATION APPROVAL FOR ACADEMIC INTEREST	iv
ADVISOR APPROVAL FOR JOURNAL/INSTITUTION'S REPOSITORY	v
PLAGIARISM REPORT	vi
GPTZERO REPORT	vii
ABSTRACT	viii
ABSTRAK	ix
ACKNOWLEDGEMENT	x
TABLE OF CONTENTS.....	xii
LIST OF FIGURE	xv
LIST OF TABLE	xvi
CHAPTER I.....	1
INTRODUCTION.....	1
1.1 Background of The Problem	1
1.2 Research Question	5
1.3 Research Objectives	5
1.4 Research Benefit	5
1.5 Scope and Limitations.....	7
CHAPTER 2	8
LITERATURE REVIEW.....	8
2.1 Insurance.....	8
2.2 Claim of Health Insurance.....	9
2.3 Sample Space	9
2.4 Event	9
2.5 Probability.....	10
2.6 Random Variable	10
2.6.1 Cumulative Distribution Function (CDF)	11
2.6.2 Probability Density Function (PDF)	11

2.6.3 Expected Value, Variance, Covariance, and Correlation	12
2.7 Time Series	13
2.7.1 Autoregressive Process (AR).....	14
2.7.2 Moving Average (MA).....	14
2.7.3 Autoregressive Moving Average (ARMA) Model.....	15
2.7.4 Autoregressive Integrated Moving Average (ARIMA) Model	15
2.8 Hypothesis Testing	16
2.9 Stationary	17
2.10 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).....	19
2.11 Normality Test	20
2.12 The Ljung-Box Test.....	20
2.13 Error Analysis.....	21
2.13.1 Root Mean Square Error (RMSE)	21
2.13.2 Mean Square Error (MSE)	22
2.13.3 Mean Absolute Error (MAE).....	22
2.13.4 Mean Absolute Percent Error (MAPE)	23
2.14 Previous Research	24
2.15 Research Gap	26
CHAPTER 3	27
METHODOLOGY	27
3.1 Researched Design.....	27
3.2 Data Analysis Design	27
3.3 Research Flowchart.....	28
CHAPTER IV	29
RESULT AND DISCUSSION.....	29
4.1 Determining Time Series Model Based on Claim Amount.....	29
4.1.1 Stationarity	29
4.1.2 Identify ARIMA Model	30
4.1.3 Parameter Testing.....	32
4.1.4 Residual Testing.....	33

4.2 Forecasting	34
4.3 Error Testing	36
CHAPTER V	37
CONCLUSION AND RECOMMENDATION	37
5.1 Conclusion	37
5.2 Recommendation	39
RERERENCES	40

LIST OF FIGURE

Figure 1.1 Size of Claim from September 2022 to August 2023 in Rupiah	3
Figure 4.1 PACF and ACF Plot for Outpatient Benefit.....	30
Figure 4.2 Model 2 Q-Q Normal Plot	34
Figure 4.3 Forecasting Plot for 5 Weeks	35

LIST OF TABLE

Table 4.1 ARIMA(3,1,1) Model.....	31
Table 4.2 Parameter Estimation Result of Claim Amount Data.....	32
Table 4.3 Residual Analysis Result.....	33
Table 4.4 Actual Data for the Next 5 Weeks	34
Table 4.5 Forecasting Result of the Data	35
Table 4.6 Error Value Estimation Parameters Result	36