ISSN 2278-3091
Volume 8, No.3, May - June 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse33832019.pdf
https://doi.org/10.30534/ijatcse/2019/33832019

Scholarship Grants Prediction using Autoregressive Integrated Moving Average (ARIMA) Algorithm

Elvis P. Patulin
College of Arts and Sciences, Surigao State College of Technology, Surigao City, Philippines,
elvispatulin@yahoo.com

ABSTRACT

Every university has its methods to realize scholarship assessment. Schools establish scholarships as a motivation for students that have outstanding and exceptional achievements. It is a form of incentive and benefit to inspire and embolden students and improve schools' learning principles. From that point of view, an optimal way is considered to assess and distribute scholarship grants. In this paper, the use of Autoregressive Integrated Moving Average (ARIMA) was observed in predicting future scholarship grants that can be availed in Surigao State College of Technology, Surigao City, Philippines. The output of this study can be used as input for management decision as to budget allocation by the sponsoring agent of the scholarship program. It is also hoped that this study could contribute on the literature of time series forecasting and educational data mining in general.

Key words: ARIMA, EDM, Forecasting, Prediction

1. INTRODUCTION

Almost all major colleges and state universities make an assessment on students' comprehensive quality and set different rewards and regulations for the various level in order to stimulate students' interest to study and participate in extracurricular activities. The main reward system that is used is the providing of financial incentives such as scholarship grants [1].

Data mining also termed as Knowledge Discovery in Databases (KDD), is a medium of discovering novel and potentially valuable information from large amounts of data [2]-[4]. The field of Educational Data Mining is fresh, new, and developing in the field of education sector which can also be applied in areas such as government, accounts, sports, transportation and a lot more [5]. This newly emerging field concerns with developing methods that discover knowledge from data originating from educational environments. Data mining techniques such as decision trees, Naïve Bayes, K-Nearest neighbor, neural networks, and many others are instrumental in extracting data from the datasets [6].

This study forecast the indexed data of the different scholarship programs offered in Surigao State College of Technology, the only State College in the province of Surigao del Norte, Philippines as to the number of future scholarship grantees using ARIMA algorithm. From the obtained results, this will also provide a predicted increase and decrease of the grantees in the next five years. This is to monitor and to identify which scholarship grant will exhibit an increase in the number of grantees within targeted time as a preparation for the budget allocation by the sponsoring agents.

2. LITERATURE REVIEW

Recently, a lot of data are obtained in industry, agriculture, medical, education areas and so on. KDD technology develops gradually with databases. Data mining is the key step of KDD. Data mining extracts patterns from large data such that they are certain, previously unknown, and potentially useful for the specific application [5].

The literature on representations and distance measures for time-series, clustering [7], [8], and classification is extensive [9]-[11]. Time series analysis method is a kind of data mining method, which is a sequence of data points, typically consisting of successive measurements made over a time interval. It is a method for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values [12].

Prediction can be made utilizing autoregressive integrated moving average (ARIMA) algorithm that used historical data in predicting cases such as in education, society, climate, health, etc. In other countries, [13] used ARIMA in forecasting incidence of hemorrhagic fever with renal syndrome in China. Meanwhile, [14] also used ARIMA model in forecasting dengue hemorrhagic fever cases in Southern Thailand. Further, the study of [15] indicated the potential and effectiveness of using ARIMA modeling in the prediction of travel time to the urban roadway.

3. METHODOLOGY

3.1 Datasets

The data that were used are the following but are not limited as to wit: different scholarship grants that can be availed in SSCT, the gender of the grantees, and the number of grantees per scholarship grants per year from the year 2014-2017. The dataset was retrieved from the Student Financial Assistance Unit (SFAU) of Surigao State College of Technology, Surigao City, Philippines.

3.2 Scholarship Forecasting

The R Language which runs in RStudio software was utilized for the trend analysis of indexed data of the different scholarship grants offered in Surigao State College of

Technology. Initially, the time series graph obtained in every cluster showing the behavior of indexed scholarship data. The study utilized the ARIMA(1,0,0) model to forecast the possible increase of scholarship grantees on the following scholarship programs: Academic scholarship, athletic scholarship, barangay scholarship, choir, city scholarship, CSSGP, dance troupe, ESGP-PA, LGU Basilisa, LGU Claver, PGMC, provincial eskolaran, StuFAPS, TMC, TD01, TD02, and TD022 for the next five years. The main concept of the algorithm is that a corresponding mathematical model describes a group of orderly time series data formed over time, and then future increase is forecasted. Figure 1 shows the algorithm flowchart of ARIMA.

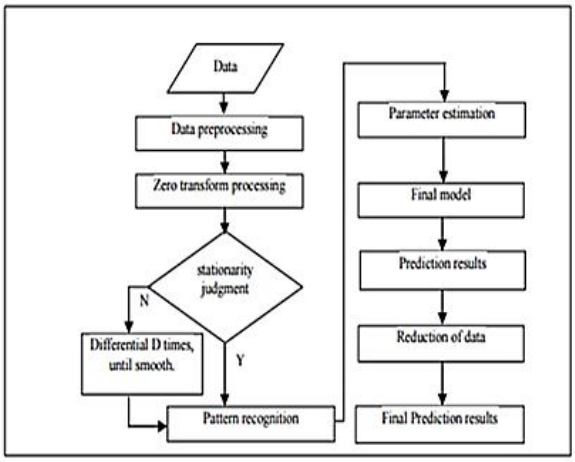

Figure 1: ARIMA algorithm flowchart

4. RESULTS AND DISCUSSION

Figures 2-18 showed the graph of the predicted grantees for each scholarship grants from 2017 to 2022 having a high-95 % prediction.

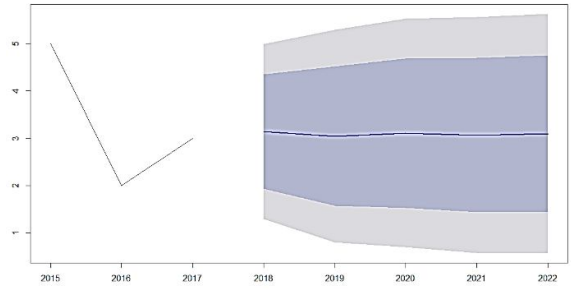

Figure 2: Forecasted grantee for Academic Scholarship grant from 2017-2022.

A steady pattern is shown for the forecasted grantees for the Academic Scholarship grant from the year 2017 to 2022 having been forecasted an average of 3 grantees per year with the highest number of 5 on the year 2022 is evident in the Figure 2.

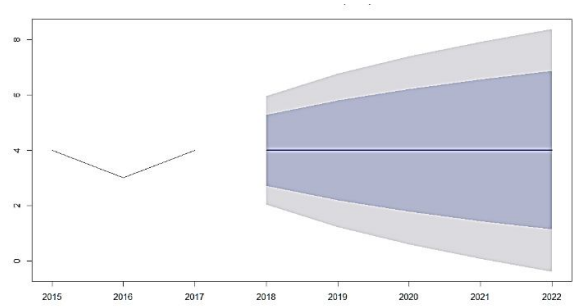

Figure 3: Forecasted grantees for Athletic Scholarship grant from 2017-2022

The predicted grantees for the Athletic Scholarship grant showed data stability from 2017 to 2022. The prediction signifies that the forecasted data reach the same number of the grantee in SSCT during the year 2015 to 2017, as shown in Figure 3.

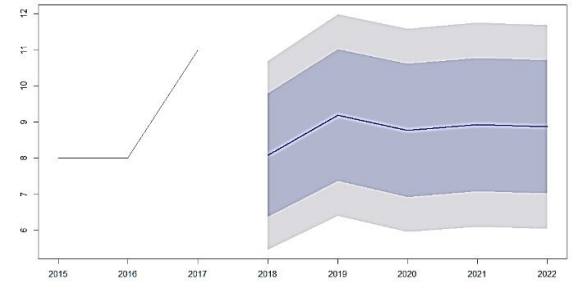

Figure 4: Forecasted grantee for Barangay Scholarship grant from 2017-2022

The forecasted grantees for the Barangay Scholarship grant from 2017 to 2022 in SSCT showed a decreasing pattern from 2017 to 2018. Moreover, a prediction of nine with high-11 grantees in the year 2019 is depicted and data stability from the year 2020 to 2022 as evident in Figure 4.

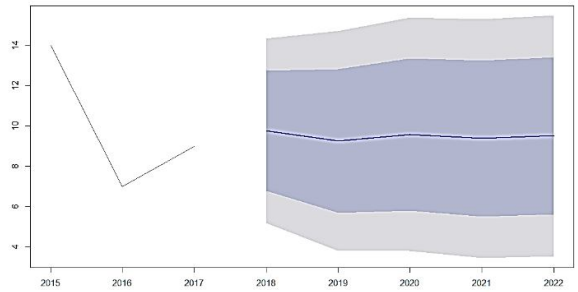

Figure 5: Forecasted grantee for Choir Scholarship grant from 2017-2022

The forecasted grantees for the Choir Scholarship grant from 2017 to 2022 in SSCT showed a slightly increased pattern from 2017 to 2018 as shown in Figure 5 graph of forecast.

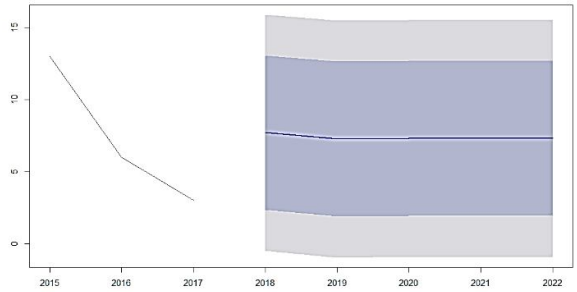

Figure 6: Forecasted grantee for CSSGP Scholarship grant from 2017-2022

The forecasted grantees for the CSSGP Scholarship grant from 2017 to 2022 showed an increased forecast pattern from 2017 to 2018 and a steady trend from 2018 to 2022 based on Figure 6.

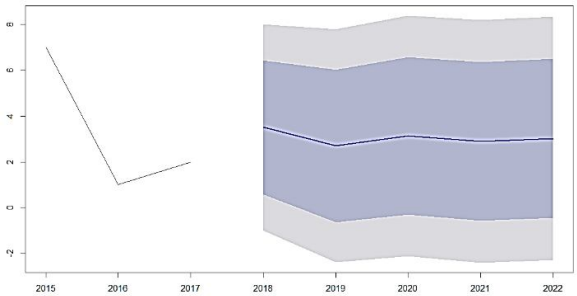

Figure 7: Forecasted grantee for Dance Troupe Scholarship grant from 2017-2022

The forecasted grantees for Dance Troupe Scholarship grant from 2017 to 2022 showed an increased pattern from the year 2017 to 2018 and a slightly reduced pattern from 2018 to 2019 as evident in Figure 7.

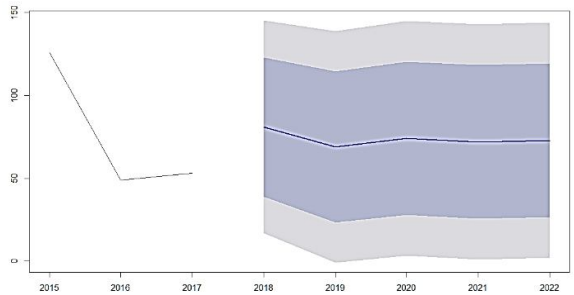

Figure 8: Forecasted grantee for ESGP-PA Scholarship grant from 2017-2022

The predicted grantees for the ESGP-PA showed an increased pattern from 2017 to 2018, and there is a slightly decreased pattern from the year 2018 to 2019. There is a stability of grantees in ESGP-PA from the year 2020 to 2022, as evident in Figure 8.

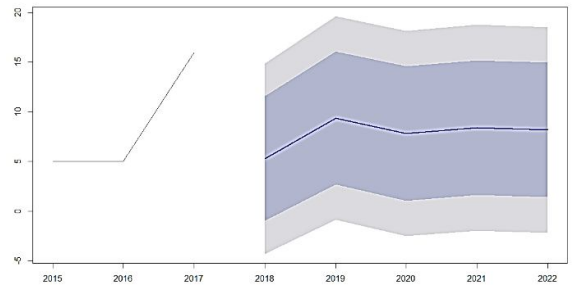

Figure 9: Forecasted grantee for LGU Basilisa Scholarship grant from 2017-2022

Forecasted grantees for the LGU Basilisa Scholarship grant from 2017 to 2022 showed a rapidly decreased pattern from 2017 to 2018. Moreover, there is an increased design from 2018 to 2019, as evident in the graph. Further, there is a slightly decreased pattern from the year 2019 to 2020, as shown in Figure 9.

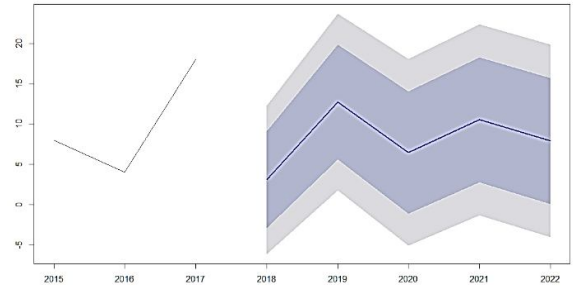

Figure 10: Forecasted grantee for LGU Claver Scholarship grant from 2017 to 2022

Forecasted grantees for the LGU Claver Scholarship grant from 2017 to 2022 showed a rapidly decreased pattern from 2017 to 2018. Moreover, there is an increased design of the grant from 2018 to 2019, as evident in the graph. From the year 2019 to 2020, it is evident that there is a slightly decreased pattern of the grantees. Further, there is a slightly increased pattern from the year 2020 to 2021 and a decrease after, as shown in Figure 10.

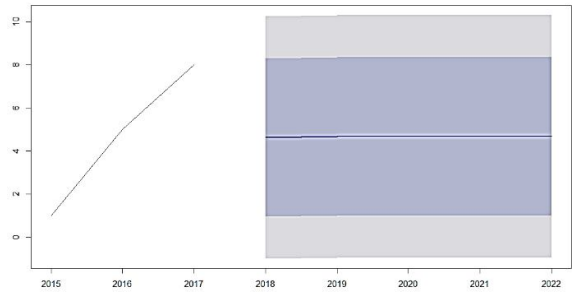

Figure 11: Forecasted grantee for PGMC Scholarship grant from 2017 to 2022

Forecasted grantees for PGMC showed a stability pattern from 2018 to 2022 but a rapid decrease from the year before, as shown in Figure 11.

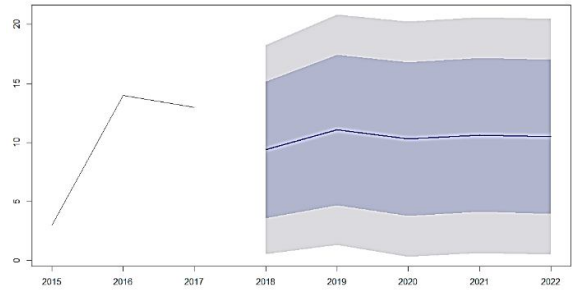

Figure 12: Forecasted grantee for StuFAPS Scholarship grant from 2017 to 2022

Forecasted grantees for the StuFAPS Scholarship grant from 2017 to 2022 showed a rapidly decreased pattern from 2017 to 2018. Moreover, there is an increased design from 2018 to 2019, as evident in the graph. Further, there is a slightly decreased pattern from the year 2019 to 2020, and a stable graph in the following years is shown in Figure 12.

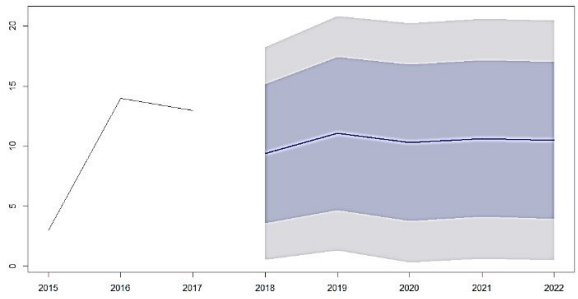

Figure 13: Forecasted grantee for TMC Scholarship grant from 2017 to 2022

Forecasted grantees for the TMC Scholarship grant from 2017 to 2022 showed a decreased pattern from 2017 to 2018. There is an increase of grantees from 2018 to 2019, as evident in Figure 13. Further, there is a slightly decreased pattern from the year 2019 to 2020 and a stable graph right after 2020.

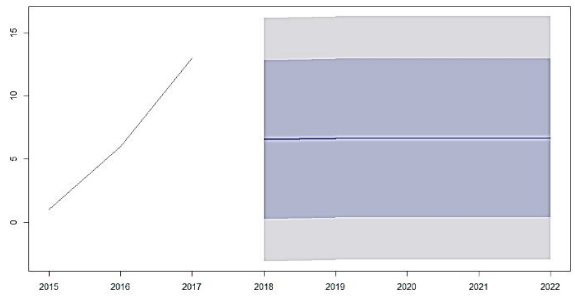

Figure 14: Forecasted grantee for TD-01 Scholarship grant from 2017 to 2022

Forecasted grantees for Tulong Dunong-01 showed a decrease of the grantees from 2017 to 2018 and a stability pattern from 2018 to 2022, as shown in Figure 14.

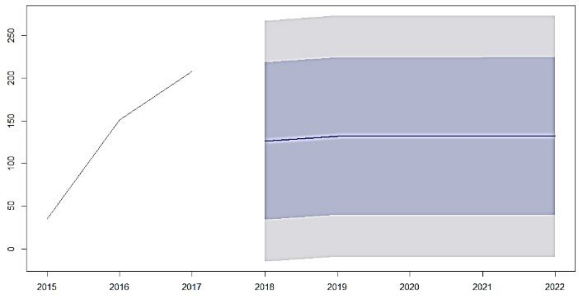

Figure 15: Forecasted grantee for City Scholarship grant from 2017 to 2022

Forecasted grantees for City Scholarship grant showed a rapid decrease from the year 2017 to 2018 and a stability pattern from 2018 to 2022 based on Figure 15.

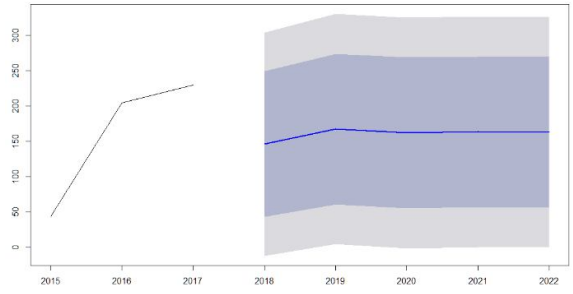

Figure 16: Forecasted grantee for Provincial Eskolaran grant from 2017 to 2022

The forecasted grantees for Provincial Eskolaran from 2017 to 2022 showed a decreasing pattern from the year 2017 to 2018. Moreover, there is a slightly increased pattern from 2018 to 2019 and a stable pattern onwards based in Figure 16.

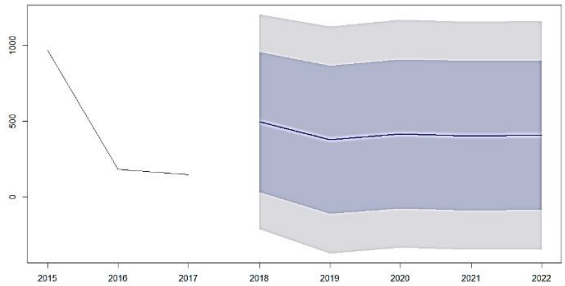

Figure 17: Forecasted grantee for TD-022 Scholarship grant from 2017 to 2022

The forecasted grantees for Tulong Dunong-02 under Cong. Bag-ao, Cong. Barbers and Romarate Scholarship grant from 2017 to 2022 showed an increased pattern from the year 2017 to 2018. Moreover, there is a slightly reduced pattern from 2018 to 2019 and a stable pattern from 2020 to 2022, as shown in Figure 17.

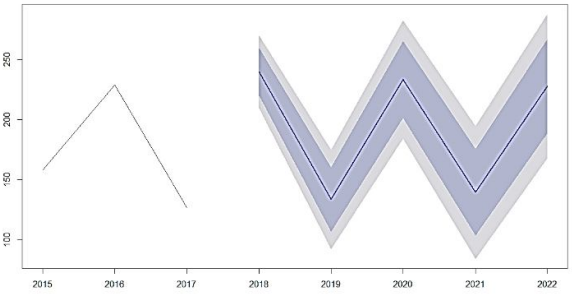

Figure 18: Forecasted grantee for TD-0222 Scholarship grant from 2017 to 2022

Forecasted grantees for the Tulong Dunong-02 under Congressman Matugas Scholarship grant from 2017 to 2022 showed a rapidly increased pattern from 2017 to 2018. Moreover, there is a decreased design of the grant from 2018 to 2019, as evident in the graph. From the year 2019 to 2020, it is evident that there is a rapidly increased pattern of the grantees. Further, there is a slightly decreased pattern from the year 2020 to 2021 and an increase after, as shown in Figure 18.

5. CONCLUSION

Predicting grantees for each scholarship programs can be beneficial to the sponsoring agents since it will give them insight as to the number of their future grantees. This allows them to better prepare for budget allocation. In the study, the use of ARIMA algorithm with ARIMA(1,0,0) model helped in predicting the possible grantees in each scholarship grants. The simulation result showed the effectiveness of the ARIMA algorithm in time series prediction. It is recommended that the use of ARIMA algorithm on other time series datasets will be observed for future researches.

REFERENCES

[1] W. Wei, J. Han, J. Kong, and H. Xia, "Prediction of the Scholarship Using Comprehensive Development," *Proc. - 4th Int. Conf. Enterp. Syst. Adv. Enterp. Syst. ES 2016*, pp. 183–188, 2017. https://doi.org/10.1109/ES.2016.30

[2] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012.

[3] A. J. P. Delima, "An Experimental Comparison of Hybrid Modified Genetic Algorithm-based Prediction Models," *Int. J. Recent Technol. Eng.*, vol. 8, no. 1, pp. 1756–1760, 2019.

[4] A. J. P. Delima, A. M. Sison, and R. P. Medina, "Variable Reduction-based Prediction through Modified Genetic Algorithm," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 5, pp. 356–363, 2019. https://doi.org/10.14569/IJACSA.2019.0100544

[5] P. Kaur, M. Singh, and G. S. Josan, "Classification and Prediction Based Data Mining Algorithms to Predict Slow Learners in Education Sector," *Procedia Comput. Sci.*, vol. 57, pp. 500–508, 2015. https://doi.org/10.1016/j.procs.2015.07.372

[6] S. K. Yadav and S. Pal, "Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification," *World Comput. Sci. Inf. Technol. J. WCSIT*, vol. 2, no. 2, pp. 51–56, 2012.

[7] S. B. B, K. K. V, and A. N. Ahmed, "Semantically enriched Tag clustering and image feature based image retrieval system," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 1, pp. 138–141, 2019.

[8] J. Goyal and B. Kishan, "Progress on Machine Learning Techniques for Software Fault Prediction," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 2, pp. 305–311, 2019. https://doi.org/10.30534/ijatcse/2019/33822019

[9] E. Keogh and S. Kasetty, "On the need for time series data mining benchmarks," *Proc. eighth ACM SIGKDD Int. Conf. Knowl. Discov. data Min. - KDD '02*, p. 102, 2002. https://doi.org/10.1145/775047.775062

[10] X. Wang, A. Mueen, H. Ding, G. Trajcevski, P. Scheuermann, and E. Keogh, "Experimental comparison of representation methods and distance measures for time series data," *Data Min. Knowl. Discov.*, vol. 26, no. 2, pp. 275–309, 2013. https://doi.org/10.1007/s10618-012-0250-5

[11] T. Warren Liao, "Clustering of time series data - A survey," *Pattern Recognit.*, vol. 38, no. 11, pp. 1857–1874, 2005. https://doi.org/10.1016/j.patcog.2005.01.025

[12] W. Hongbin, D. Shuangyu, D. Jianzhuo, R. Ming, and D. Ming, "Study on Condition Pre-warning Method of Power Transformer based on Load Time Series Model *," pp. 223–227, 2015. https://doi.org/10.1109/STA.2015.7505103

[13] Q. Liu, X. Liu, B. Jiang, and W. Yang, "Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model," 2011. https://doi.org/10.1186/1471-2334-11-218

[14] S. Promprou, M. Jaroensutasinee, and K. Jaroensutasinee, "Forecasting dengue haemorrhagic fever cases in Southern Thailand using ARIMA Models," *Dengue Bull.*, vol. 30, pp. 99–106, 2006.

[15] D. Billings and Y. Jiann-Shiou, "Application of the ARIMA Models to Urban Roadway Travel Time Prediction-A Case Study. Systems, Man and Cybernetics, 2006. SMC'06," *IEEE Int. Conf.*, pp. 2529–2534, 2006. https://doi.org/10.1109/ICSMC.2006.385244