ISSN 2278-3091
Volume 9, No.1.1, 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse1391.12020.pdf
https://doi.org/10.30534/ijatcse/2020/1391.12020

Crime Data Forecasting using Exponential Smoothing

Everly Chua[1], Gilbert Tumibay[2]
[1]Graduate School, Angeles University Foundation, Holy Angel University, Philippines, echua@hau.edu.ph
[2]Graduate School, Angeles University Foundation, Philippines, tumibay.gibo@auf.edu.ph

ABSTRACT

Crime data analysis to predict the crime that is likely to happen in the future, can be established using mathematical models. Such forecasting is done through data mining techniques to extract relevant information and reveal patterns from existing sets of data. This study investigated forecasting models using Exponential Smoothing (ES) for crime incidents geared towards the development of a resource allocation recommendation system. Comparison of the ES models using different alpha values from 0.2 to 0.7 were subjected to forecast accuracy test using Mean Absolute Deviation (MAD). The Quantitative Forecasting method was used on the dataset which consisted of the crime reports recorded from 2016 to 2018 from the six (6) Angeles City Police Stations (ACPS) under Angeles City Police Office (ACPO) in Region 3. The application of data cleaning process allowed the researchers to pinpoint outliers from the crime data and the crime incidents, summarized according to types of crimes committed within the jurisdiction of the police stations in the locale. The Exponential Smoothing with various models using different alpha values was used. Results showed that such method used is significant in the monthly analysis of crime data. Each of the crime incident types indicate the need for a different alpha value to be used. The police station's resource allocation recommendation system with crime incidents forecasting was made using prototyping methodology. The recommendation feature can help the ACPO in planning its allocation of resources for more efficient logistic management.

Key words: Exponential smoothing, Crime Data, Crime forecasting, Recommendation System.

1. INTRODUCTION

Law enforcement agencies store information about reported crimes and this information is made publicly available in the spirit of open-data. Open data allows the state, local, and all levels of federal government to create, promote, and execute information-based policies since it aims to foster innovation, efficiency, and effectiveness in government services. This help foster collaboration across and within public agencies and departments [1].

Police databases hold a large amount of crime data that can be used to provide information about current and future crime trends and patterns. A good theoretical understanding on the analysis of past crime data is needed to provide practical crime prevention solutions. These approaches involve relating past crime trends with factors that influence the future scope of crime. The analysis of crime related data aims to predict the crime that is likely to happen, and when and where it could possibly happen. Crime is naturally unpredictable, not random, and does not occur consistently in space and time [2].

The use of crime data for predictive analysis in criminological applications is often referred to as predictive policing. The first generation of predictive policing technologies embodies the beginning of a fundamental transformation of how law enforcement prevents crime.

The use of predictive analytics has begun shaping policing strategies in some developed countries. Police districts all over America have adopted predictive policing strategies that embody the goal of policing - stopping crime before it happens. Data from past crimes with details on crime types and locations are fed into a computer algorithm to identify targeted city blocks with a daily and sometimes hourly forecast of crime [3].

Crime prediction relies heavily on extracting relevant information from crime data that typically includes crime type, the committed date and time of the crime, the date and time the crime was reported, and the location of the crime incidents. Crime prediction is the process of finding out crime rate change from one year to the next and forecast the observed changes into the future. Crime predictions can be made through qualitative and quantitative methods using different crime prediction techniques [4].

Qualitative forecasting is often used when there is little or no historical data available on which to base the forecast, while the Quantitative forecasting method makes formal use of historical data and a forecasting model. This method was used since the researchers were able to obtain historical data [5].

The availability of crime data offers a wide array of studies for researches in the field of criminology. The use of prediction algorithm is highly dependent on the approach of a specific research.

In Japan, crime prediction has drawn increasing attention but prediction is a challenge since the crime rate in the country is considerably lower. The study introduced Risk Terrain Modeling (RTM) as a suitable method that does not rely on past crime data and depends mainly on environmental factors associated with crime [6].

Visualization spatial patterns in murder and physical injury in Metro Manila, Philippines was done through choropleth maps. Some demographic variables were investigated to see if there was any relationship between both crimes while accounting for possible spatial auto-correlation using spatial lag models. A recommendation that a system for crime monitoring should include the demographic variables to aid in resource allocation and program planning for better crime prevention and security management [7].

In the study of Malleson et al., the existence of seasonality for a number of different crime types and the variations of seasonality across space were analyzed. Results showed that various crime types do not only exhibit seasonal patterns, but seasonal patterns have relatively distinct spatial patterns [8].

While these researches focused on the forecasting of crime incidence based on space, time and crime type, several criminological researches are consistent in their support for the contribution of populations in crime incidence forecasting.

The study of Awal et al., presented the use of Linear Regression to forecast future crime trends in Bangladesh. In this study, the linear regression model was applied from the crime data of previous years. Different types of crime were forecasted for the year 2016. Results showed the accuracy of linear regression in forecasting the future crime trends. From the experimentation conducted, it was also observed that most of the crimes increased along with the growth of population. The knowledge discovered from crime data analysis may assist police department of Bangladesh and law enforcement agencies to forecast future crime trends for prevention. A recommendation for future study to forecast the location of crime occurrence so that prior actions can be done to prevent crime [9].

Most developed countries such as the United States of America and Europe have already integrated predictive policing innovations [4]. Despite the considerable advances in the field of criminology in the Philippines over the past years, the pace of growth has been quite slow particularly with predictive policing. In the effort to strengthen the Philippine National Police's (PNP) Investigative Functions, the PNP has implemented the Directorate for Investigation and Detective Management (DIDM) IT Solution, a web application used by the Investigation and Detective Management (IDM) unit. One of the core components is the enhanced e-Blotter or Crime Information Reporting and Analysis System (CIRAS) which serves as a crime database and is interfaced with Geographic Information System (GIS). It was designed to evolve into a Qualitative Crime Analysis Management Tool; identify crime hot spots along with other trends and patterns; use spatial (space) and time series for analysis [10].

With the abovementioned features, it can only provide the crime hotspots through the GIS based on the data input. Also, they are required to provide statistical summary of the reported crime on a weekly, monthly, and quarterly basis. This is done manually by mapping the data inputs from the CIRAS to a standard excel format provided by the assigned headquarters. Since provision of real time statistical summary report is not present with the current module, it is likely unable to provide analysis on the trends and patterns using space and time, thus, the module lacks forecasting capability. Rather than performing proactive measurements to prevent crime, they are only able to react to the crime when it has been reported. This is due to the unavailability of the feature that may provide data analysis to forecast the occurrence of crime offenses to a specific area on a specific month it may happen.

This study aims to propose a system that is able to perform statistical analysis using spatial (space) and temporal (time) component of the crime data generated from the CIRAS in forecasting crime occurrences as basis for recommendations in allocating resources. The ability to forecast the occurrence of crime incidents and its locations may serve as a valuable source of information for law enforcement agencies in the preparation of operational plans and in decision-making.

1.1 Research Objectives

Forecasting is a sub-discipline of prediction in which it predicts the future, on the basis of time-series data [11]. There are a number of popular models that can be used with time series forecasting [12]. Some of these models are considered simple while others are complex. The study of Green and Armstrong proved that forecasting procedures should be simple enough for forecast users to understand [13].

Generally, the objective of the study was to investigate forecasting model that best fit in forecasting crime incidents geared towards the development of a recommendation system for Angeles City Police Office (ACPO) allocation of resources. Specifically, it aimed to:

1. Perform pre-processing procedures to analyze the spatiotemporal data related to crime incidents;
2. Apply forecasting model on the processed spatiotemporal data;
3. Compare the forecasted results to determine which forecasting model to use in the police stations resource allocation recommender system;
4. Design a police stations resource allocation recommender system with crime incidents forecasting.

2. RESEARCH METHODOLOGY

Quantitative forecasting method was used in this study. The model summarizes patterns in the data and shows statistical relationship between the previous and current values of the observed variable. The model was then used to project the patterns in the data into the future [5].

The crime data collected were from the six (6) police stations under ACPO. Data cleaning methods were applied in processing the spatiotemporal elements of the crime data. The researchers considered spatiotemporal context information, such as date, time and locations of the reported crime incidents for forecasting. Graphs and maps allowed the researchers to visualize the attributes of crime incidents linked to spatial and temporal data. Alpha values on exponential smoothing was used and compared to analyze the best forecasting model to be used in the recommendation system to be developed.

2.1 Data Preparation

Data cleaning tasks were performed by the researchers using MS Excel tools. The objective of data cleaning tasks is to have a dataset that is uniform and were merged with other related datasets from the various police stations. Pivot table and filters were used to process and summarize the crime data from January 2016 to June 10, 2019. There were 6,847 crimes from the six (6) police stations under ACPO. Simple descriptive graphs and maps were used to visualize attributes of crime incidents linked to spatial and temporal data. This helped the researchers in the analysis of spatial and temporal patterns of crime incidents.

2.2 Modeling Selection and Fitting

Selecting the appropriate modeling technique is needed to generate test scenario for validating the model's quality.
Exponential Smoothing (ES) is a time series forecasting method for univariate data. The ES assigns weights that decreases exponentially as the observation gets older. In this model, recent observations are given more weight in forecasting than the older observations [14]. The formula used in this study was:

$$Ft = F(t-1) + \alpha\,(A(t-1) - F(t-1)) \tag{1}$$

F(t -1) represents the forecast for the previous period. *A(t -1)* is the actual demand for the period, and α represents weight between 0 and 1. The closer to zero, the smaller the weight. Alpha values (α) can be set between 0.1 to 0.9. Comparison of the alpha values in ES was used to determine which model to use in forecasting crime incidents.

2.3 Model Validation

Forecast error is the difference between the forecast and actual value for a given period using the following computation:

$$Et = At - Ft \tag{2}$$

where Et is the forecast error for period t; *At* is the actual value for period t; and *Ft* is the forecast for period t. However, the error for one-time period does not tell much on the accuracy of the forecasting model.

The accuracy of the model can be assessed if the performance of the forecast is measured over time. Tracking the model's performance is important in monitoring forecast errors [14]. Thus, the Mean Absolute Deviation (MAD) was used in this study.

$$MAD = \frac{\Sigma\,|\,Et\,|}{n} \tag{3}$$

where Σ, is the summation; t, is the observation; and n, is the number of observations; *Et* is the difference between the Actual value (*At*) and forecast value (*Ft*).

2.4 System Development

Prototyping was used in the development of the recommendation system for ACPO's resource allocation. It consisted of the following phases: initiation, prototyping cycle and the production phases. The initiation phase is where preliminary data gathering was conducted to determine the contact person in the locale. In the Requirements Gathering phase, the researcher requested for data on crime incidents and sample forms and reports from the locale. Informal interviews were also conducted upon receiving the data and forms from various police station personnel. The gathered data and documents were subjected to document analysis for the researcher to move on the second phase in the prototyping cycle. The next phase of the prototyping cycle is the Objectives, Functions and Performance Requirement.

This is where the product requirements were understood from the locale's perspective. In this phase, detailed communication was conducted with the locale's representative to understand their processes, expectations and exact requirements. Having a clear understanding on the detailed system requirements, the researcher moved on to the Prototype phase. User Interface was developed since it is the junction between the system and its users. It determines the means by which the user and a computer system interact. The researcher will then proceed with the User Acceptance Test (UAT). After the completion of the prototype, the researcher will submit an implementation plan to the locale.

3. RESULTS AND DISCUSSIONS

A. Data Processing

The historical data of the crime incidents was subjected to data pre-processing before analysis and forecasting.

Figure 1 shows the crime reports from the six (6) police stations. The said report details the crimes that were recorded from January 2016 to June 2019. The reports from the precincts have MS Excel formatting on the datasets that would affect the structure of records.

BLOTTER	DATE ENCO	PRO	PPO	STATION	BARANGAY	STREET	DATE REPORTE	TIME REPORTE	DATE COMMITTE	TIME COMMITTEI	OFFENSE
037805-201604-6813		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-03-25	13:30:00	2016-03-25	11:30:00	THEFT
037801-201604-19540		PRO 3	ANGELES CPO	Police Station 1	Pulungbulo		2016-01-01	18:30:00	2016-01-01	16:30:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037805-201604-5909		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-01	02:00:00	2016-01-01	01:00:00	PHYSICAL INJURIES
037805-201604-5920		PRO 3	ANGELES CPO	Police Station 5	Margot		2016-01-02	16:30:00	2016-01-01	22:30:00	ROBBERY
037805-201604-5942		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-05	01:00:00	2016-01-04	23:30:00	HOMICIDE
037805-201604-5991		PRO 3	ANGELES CPO	Police Station 5	Sapangbato		2016-01-10	15:00:00	2016-01-10	15:00:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037803-201604-3784		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-01	13:15:00	2016-01-01	13:00:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037803-201604-3785		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-01	03:40:00	2016-01-01	03:20:00	PHYSICAL INJURIES
037802-201604-10093		PRO 3	ANGELES CPO	Police Station 2	Amsic		2016-01-06	01:30:00	2016-01-06	00:30:00	ROBBERY
037803-201604-3787		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-02	01:15:00	2016-01-02	13:00:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037805-201604-5997		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-11	00:55:00	2016-01-10	21:30:00	THEFT
037803-201604-3791		PRO 3	ANGELES CPO	Police Station 3	Mining		2016-01-04	20:30:00	2016-01-04	20:00:00	PHYSICAL INJURIES
037802-201604-10109		PRO 3	ANGELES CPO	Police Station 2	Amsic		2016-01-06	16:00:00	2016-01-06	15:30:00	UNJUST VEXATION
037803-201604-WCPD211		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-04	13:30:00	2016-01-04	09:00:00	CHILD ABUSE ACT (R.A. 7610)
037802-201604-WCPD 0226		PRO 3	ANGELES CPO	Police Station 2	Sta Teresita		2016-01-10	07:00:00	2016-01-10	04:00:00	CHILD ABUSE ACT (R.A. 7610)
037805-201604-6006		PRO 3	ANGELES CPO	Police Station 5	Margot		2016-01-11	21:00:00	2016-01-11	21:00:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037802-201604-10179		PRO 3	ANGELES CPO	Police Station 2	Pampang		2016-01-13	23:30:00	2016-01-13	22:50:00	HOMICIDE
037805-201604-6013		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-12	14:00:00	2016-01-12	14:00:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037803-201604-3792-A		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-05	02:00:00	2016-01-05	01:15:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037803-201604-3792-B		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-05	02:00:00	2016-01-05	00:45:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037803-201604-3793		PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul		2016-01-06	15:45:00	2016-01-06	15:30:00	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)
037802-201604-10217		PRO 3	ANGELES CPO	Police Station 2	Lourdes Northwest		2016-01-18	00:03:00	2016-01-17	22:45:00	MURDER
037802-201604-10224		PRO 3	ANGELES CPO	Police Station 2	Lourdes Northwest		2016-01-18	16:30:00	2016-01-16	16:00:00	THE ANIMAL WELFARE ACT OF 1998 (R.A. 8485)
037802-201604-WCPD 0227		PRO 3	ANGELES CPO	Police Station 2	Pampang		2016-01-18	11:00:00	2016-01-17	08:00:00	ANTI-VIOLENCE AGAINST WOMEN & THEIR CHILDREN (R.A. 9262)
037805-201604-6024		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-13	08:25:00	2016-01-13	05:00:00	THEFT
037805-201604-6042		PRO 3	ANGELES CPO	Police Station 5	Cutcut		2016-01-14	23:30:00	2016-01-14	22:10:00	ROBBERY

Figure 1: ACPS Crime Reports

Since formatting is going to be a problem when unhandled at an early stage, it was removed to avoid errors in reading the data and applying computations in MS Excel as seen in Figure 2.

BLOTTER NUMBER	PRO	PPO	STATION	BARANGAY	DATE REPORTED	TIME REPORTED	DATE COMMITTED	TIME COMMITTED	INCIDENT TYPE	OFFENSE	OFF
037805-201604-681	PRO 3	ANGELES CPO	Police Station 5	Cutcut	25/03/2016	13:30:00	25/03/2016	11:30:00	Others	THEFT	
037801-201604-195	PRO 3	ANGELES CPO	Police Station 1	Pulungbulo	01/01/2016	18:30:00	01/01/2016	16:30:00	(Operation) Buy Bust	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037805-201604-590	PRO 3	ANGELES CPO	Police Station 5	Cutcut	01/01/2016	2:00:00	01/01/2016	1:00:00	Others	PHYSICAL INJURIES	
037805-201604-592	PRO 3	ANGELES CPO	Police Station 5	Margot	02/01/2016	16:30:00	01/01/2016	22:30:00	(Incident) Robbery	ROBBERY	
037805-201604-594	PRO 3	ANGELES CPO	Police Station 5	Cutcut	05/01/2016	1:00:00	04/01/2016	23:30:00	Others	HOMICIDE	
037805-201604-599	PRO 3	ANGELES CPO	Police Station 5	Sapangbato	10/01/2016	15:00:00	10/01/2016	15:00:00	Others	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037803-201604-378	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	01/01/2016	13:15:00	01/01/2016	13:00:00	Others	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037803-201604-378	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	01/01/2016	3:40:00	01/01/2016	3:20:00	Others	PHYSICAL INJURIES	
037802-201604-100	PRO 3	ANGELES CPO	Police Station 2	Amsic	06/01/2016	1:30:00	06/01/2016	0:30:00	(Incident) Robbery	ROBBERY	
037803-201604-378	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	02/01/2016	1:15:00	02/01/2016	13:00:00	(Operation) Check Point	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037805-201604-599	PRO 3	ANGELES CPO	Police Station 5	Cutcut	11/01/2016	0:55:00	10/01/2016	21:30:00	Others	THEFT	
037803-201604-379	PRO 3	ANGELES CPO	Police Station 3	Mining	04/01/2016	20:30:00	04/01/2016	20:00:00	(Incident) Shooting	PHYSICAL INJURIES	
037802-201604-101	PRO 3	ANGELES CPO	Police Station 2	Amsic	06/01/2016	16:00:00	06/01/2016	15:30:00	Others	UNJUST VEXATION	
037803-201604-WC	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	04/01/2016	13:30:00	04/01/2016	9:00:00	Others	CHILD ABUSE ACT (R.A. 7610)	
037802-201604-WC	PRO 3	ANGELES CPO	Police Station 2	Sta Teresita	10/01/2016	7:00:00	10/01/2016	4:00:00	Others	CHILD ABUSE ACT (R.A. 7610)	
037805-201604-600	PRO 3	ANGELES CPO	Police Station 5	Margot	11/01/2016	21:00:00	11/01/2016	21:00:00	(Operation) Buy Bust	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037802-201604-101	PRO 3	ANGELES CPO	Police Station 2	Pampang	13/01/2016	23:30:00	13/01/2016	22:50:00	Others	HOMICIDE	
037805-201604-601	PRO 3	ANGELES CPO	Police Station 5	Cutcut	12/01/2016	14:00:00	12/01/2016	14:00:00	(Operation) Buy Bust	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037803-201604-379	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	05/01/2016	2:00:00	05/01/2016	1:15:00	(Operation) Check Point	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037803-201604-379	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	05/01/2016	2:00:00	05/01/2016	0:45:00	(Operation) Check Point	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037803-201604-379	PRO 3	ANGELES CPO	Police Station 3	Pulung Maragul	06/01/2016	15:45:00	06/01/2016	15:30:00	(Operation) Buy Bust	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	
037802-201604-102	PRO 3	ANGELES CPO	Police Station 2	Lourdes Northwe	18/01/2016	0:03:00	17/01/2016	22:45:00	(Incident) Shooting	MURDER	
037802-201604-102	PRO 3	ANGELES CPO	Police Station 2	Lourdes Northwe	18/01/2016	16:30:00	16/01/2016	16:00:00	Others	THE ANIMAL WELFARE ACT OF 1998 (R.A. 8485)	
037802-201604-WC	PRO 3	ANGELES CPO	Police Station 2	Pampang	18/01/2016	11:00:00	17/01/2016	8:00:00	Others	ANTI-VIOLENCE AGAINST WOMEN & THEIR CHILDREN (R.A. 9262)	
037805-201604-602	PRO 3	ANGELES CPO	Police Station 5	Cutcut	13/01/2016	8:25:00	13/01/2016	5:00:00	Others	THEFT	
037805-201604-604	PRO 3	ANGELES CPO	Police Station 5	Cutcut	14/01/2016	23:30:00	14/01/2016	22:10:00	(Incident) Robbery	ROBBERY	

Figure 2: Unformatted Data

Figure 3 shows the structured data. Data were structured where unnecessary fields for computation such as street, time and date reported, incident type, stages of felony and offense category were removed. Fields that are linked to spatial and temporal analysis were retained. These are the stations, barangay, and committed crime date and time.

STATION	BARANGAY	OFFENSE	DATE COMMITTED	TIME COMMITTED
Police Station 5	Cutcut	THEFT	2016-03-25	11:30:00
Police Station 1	Pulungbulo	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-01	16:30:00
Police Station 5	Cutcut	PHYSICAL INJURIES	2016-01-01	1:00:00
Police Station 5	Margot	ROBBERY	2016-01-01	22:30:00
Police Station 5	Cutcut	HOMICIDE	2016-01-04	23:30:00
Police Station 5	Sapangbato	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-10	15:00:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-01	13:00:00
Police Station 3	Pulung Maragul	PHYSICAL INJURIES	2016-01-01	3:20:00
Police Station 2	Amsic	ROBBERY	2016-01-06	0:30:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-02	13:00:00
Police Station 5	Cutcut	THEFT	2016-01-10	21:30:00
Police Station 3	Mining	PHYSICAL INJURIES	2016-01-04	20:00:00
Police Station 2	Amsic	UNJUST VEXATION	2016-01-06	15:30:00
Police Station 3	Pulung Maragul	CHILD ABUSE ACT (R.A. 7610)	2016-01-04	9:00:00
Police Station 2	Sta Teresita	CHILD ABUSE ACT (R.A. 7610)	2016-01-10	4:00:00
Police Station 5	Margot	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-11	21:00:00
Police Station 2	Pampang	HOMICIDE	2016-01-13	22:50:00
Police Station 5	Cutcut	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-12	14:00:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-05	1:15:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-05	0:45:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016-01-06	15:30:00
Police Station 2	Lourdes Northwest	MURDER	2016-01-17	22:45:00

Figure 3: Structured Data

As shown in Figure 4, data types in date format were converted into strings to separate the months, days and year when the crimes had been committed. The conversion is necessary to make the application of pivot and filtering easier for the computation.

STATION	BARANGAY	OFFENSE	Year	Month	Date	Day	TIME COMMITTED
Police Station 5	Cutcut	THEFT	2016	Mar	25	Friday	11:30:00
Police Station 1	Pulungbulo	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	1	Friday	16:30:00
Police Station 5	Cutcut	PHYSICAL INJURIES	2016	Jan	1	Friday	1:00:00
Police Station 5	Margot	ROBBERY	2016	Jan	1	Friday	22:30:00
Police Station 5	Cutcut	HOMICIDE	2016	Jan	4	Monday	23:30:00
Police Station 5	Sapangbato	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	10	Sunday	15:00:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	1	Friday	13:00:00
Police Station 3	Pulung Maragul	PHYSICAL INJURIES	2016	Jan	1	Friday	3:20:00
Police Station 2	Amsic	ROBBERY	2016	Jan	6	Wednesday	0:30:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	2	Saturday	13:00:00
Police Station 5	Cutcut	THEFT	2016	Jan	10	Sunday	21:30:00
Police Station 3	Mining	PHYSICAL INJURIES	2016	Jan	4	Monday	20:00:00
Police Station 2	Amsic	UNJUST VEXATION	2016	Jan	6	Wednesday	15:30:00
Police Station 3	Pulung Maragul	CHILD ABUSE ACT (R.A. 7610)	2016	Jan	4	Monday	9:00:00
Police Station 2	Sta Teresita	CHILD ABUSE ACT (R.A. 7610)	2016	Jan	10	Sunday	4:00:00
Police Station 5	Margot	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	11	Monday	21:00:00
Police Station 2	Pampang	HOMICIDE	2016	Jan	13	Wednesday	22:50:00
Police Station 5	Cutcut	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	12	Tuesday	14:00:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	5	Tuesday	1:15:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	5	Tuesday	0:45:00
Police Station 3	Pulung Maragul	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	2016	Jan	6	Wednesday	15:30:00
Police Station 2	Lourdes Northwest	MURDER	2016	Jan	17	Sunday	22:45:00

Figure 4: Data type Conversion

Figure 5 shows the merge data into single data set. This pre-processing task is necessary to combine all reports into a single data set. The reports came from six (6) police stations in Angeles City and have separate files from 2016 to 2018.

Search
(All)
2016
2017
2018
2019
(blank)
Select Multiple Items
OK
Cancel

Year	(Multiple Items)												
		26											
Count of OFFENSE													
Row Labels		**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**	**Grand Total**
ACCESS DEVICES REGULATION ACT 1998 (R.A. 848				1									2
ACTS OF LASCIVIOUSNESS		6	3	3	4	5	3	2	1	7	5	2	44
ADDITIONAL PENALTIES FOR OTHER OFFENSES											1		1
ADULTERY				1	2					1			5
ALARMS AND SCANDAL		5	1	4	1	3	7	3	9	1	2	5	43
ALARMS AND SCANDALS			1			1				1			3
ANTI FENCING LAW (P.D. 1612)		1	2						2	1	1		7
ANTI-CARNAPPING ACT (R.A. 6539) MC		30	27	18	22	31	22	22	20	25	30	26	299
ANTI-CARNAPPING ACT (R.A. 6539) MV		3	1	1	2	2	2	1			2	3	19
ANTI-ELECTRICITY PILFERAGE ACT (R.A. 7832)								1					1
ANTI-SEXUAL HARASSMENT ACT (R.A. 7877)						1				2			3
ANTI-TRAFFICKING IN PERSONS (R.A. 9208)	2			1			4	2	1	1	1		12
ANTI-VIOLENCE AGAINST WOMEN & THEIR CHILDREN (R.A. 9262)	14	20	17	20	18	12	16	12	14	17	5	11	176
ARSON			1		1	1				2	1		6
ASSAULT/RESISTANCE TO AUTHORITY & THEIR AGENTS				1				1			1		3
BP BLG 881 OMNIBUS ELECTION CODE OF THE PHILIPPINES	3	2	5	1	4	1							16

Figure 5: Merged Datasets

Figure 6 shows the output after the researchers selected and removed the blank cells. All crime incidents with blank cells had to be removed as it will affect the accuracy of the computations in monthly data. Out of 6,847 recorded crimes, only 5,387 were used in this study.

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	TOTAL
1	ANTI-CARNAPPING ACT (R.A. 6539) MC	26	30	27	18	22	31	22	22	20	25	30	26	299
2	ANTI-VIOLENCE AGAINST WOMEN & THEIR CHILDREN (R.A. 9262)	14	20	17	20	18	12	16	12	14	17	5	11	176
3	CHILD ABUSE ACT (R.A. 7610)	20	15	19	12	12	18	14	13	15	13	9	9	169
4	COMPREHENSIVE DANGEROUS DRUGS ACT OF 2002 (R.A. 9165)	80	39	80	108	68	95	121	160	171	102	94	79	1197
5	DIRECT ASSAULT	5	7	3	8	5	9	9	14	11	7	13	11	102
6	ESTAFA	17	18	20	9	17	9	19	17	8	11	7	5	157
7	GRAVE THREATS	7	1	7	6	5	2	11	10	9	10	5	3	76
8	HOMICIDE	9	7	7	12	9	7	12	19	8	17	15	18	140
9	HOMICIDE (RECKLESS IMPRUDENCE RESULTING)	2	4	9	7	5	4	4	3	6	6	6	8	64
10	MALICIOUS MISCHIEF	9	8	5	7	4	11	3	8	7	3	4	8	77
11	MALICIOUS MISCHIEF (RIR TO DAMAGE TO PROPERTY)	19	7	13	8	4	11	14	4	7	15	21	23	146
12	MURDER	9	14	7	7	7	10	7	11	9	12	9	10	112
13	PHYSICAL INJURIES	25	12	18	7	17	13	14	12	17	14	7	15	171
14	PHYSICAL INJURY (RECKLESS IMPRUDENCE RESULTING)	99	90	71	53	73	68	69	77	57	69	81	83	890
15	RAPE (Art. 266-A RC & R.A.8353)	12	9	10	6	13	9	9	10	9	10	9	8	114
16	ROBBERY	58	56	61	80	41	64	46	57	52	38	46	27	626
17	THEFT	90	85	90	106	68	73	73	51	64	61	54	56	871
	Grand Total	**501**	**422**	**464**	**474**	**388**	**446**	**463**	**500**	**484**	**430**	**415**	**400**	**5387**

Figure 6: Selecting and Treating Blank Cells

Table 1 shows the 17 crimes used in this study. The researchers made sure that values always occurred in some particular months. There were 97 distinct crime incident types that were classified, however, the study focused on the 17 crime types that frequently occurred based from the historical data.

Table 1: Crime Types

	Crime Type1
1	Anti-Carnapping – Motorcycle
2	Anti-Violence against Women and their Children (R.A. 9262)
3	Child Abuse Act (R.A. 7610)
4	Comprehensive Dangerous Drugs Act 2002
5	Direct Assault
6	Estafa
7	Grave Threats
8	Homicide
9	Homicide – (Reckless Imprudence Resulting to)
10	Malicious Mischief
11	Malicious Mischief – (Reckless Imprudence Resulting to)
12	Murder
13	Physical Injuries
14	Physical Injuries- Reckless Imprudence Resulting to)
15	Rape
16	Robbery
17	Theft

B. Data Analysis

Through simple graph visualization, it can be seen that occurrence of crime incidents is linked to spatial and temporal attributes.

Figure 7 shows the distribution of the crime on a yearly basis according to barangay and police stations' area of jurisdiction.

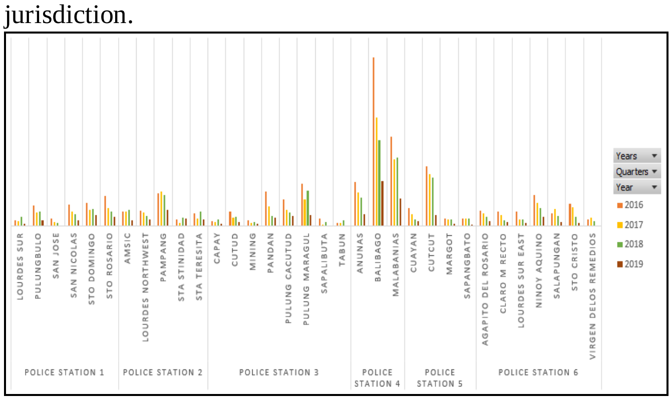

Figure 7: Crime Clusters per Year

Figure 8 shows the trend of crime incidents that occurred in a particular area.

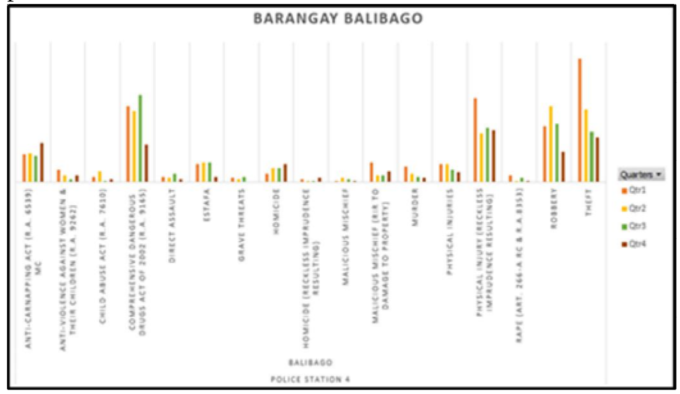

Figure 8: Crime Clusters per Area

C. Exponential Smoothing

Figure 9 shows the forecast computation models of exponential smoothing using the alpha values 0.2, 0.3, 0.4, 0.5, 0.6, 0.7. Results showed that the alpha value of 0.4 is the forecast model that is best to use for the crime type robbery as far as the MAD is taken into consideration.

Alpha (α)	0.2	0.3	0.4	0.5	0.6	0.7
MAD	4.86369	4.67235	4.6566	4.80933	5.01968	5.247
Year	MONTH	ROBBERY	Forecast	Error	Absolute Error	
2016	Jan	24	24			
	Feb	30	24.00	6.00	6.00	
	Mar	30	25.20	4.80	4.80	
	Apr	31	26.16	4.84	4.84	
	May	22	27.13	-5.13	5.13	
	Jun	34	26.10	7.90	7.90	
	Jul	24	27.68	-3.68	3.68	
	Aug	27	26.95	0.05	0.05	
	Sep	24	26.96	-2.96	2.96	
	Oct	25	26.37	-1.37	1.37	
	Nov	15	26.09	-11.09	11.09	
	Dec	10	23.87	-13.87	13.87	
2017	Jan	17	21.10	-4.10	4.10	
	Feb	19	20.28	-1.28	1.28	
	Mar	17	20.02	-3.02	3.02	

Figure 9: Robbery

Table 2 shows the comparative analysis of the alpha values as far as Mean Absolute Deviation (MAD) is taken into considerations. The table indicates that the alpha values of 0.2 and 0.4 yielded the most number times with the lowest MAD. It was however noted that the alpha value of 0.7 is not mathematically far in terms of the number of times it yielded that lowest MAD among the 17 crime incident types. Each of the crime incident type indicates a need to use a different alpha value to achieve the most accurate results in forecasting.

Table 2: Comparison of Alpha Values

	Alpha Values *(α)*						
	0.2	*0.3*	*0.4*	*0.5*	*0.6*	*0.7*	*MAD*
1	4.01	3.75	3.63	3.52	3.42	3.40	0.7
2	1.76	1.63	1.61	1.63	1.70	1.78	0.4
3	2.07	1.88	1.79	1.78	1.82	1.87	0.5
4	14.32	13.48	13.27	13.05	12.79	12.68	0.7
5	1.51	1.57	1.61	1.65	1.69	1.72	0.2
6	2.12	2.15	2.21	2.28	2.34	2.43	0.2
7	1.48	1.50	1.51	1.53	1.55	1.57	0.2
8	2.09	2.16	2.20	2.22	2.28	2.39	0.2

9	1.17	1.15	1.13	1.16	1.21	1.25	0.4
10	1.09	1.07	1.11	1.37	1.17	1.21	0.2
11	3.51	3.29	3.16	3.05	2.96	2.92	0.7
12	1.49	1.57	1.67	1.76	1.86	1.97	0.2
13	2.42	2.39	2.38	2.40	2.47	2.56	0.4
14	7.34	7.05	6.95	6.97	7.00	7.13	0.4
15	1.46	1.38	1.33	1.34	1.35	1.35	0.4
16	4.86	4.67	4.66	4.80	5.02	5.24	0.4
17	5.86	5.38	5.21	5.14	5.08	5.05	0.7

D. Recommendation System

This study was conducted to develop a recommendation system for allocation of resources that included an online information system for Angeles City Police Stations. The resource allocation recommendation system provides the most relevant information on crime incidents through the patterns discovered in the dataset used in this study.

Figure 10 shows that the system can handle the database of the reported crime incidents to get a real-time update on the reported crime in every police station.

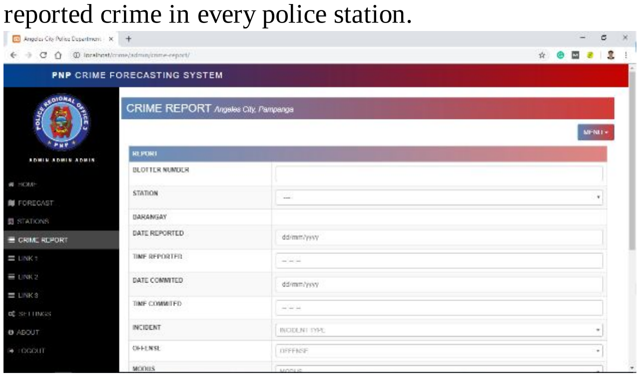

Figure 10: Crime Report - Input Module

Figure 11 indicates that the system can give details of the reported crimes. Each police station can access its own reports, while the Angeles City police headquarters can view consolidated reports.

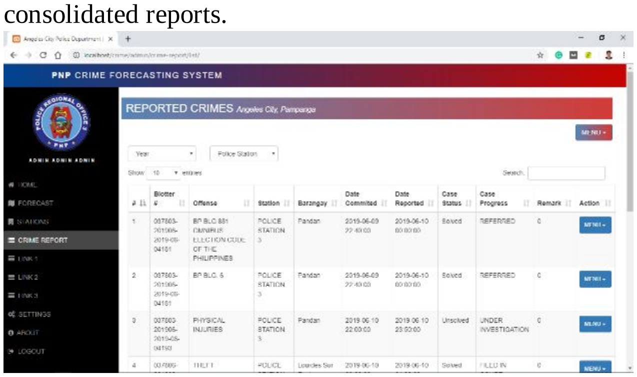

Figure 11: Crime Report - Reported Crimes Module

Figure 12 indicates that the system can show statistical reports of the crimes. This is important in identifying patterns or sudden increase of a crime in an area. Each police station can access its own reports while the Angeles City police headquarters can view consolidated reports.

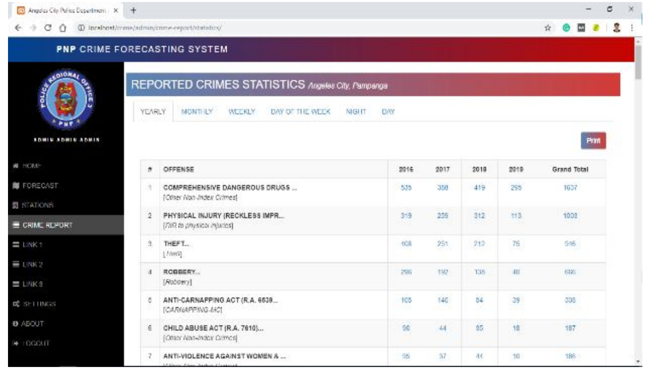

Figure 12: Crime Report – Statistics Module

Figure 13 graphically and numerically indicates that the system can compute forecasted crime incidents. This is important in identifying areas and periods with sudden increase of crime incidents.

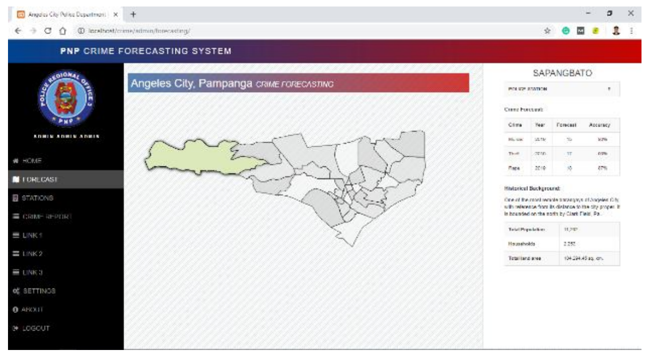

Figure 13: Forecast Module

Figure 14 indicates that the system can give recommended resource allocation based on the population and the number of crimes that happened in various police stations. This is important in allocating resources without bias on any station commander. Each police station can access its own recommended resources, while the Angeles City Police Headquarters can view consolidated recommended allocation.

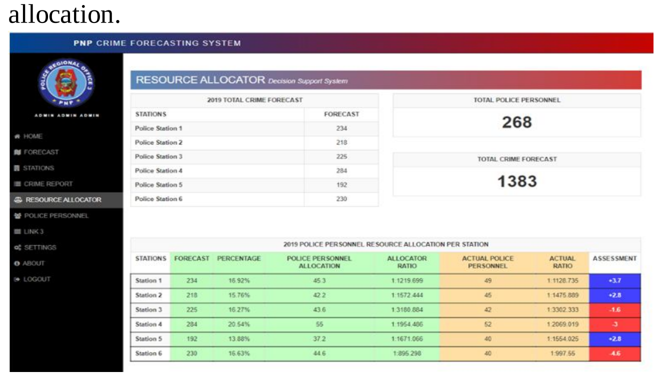

Figure 14: Resource Allocation Module

4. CONCLUSION

The pre-processing procedures were used by the researchers to analyze the spatiotemporal data related to crime incidents. A type of regression model called exponential smoothing was applied on the processed spatiotemporal data. The alpha values considered were 0.2 to 0.7. Each of the crime incident types indicated the need to for different alpha value to be used. This was based on the comparison of the forecasted results. The ES models used in the police stations resource allocation recommender system dynamically suggest alpha values between 0.2 to 0.7 depending on the most accurate forecast using MAD. The police station's resource allocation recommendation system with crime incidents forecasting was built using prototyping methodology. This proposed recommendation system, unlike the existing one would automatically produce statistical reports on all crimes reported. Such real-time update on crime reporting is deemed to be more efficient in generating forecast. The recommendation feature can help the ACPO in planning their allocation of resources. Each police station may come up with operational plans on the deployment of patrols for more efficient logistic management.

Based on this study using simple exponential smoothing method, further studies in the analysis of trends and seasonality on crime data can be conducted applying other time-series forecasting methods that would consider trend and seasonality of crime occurrences.

ACKNOWLEDGEMENT

The study was made possible through the support provided by the Angeles City Police Office Region 3, specifically, the Investigation and Detective Management Branch unit, for providing the compilation of crime reports from the six (6) Angeles City Police Stations and the Personnel and HR Development Branch unit for releasing the strength profile of each police station.

REFERENCES

1. E. Tran and G. Scholtes. **Open data literature review**, 2016.
2. W. Ding, M. W. Ward, C.-H. Yu, and M. Morabito. **Crime forecasting using data mining techniques**, 2011 IEEE 11th Int. Conf. Data Min. Work., vol. 1, pp. 779–786, 2011.
3. A. G. Ferguson. **Policing predictive policing**, vol. 94, no. 5. 2017.
4. N. Dubey and S. K. Chaturvedi. **A survey paper on crime prediction technique using data**, vol. 4, no. 3, pp. 396–400, 2014.
5. D. C. Montgomery. ***Time Series Analysis and Forecasting***, 2nd ed., Canada: John Wiley and Sons, 2015, pp. 4-5.
6. T. Ohyama and M. Amemiya. **Applying crime prediction techniques to Japan**: **A comparison between risk terrain modeling and other methods**, in *Eur. J. Crim. Policy Res.*, vol. 24, no. 4, pp. 469–487, 2018.
   https://doi.org/10.1007/s10610-018-9378-1
7. V. J. Mojica, A. Choi, R. N. Leong, and F. Co. **Spatial analysis of violent crimes in Metro Manila, Philippines**, *Int. J. Comp. Appl. Crim. Justice*, vol. 43, no. 1, pp. 29–47, 2019.
8. N. Malleson and M. A. Andresen. **Journal of criminal justice exploring the impact of ambient population measures on London crime hotspots**, *J. Crim. Justice*, vol. 46, pp. 52–63, 2016.
   https://doi.org/10.1016/j.jcrimjus.2016.03.002
9. A. Awal, J. Rabbi, S. I. Hossain, and M. M. A. Hashem. **Using linear regression to forecast future trends in crime of Bangladesh**, thesis, Dept. Comp. Sci. Eng., Khulna University of Eng. Tech., Bangladesh.
10. I. Reporting, C. I. Database, and C. Management. **PNP next generation investigation**, 2016.
11. M. Doring. **Data science blog**, 2018. [Online]. Available:
   https://www.datascienceblog.net/post/machine-learning/forecasting_vs_prediction/, [Accessed: 19-Aug-2019].
12. A. Sanghani and N. Bhatt. **A review of soft computing techniques for time series forecasting**, December, 2016.
   https://doi.org/10.17485/ijst/2016/v9iS1/99604
13. K. C. Green and J. S. Armstrong. **Simple versus complex forecasting: The evidence**, in *J. Bus. Res.*, 2015.
   https://doi.org/10.2139/ssrn.2643534
14. C. Zaiontz. **Simple exponential smoothing**, 2019. [Online]. Available:
   http://www.real-statistics.com/time-series-analysis/basic-time-series-forecasting/simple-exponential-smoothing/, [Accessed: 01-Aug-2019].