APPLYING DATA ENVELOPMENT ANALYSIS AND CLUSTERING ANALYSIS IN ENHANCING THE PERFORMANCE OF PHILIPPINE NATIONAL POLICE-DISTRICT VI IN THE PROVINCE OF CAVITE

Mengvi P. Gatpandan , Shaneth C. Ambat

School of Graduate Studies, AMA University,Quezon City, Philippines

ABSTRACT

Data envelopment analysis is a technique or method for assessing and evaluating the relative performance of organizational entities where the manifestation of multiple inputs and outputs makes comparison difficult. Efficiency was measured through data envelopment analysis in Philippine National Police District VI in the Province of Cavite to measure the performance of decision-making units in terms of their resources. Clustering is the process of grouping and analyzing the list of objects which have similar characteristics. Clustering algorithm is used in this study to help identify crime pattern. The clustering algorithm was implemented in the application software Crime Management System (CriMS) to predict the crime pattern to help Philippine National Police District VI in the Province of Cavite in decreasing the total number of crime volume and increase the number of crimes solved to countervail security concerns of an individual, community, and the state. Further studies must be conducted to determine the usefulness of the application software by leading to an empirical study on the rule set used to determine the predictive accuracy and/or software productivity.

KEYWORDS

Data Envelopment Analysis, Data Mining, Clustering Techniques, K-means Algorithm,

1. INTRODUCTION

The population shift from rural to urban areas has focused the lenses of urbanization such as the endogenous, modernization and world system. This shift has tremendously posed enormous challenges and pressing serious problems which considerably threaten security issues due to rising urban crimes, expansion of slums, environmental degradation: pollution and vulnerability to flooding [1] [2].

In this changing society, the continuous measures to resist unlawful acts punishable by a state which in any form an act harmful not only to a certain individual but also to the society or to the state. The Philippine Development Plan 2011-2016 recognizes the need for a safe and secure environment as an important factor in fostering investment and the country's economic growth. The protection of the state public order and stability envisions casting out the index and high profile crimes are consistent with the R.A. No. 3815 or the Revised Penal Code of the Philippines which penalizes crimes committed against legal code or law.

The government has the sole power by the mandate of its legislations and crime regulatory bodies to severely restrict one's liberty for the commission of crimes. Philippine National Police (PNP) has defined crime classification as an index and non-index crimes. Index crimes involve 1) crimes against persons such as a) murder, b) homicide, c) physical injury and d) rape, and 2) crimes against property such as a) robbery, b) theft, c) car-napping/carjacking and, d) cattle rustling while non-index crimes are violations of special and private laws such as local ordinances.

DOI:10.5121/ijitca.2017.7201

Data envelopment analysis (DEA) is a technique or method for assessing and evaluating the relative performance of organizational entities where the manifestation of multiple inputs and outputs make comparison difficult [3]. Norman & Stoker [4] noted that DEA is a powerful, analytical technique for evaluating the performance of organizational units in the private or public sector. On their critics and review, Lertworasirikul, Fang, Joines, & Nuttle [5] posited that "data envelopment analysis (DEA) models require crisp input/output data especially in evaluating the performance of activities or organizations services". Charnes, Cooper, & Rhodes [6] used DEA in measuring the efficiency of decision-making units.

Berry & Linoff [7] defines data mining acquaints new methodologies and techniques for a better and informed decision through a careful examination of large databases and considered as knowledge mining from data [8] towards the development of models about aggregated data [9].
A cluster is a group of objects with similar attributes or object characteristics. Data Clustering is the process of grouping and analyzing the list of objects which have similar characteristics. It was exemplified [10] that clustering is "unsupervised learning" and a technique for grouping similar data points [11]. Recent scholarly works show that data mining can aid in crime detection problems and speed up the crime resolution. A clustering algorithm helps identify crime patterns and helps to improve in decreasing the number of crime to countervail security concerns of an individual, community, and the state.

The researchers expressed and presented new detection through the criminological enterprise of crime rampant to the Philippine National Police – District VI in the Province of Cavite to aid the organization and the community. The researchers asked permission to conduct the research study in the Cavite Provincial Police Office (Cavite PPO) to the Officer-in-charge through the Chief of PIDMB that handled the Statistics on Criminal Cases of Index Crime of all Police Station in the Province of Cavite. The historical data available in Cavite PPO and lenses of the research study started from the year 2014 up to the year 2016.

2. LITERATURE REVIEW

Several scholarly works about measuring the efficiency using data envelopment analysis in different organization of DMUs were used. Chan & Karim [12] used a two-stage estimation DEA technique in East Asian economies to determine the relationship between the financial market regulation, country governance, and efficiency of commercial banks during the period of 2001-2008. Kinachi, Najjari, & Alp [13] used DEA and stochastic frontier analysis methods for the scores efficiency and hydroelectricity centers rank to measure the efficiency of 32 Iranian electricity industry. The study used an input-output oriented model using CCR-model and BCC-model in DEA. Osman, Berbary, Sidani, Al-Ayoubi, & Emrouznejad [14] focused on the performance and appraisal evaluation for nurses using a data envelopment analysis. The assessment and relative performance of nurses was and useful for both nurses and hospital in the age of clinical supremacy. Ulucan [15] measured the efficiency in higher education institutions using DEA in Turkish universities using multiple inputs and outputs. Estrada, Song, Kim, Namn, & Kang [16] focused on a dynamic method in benchmarking to identify and measure the inefficient DMUs to improve the efficiency progressively from the dataset collected from the Canadian Bank branches. The study proposed an active method of stepwise benchmarking for the inefficient DMUs.

Predictive analysis uses data mining techniques. Data mining is extracting hidden knowledge, useful and meaningful pattern and trends in a large data set in which organization uses it for decision-making purpose. Grubesic, Wei, & Murray [17] Cluster analysis continues to be an important exploratory technique in scientific inquiry. Several scholarly works used it widely in public health, ecology, geography, and many other fields. Clustering data mining technique can be used for crime detection and crime prevention. Jin-ho & Seung-Ryul [18] used text mining to extract and treat useful information based on the natural language processing, opinion mining to

assigned positive and negative or neutral preference to social media unstructured data, and social network analysis to discover viral objects by measuring a user's reputation or influence based on their connection network and clustering analysis. Ceccato, & Uittenbogaard [19] assessed the crime rates in underground stations using space-time variation analysis to extract patterns for the crime. Tayebi, Ester, Glasser, & Brantingham [20] specialized on spatial crime analysis that focused on crime hotspot areas with disproportionality in which location has higher crime density. Rajagopal [21] research study is all about customer clustering. Segmentation is one of the most important factors used in the study of marketing. Phillips & Lee [22] used to develop a model for density change among spatial regions using density tracing-based approach for the large aggregated crime datasets. Semwal, Vijay Bhaskar, et al., [23] used clustering analysis to identify the gait pattern classification.

3. METHODOLOGY

The research study employs the quantitative method of research. The quantitative research involves formal, objective information with mathematical quantification. During the developmental phase, Cross Industry Standard Process for Data Mining (CRISP-DM) process and model will be utilized.

3.1. RESEARCH DESIGN

Data Envelopment Analysis was used to determine and assess the performance efficiency of the four (4) Police Stations of District VI in the province of Cavite as DMUs in terms of organizational resources such as manpower, physical, financial and technology, and the total volume of index crimes committed per year. DEA assess and evaluate the efficiency of DMUs using the illustrations below:

$$Performance = \frac{virtual\ output}{virtual\ input} = \frac{u_1 y_{1o} + \ldots + u_s y_{so}}{v_1 x_{1o} + \ldots + v_m x_{mo}} \qquad \textit{Eq. (1)}$$

where:

x and y are resp. the input and output vectors and

u_s $output's$ weight, v_m $output'_m s$ weight.

Data Clustering is the process of grouping and analyzing the list of objects which have similar characteristics. The research study will use the partition as the clustering technique and the k-means method as the clustering algorithm. Devi & Rajagopalan [24] k-means clustering is used to cluster observations into groups or clusters of associated observations without any prior idea or knowledge of those relationships. According to Seddawy, Khedr, & Sultan [25], k-means algorithm works as follows:

1. Initialize the center of the cluster

$$\mu_i = \text{some value}, i = 1, \ldots, k \qquad \textit{Eq. (2)}$$

2. Attribute the closest cluster to each data point

$$c_i = \{j : d(\mathbf{x}_j, \mu_i) \leq d(\mathbf{x}_j, \mu_l), l \neq i, j = 1, \ldots, n\} \qquad \textit{Eq. (3)}$$

3. Set the position of each to the mean of all data points belonging to that cluster

$$\mu_i = \frac{1}{|c_i|} \sum_{j \in c_i} \mathbf{x}_j, \forall i \qquad \textit{Eq. (4)}$$

4. Repeat Steps 2-3 until convergence.

Notation: l**c**l = number of elements in c

A CRISP-DM methodology was used as a research design of the study. CRISP-DM provides a structured approach to planning a data mining project in the analytical task in the process of criminal analysis. The researchers used the data mining framework shown below.

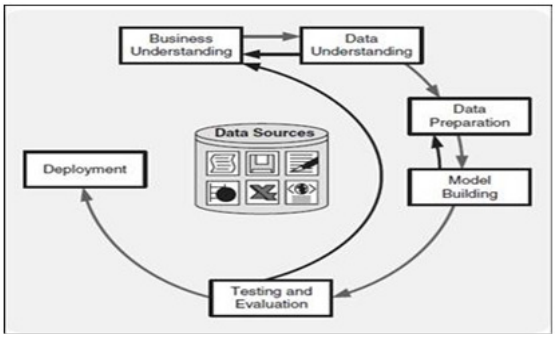

Figure 1.0 CRISP-DM Model

Business Understanding. It focuses on understanding and identifying the project objectives and requirements from a business perspective, and converts the knowledge in a data mining problem definition and provides initial plan design to achieve the desired objectives. The researchers' goal is to help the Philippine National Police – District VI in the Province of Cavite to decrease the index crime rate.

Data Understanding. Initial data gathered and collected from the four (4) police stations at Cavite Provincial Police Office (PPO) situated at Imus, Cavite then determines data quality problem to discover interesting subsets to form an assumption for hidden information.

Data Preparation. Identify, select and prepare data attributes needed to data mining and cleanse data for modeling tools.

Modeling. Selection of clustering modeling techniques, generate test design, form, and assess the model. The k-means algorithm was used for identifying the index crime patterns.

Evaluation. Evaluate and check the model results if it generates the desired results of the study.

Deployment. Provide an organized and presented data report for the intended beneficiaries which includes detailed findings, explanation of models, and others to discuss the initial data mining goals have been met.

3.2 OTHER TOOLS USED IN THE RESEARCH STUDY

MaxDEA software was used to obtain measures of productivity and efficiency to conduct data envelopment analysis. Also, SQL server 2012 with Microsoft Excel 2013 was used to determine crime pattern using clustering technique. Further, Microsoft.Net 4.5 was used for environment framework, Visual Studio 2013 for integrated development environment, Visual C# as programming language, ASP.Net MV5 as development framework for single page application in the web-based environment, SQL server 2012 for data store/persistence, Internet Information Services for web server, jquery 1.10 for javascript framework, jquery easyui for front-end user interface framework, google charts and OLAP pivot graph for charting and visualization, Entity framework 6 for domain entities/models and dapper micro-form for access.

4. RESULTS AND DISCUSSION

The efficiency of the Philippine National Police – District VI was determined through Data Envelopment Analysis using Radial Measure of Efficiency with Input-oriented orientation and a

Rate-to-Scale (RTS) using Scale efficiency. The same result was obtained using input-oriented and output-oriented Chames, Cooper and Rhodes (CCR) model.

Table 1.0 : The variables used in measuring the efficiency of the PNP

Municipality 1, Municipality 2, Municipality 3 and Municipality 4	Personnel	Number of Crimes Cleared and Number of Crimes Solved
	Mobile Units	
	Radio	
	Computer	
	Printer	
	Budget	
	Firearms	
	Total Crimes	

Table 1.0 presents the variables used in measuring the efficiency of the Philippine National Police District VI such as Municipality 1, Municipality 2, Municipality 3 and Municipality 4 as DMUs, personnel, mobile units, radio, computer, printer, budget, firearms and total crimes as inputs, and a number of crimes cleared and crimes solved as output.

Table 2.0 : 3-Year Scale Efficiency Results of the PNP – District VI in the province of Cavite

DMU	2014		2015		2016		Average	Ranking
	Technical Efficiency Score	Scale Efficiency Score	Technical Efficiency Score	Scale Efficiency Score	Technical Efficiency Score	Scale Efficiency Score		
Municipality 1	0.27	0.27	0.59	0.59	1.00	1.00	0.62	3
Municipality 2	0.82	0.82	1.00	1.00	0.72	0.72	0.85	2
Municipality 3	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1**
Municipality 4	1.00	1.00	0.56	0.67	0.94	0.94	0.85	2

Table 2.0 presents the 3-year Scale Efficiency Results of the Philippine National Police – District VI in the province of Cavite. Municipality 1 got the average score of 0.62 scale efficiencies and ranked as the lowest inefficient DMU; Municipality 2 and Municipality 4 got the average score of 0.85 scale of efficiency and ranked as 2nd inefficient DMU. Among the 4 DMUs, Municipality 3 got an average score of 1.00 scale of efficiency in 3 years. Therefore, Municipality 3 was the efficient DMU of the Philippine National Police – District VI in the Province of Cavite.

The following shows the clustering technique using data mining tools in MS SQL Server 2012 using Microsoft Excel 2013 with the raw data of criminal cases of District VI for three (3) years historical data from January 2014 up to December 2016.

Table 3.0 : Cluster Characteristics of Municipality 1 by Crime Category Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
CrimeCategory	THEFT	34%	MURDER	30%	RAPE	24%
CrimeCategory	ROBBERY	28%	THEFT	17%	PHYSICALINJURIES	24%
CrimeCategory	MURDER	16%	RAPE	17%	MURDER	18%
CrimeCategory	PHYSICALINJURIES	9%	ROBBERY	17%	THEFT	18%
CrimeCategory	RAPE	6%	PHYSICALINJURIES	13%	MOTORNAPPING	12%
CrimeCategory	CARNAPPING	3%	MOTORNAPPING	4%	ROBBERY	6%
CrimeCategory	HOMICIDE	3%				

Table 3.0 shows the cluster characteristics of Municipality 1 by Crime Category year 2014 – 2016. The summary shows that *theft* ranked the highest in the year 2014 with 34% probability, *murder* in the year 2015 with 30% probability and *rape* in the year 2016 with 24% probability.

Rule Set for Year 2016

Cluster 1: CrimeCategory=PHYSICALINJURIES, CrimeCategory=RAPE,

CrimeCategory=MOTORNAPPING

Cluster 2: CrimeCategory=MURDER, CrimeCategory=THEFT, CrimeCategory=ROBBERY

Table 4.0 : Cluster Characteristics of Municipality 1 by Barangay Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
Barangay	Buho	9%	Talon	13%	Talon	24%
Barangay	Maitim I	9%	Barangay II	9%	Halang	12%
Barangay	Pangil	9%	Maitim I	9%	Buho	12%
Barangay	Maymangga	6%	Halang	9%	Bucal	12%
Barangay	Loma	6%	Buho	9%	Barangay III	6%
Barangay	Halang	6%	Salaban	9%	Dagatan	6%
Barangay	Barangay VII	6%	Barangay XI	9%	Barangay XI	6%
Barangay	Barangay VI	3%	Pangil	4%	Tamacan	6%
Barangay	Talon	3%	Barangay V	4%	Salaban	6%
Barangay	Poblacion 6	3%	Loma	4%	Barangay IV	6%
Barangay	Barangay VIII	3%	Barangay X	4%	Cabuco	6%
Barangay	Salaban	3%	Tamacan	4%		
Barangay	Dagatan	3%	Barangay VII	4%		
Barangay	Barangay IX	3%	Dagatan	4%		
Barangay	Poblacion 5	3%	Barangay III	4%		
Barangay	Poblacion 10	3%				
Barangay	Banaybanay	3%				
Barangay	Tamacan	3%				
Barangay	Barangay I	3%				
Barangay	Barangay XII	3%				
Barangay	Barangay II	3%				
Barangay	Minantok Kanluran	3%				

Table 4.0 shows the cluster characteristics of Municipality 1 by Barangay year 2014 – 2016. The summary shows that *Buho* ranked the highest in the year 2014 with 9% probability and Talon in the year 2015 and 2016 with 13% and 24% probability.

Table 5.0 : Cluster Characteristics of Municipality 1 by Seasonal (Month) Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
MonthComttd	July	13%	April	17%	August	18%
MonthComttd	November	13%	August	17%	April	18%
MonthComttd	March	9%	November	13%	September	12%
MonthComttd	December	9%	December	9%	June	12%
MonthComttd	January	9%	June	9%	May	12%
MonthComttd	April	9%	September	9%	January	12%
MonthComttd	August	6%	February	9%	March	12%
MonthComttd	September	6%	March	4%	November	6%
MonthComttd	October	6%	October	4%		
MonthComttd	May	6%	July	4%		
MonthComttd	February	6%	January	4%		
MonthComttd	June	6%				

Table 5.0 shows the cluster characteristics of Municipality 1 by Seasonal (Month) 2014 – 2016. The summary shows that *July* ranked the highest in the year 2014 with 13% probability, *April* and *August* in the year 2015 and 2016 with 17% and 18% probability.

Rule Set for Year 2014

Cluster 5: MonthComttd=June, MonthComttd=August, MonthComttd=February, MonthComttd=January, MonthComttd=October, MonthComttd=September
Cluster 1: MonthComttd=July, MonthComttd=March
Cluster 4: MonthComttd=December, MonthComttd=January, MonthComttd=August, MonthComttd=September, MonthComttd=June, MonthComttd=May

Rule Set Year 2015

Cluster 4: MonthComttd=November, MonthComttd=February, MonthComttd=June,

MonthComttd=March, MonthComttd=January
Cluster 1: MonthComttd=April, MonthComttd=July, MonthComttd=December
Cluster 3: MonthComttd=September, MonthComttd=October, MonthComttd=February, MonthComttd=July, MonthComttd=January, MonthComttd=June, MonthComttd=March, MonthComttd=December
Cluster 2: MonthComttd=August, MonthComttd=November

Table 6.0 : Cluster Characteristics of Municipality 2 by Crime Category Year 2014 - 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
CrimeCategory	THEFT	39%	THEFT	28%	THEFT	31%
CrimeCategory	PHYSICALINJURIES	24%	PHYSICALINJURIES	22%	ROBBERY	20%
CrimeCategory	ROBBERY	15%	ROBBERY	20%	PHYSICALINJURIES	18%
CrimeCategory	MURDER	10%	MURDER	10%	MURDER	10%
CrimeCategory	RAPE	9%	RAPE	8%	MOTORNAPPING	8%
CrimeCategory	CARNAPPING	2%	MOTORNAPPING	5%	RAPE	6%
CrimeCategory	HOMICIDE	1%	CARNAPPING	4%	CARNAPPING	5%
CrimeCategory			HOMICIDE	2%	HOMICIDE	2%

Table 6.0 shows the cluster characteristics of Municipality 2 by Crime Category year 2014 – 2016. The summary shows that *Theft* ranked the highest in the year 2014 - 2016 with 39%, 28%, and 31% probability.

Rule Set for Year 2014

Cluster 1: CrimeCategory=THEFT
Cluster 2: CrimeCategory=PHYSICALINJURIES, CrimeCategory=MURDER
Cluster 4: CrimeCategory=HOMICIDE, CrimeCategory=CARNAPPING, CrimeCategory=RAPE, CrimeCategory=MURDER, CrimeCategory=ROBBERY
Cluster 3: CrimeCategory=ROBBERY, CrimeCategory=RAPE, CrimeCategory=MURDER

Rule Set Year 2015

Cluster 3: CrimeCategory=PHYSICALINJURIES, CrimeCategory=RAPE, CrimeCategory=CARNAPPING, CrimeCategory=MURDER, CrimeCategory=HOMICIDE
Cluster 1: CrimeCategory=THEFT, CrimeCategory=MOTORNAPPING
Cluster 2: CrimeCategory=ROBBERY

Rule Set Year 2016

Cluster 3: CrimeCategory=ROBBERY, CrimeCategory=MOTORNAPPING, CrimeCategory=RAPE, CrimeCategory=CARNAPPING, CrimeCategory=HOMICIDE
Cluster 1: CrimeCategory=THEFT
Cluster 2: CrimeCategory=PHYSICALINJURIES, CrimeCategory=MURDER, CrimeCategory=CARNAPPING, CrimeCategory=HOMICIDE, CrimeCategory=MOTORNAPPING

Table 7.0 : Cluster Characteristics of Municipality 2 by Barangay Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
Barangay	Bacao I	9%	Manggahan	9%	Bacao I	10%
Barangay	Manggahan	7%	Bacao I	9%	Manggahan	10%
Barangay	Alingaro	6%	Tapia	6%	Panungyanan	5%
Barangay	Tejero	5%	Tejero	6%	Tapia	5%
Barangay	San Francisco	4%	Javalera	4%	Javalera	5%
Barangay	Tapia	4%	Biclatan	4%	Alingaro	4%
Barangay	Buenavista III	4%	San Francisco	4%	Tejero	4%
Barangay	Arnaldo	4%	Prinza	4%	Sta. Clara	3%
Barangay	Santiago	3%	Panungyanan	4%	Bacao II	3%
Barangay	Bacao II	3%	Santiago	3%	Vibora	3%
Barangay	Pinagtipunan	3%	San Gabriel	3%	Biclatan	3%
Barangay	Bagumbayan	3%	Bacao II	3%	Navarro	3%
Barangay	Navarro	3%	Alingaro	3%	San Francisco	3%
Barangay	Dulong Bayan	3%	Governor Ferrer	3%	Santiago	3%
Barangay	Biclatan	3%	San Juan I	3%	San Gabriel	3%
Barangay	Javalera	3%	Buenavista I	3%	Prinza	3%
Barangay	Prinza	3%	Arnaldo	3%	Buenavista II	2%
Barangay	Buenavista II	3%	Buenavista III	2%	Arnaldo	2%
Barangay	Vibora	3%	Pasong Camachile I	2%	Governor Ferrer	2%
Barangay	Sampalucan	3%	Corregidor	2%	Corregidor	2%
Barangay	Governor Ferrer	2%	Pasong Kawayan II	2%	San Juan II	2%
Barangay	Pasong Camachile I	2%	Pasong Camachile II	2%	Pasong Camachile I	2%
Barangay	San Juan I	2%	Sampalucan	2%	Pasong Camachile I	2%
Barangay	96th	2%	Bagumbayan	2%	San Juan I	2%
Barangay	Corregidor	2%	Navarro	2%	96th	2%
Barangay	Panungyanan	2%	Pinagtipunan	2%	Buenavista I	1%
Barangay	San Gabriel	2%	Buenavista II	2%	Sampalucan	1%
Barangay	Pasong Camachile II	1%	San Juan II	2%	Buenavista III	1%
Barangay	San Juan II	1%	Vibora	2%	Pasong Kawayan I	1%
Barangay	Sta. Clara	1%	Pasong Kawayan I	1%	Dulong Bayan	1%
Barangay	Buenavista I	1%	Dulong Bayan	1%	Bagumbayan	1%
Barangay	Pasong Kawayan II	1%	96th	1%	Pasong Kawayan II	1%
Barangay	Pasong Kawayan I	1%	Sta. Clara	1%	Pinagtipunan	1%

Table 7.0 shows the cluster characteristics of Municipality 2 by Barangay 2014 – 2016. The summary shows that *Bacao* ranked the highest in the year 2014 - 2016 with 9%, 9%, and 10% probability.

Rule Set for Year 2014

Cluster 4: Barangay=Tapia, Barangay=Pasong Camachile I, Barangay=Prinza, Barangay=Buenavista II, Barangay=Tejero, Barangay=Javalera, Barangay=Bacao II, Barangay=Sta. Clara, Barangay=96th, Barangay=San Juan I, Barangay=Corregidor, Barangay=Bagumbayan, Barangay=Governor Ferrer, Barangay=Buenavista I, Barangay=Pasong Kawayan II, Barangay=Mangahan, Barangay=Buenvavista III, Barangay=Vibora, Barangay=Pasong Kawayan, Barangay=San Gabriel

Cluster 1: Barangay=Pinagtipunan, Barangay=Dulong Bayan, Barangay=Bacao I, Barangay=Buenavista III, Barangay=Buenavista II, Barangay=Panungyanan, Barangay=Sampalucan, Barangay=San Juan II, Barangay=Mangahan, Barangay=Pasong Camachile II, Barangay=Governor Ferrer, Barangay=Pasong Kawayan II

Cluster 3: Barangay=Arnaldo, Barangay=Biclatan, Barangay=Santiago, Barangay=San Francisco, Barangay=Vibora, Barangay=Pasong Kawayan I, Barangay=San Gabriel, Barangay=Pasong Camachile II, Barangay=Panungyanan, Barangay=San Juan I, Barangay=Pasong Kawayan, Barangay=Sampalucan, Barangay=Sta. Clara, Barangay=Buenvavista III, Barangay=Tejero

Cluster 2: Barangay=Manggahan, Barangay=Alingaro, Barangay=Navarro, Barangay=San Juan II, Barangay=Sampalucan, Barangay=Bagumbayan, Barangay=Tapia

Rule Set for Year 2015

Cluster 8: Barangay=Arnaldo, Barangay=Pasong Camachile II, Barangay=Buenavista III, Barangay=Pinagtipunan, Barangay=Governor Ferrer, Barangay=Bacao II, Barangay=Panungyanan,Barangay=Pasong Kawayan II, Barangay=Pasong Kawayan I, Barangay=Navarro, Barangay=San Juan II, Barangay=Corregidor, Barangay=Sampalucan, Barangay=San Gabriel, Barangay=Bagumbayan, Barangay=Sta. Clara

Cluster 6: Barangay=Pasong Kawayan II, Barangay=Corregidor, Barangay=Navarro, Barangay=Buenavista II, Barangay=Vibora, Barangay=Pinagtipunan, Barangay=Sampalucan, Barangay=Pasong Camachile I, Barangay=96th, Barangay=Bagumbayan, Barangay=Bacao II, Barangay=Panungyanan, Barangay=Sta. Clara

Cluster 1: Barangay=Bacao I, Barangay=Tapia

Cluster 4: Barangay=Prinza, Barangay=San Francisco, Barangay=Buenavista III, Barangay=San Juan II, Barangay=Pasong Camachile I, Barangay=Dulong Bayan, Barangay=Buenavista II, Barangay=Vibora, Barangay=Sampalucan, Barangay=Panungyanan, Barangay=Corregidor, Barangay=96th, Barangay=Pasong Kawayan I, Barangay=Pinagtipunan, Barangay=Sta. Clara, Barangay=Pasong Kawayan II

Cluster 3: Barangay=Tejero, Barangay=Javalera, Barangay=San Juan I

Cluster 2: Barangay=Manggahan, Barangay=San Francisco, Barangay=Pasong Kawayan I,Barangay=96^{th}

Cluster 5: Barangay=Alingaro, Barangay=Biclatan, Barangay=Pasong Camachile II, Barangay=Pasong Camachile I, Barangay=Dulong Bayan, Barangay=Navarro, Barangay=Corregidor, Barangay=Bacao II, Barangay=Buenavista II, Barangay=Vibora, Barangay=Buenavista I

Cluster 7: Barangay=Santiago,, Barangay=Arnaldo, Barangay=Panungyanan, Barangay=Vibora, Barangay=San Gabriel, Barangay=Bacao II, Barangay=Sampalucan, Barangay=Pasong Camachile I, Barangay=Pinagtipunan, Barangay=Navarro, Barangay=96th, Barangay=Bagumbayan, Barangay=Buenavista II, Barangay=Pasong Kawayan II, Barangay=Sta. Clara, Barangay=Buenavista I

Rule Set for Year 2016

Cluster 6: Barangay=Sta. Clara, Barangay=San Juan II, Barangay=Bacao II, Barangay=96th, Barangay=Pasong Kawayan I,Barangay=Buenavista I, Barangay=Bagumbayan, Barangay=Buenavista II, Barangay=Pinagtipunan, Barangay=Buenavista III, Barangay=Navarro, Barangay=Dulong Bayan, Barangay=San Juan I, Barangay=Biclatan, Barangay=Corregidor, Barangay=Arnaldo, Barangay=Pasong Kawayan II, Barangay=Prinza, Barangay=Pasong Kwayan I, Barangay=Pasong Camachile I, Barangay=Sampalucan, Barangay=Governor Ferrer, Barangay=Santiago

Cluster 1: Barangay=Alingaro, Barangay=San Gabriel, Barangay=Manggahan

Cluster 4: Barangay=Tejero, Barangay=Pasong Kawayan II, Barangay=San Francisco, Barangay=Navarro, Barangay=Santiago, Barangay=Governor Ferrer, Barangay=San Juan II, Barangay=Biclatan, Barangay=Prinza, Barangay=Arnaldo, Barangay=Buenavista II, Barangay=Pasong Camachille I, Barangay=Buenavista III, Barangay=San Juan I, Barangay=Corregidor, Barangay=Pinagtipunan, Barangay=Dulong Bayan

Cluster 2: Barangay=Javalera, Barangay=Bacao I

Cluster 5: Barangay=Pasong Camachile II, Barangay=Pasong Camachile I, Barangay=Santiago, Barangay=Corregidor, Barangay=96th, Barangay=Prinza,

Barangay=Sampalucan, Barangay=Governor Ferrer, Barangay=San Francisco, Barangay=Bagumbayan, Barangay=Biclatan, Barangay=Bacao II, Barangay=Pasong Kawayan I, Barangay=Sta. Clara, Barangay=Dulong Bayan, Barangay=Buenavista I, Barangay=Buenavista III, Barangay=Pasong Camachille I, Barangay=Pasong Kwayan I, Barangay=Navarro, Barangay=Pinagtipunan, Barangay=San Juan I

Cluster 3: Barangay=Panungyanan, Barangay=Tapia, Barangay=Vibora

Table 8.0 : Cluster Characteristics of Municipality 2 by Seasonal (Month) Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
MonthComttd	March	11%	March	13%	March	16%
MonthComttd	February	10%	December	10%	September	13%
MonthComttd	May	10%	May	10%	July	12%
MonthComttd	July	9%	January	10%	January	8%
MonthComttd	April	9%	July	8%	August	8%
MonthComttd	January	9%	February	8%	October	7%
MonthComttd	June	8%	October	8%	April	7%
MonthComttd	November	7%	April	8%	November	7%
MonthComttd	October	7%	November	7%	May	6%
MonthComttd	September	6%	June	7%	February	5%
MonthComttd	December	6%	August	6%	June	5%
MonthComttd	August	6%	September	5%	December	5%

Table 8.0 shows the cluster characteristics of Municipality 2 by Seasonal (Month) 2014 – 2016. The summary shows that *March* ranked the highest in the year 2014 - 2016 with 11%, 13%, and 16% probability.

Rule Set for Year 2014

Cluster 5: MonthComttd=August, MonthComttd=October, MonthComttd=January, MonthComttd=November, MonthComttd=September, MonthComttd=December

Cluster 6: MonthComttd=September, MonthComttd=June, MonthComttd=December, MonthComttd=November, MonthComttd=August

Cluster 1: MonthComttd=September, MonthComttd=June, MonthComttd=December, MonthComttd=November, MonthComttd=August

Cluster 2: MonthComttd=March, MonthComttd=December, MonthComttd=October

Cluster 3: MonthComttd=February, MonthComttd=January

Cluster 4: MonthComttd=April, MonthComttd=October, MonthComttd=August, MonthComttd=June

Rule Set for Year 2015

Cluster 6: MonthComttd=July, MonthComttd=June, MonthComttd=November, MonthComttd=September, MonthComttd=October

Cluster 1: MonthComttd=January, MonthComttd=December

Cluster 5: MonthComttd=August, MonthComttd=February, MonthComttd=June, MonthComttd=July, MonthComttd=October, MonthComttd=November, MonthComttd=September

Cluster 2: MonthComttd=March, MonthComttd=August

Cluster 4: MonthComttd=May, MonthComttd=November, MonthComttd=October

Cluster 3: MonthComttd=April, MonthComttd=February, MonthComttd=September,MonthComttd=October

Rule Set for Year 2016

Cluster 1: MonthCom=July, MonthCom=September

Cluster 5: MonthCom=January, MonthCom=May, MonthCom=June, MonthCom=February, MonthCom=October

Cluster 6: MonthCom=November, MonthCom=February, MonthCom=October,

MonthCom=December, MonthCom=May, MonthCom=June, MonthCom=January
Cluster 2: MonthCom=March
Cluster 4: MonthCom=April, MonthCom=December, MonthCom=October, MonthCom=November, MonthCom=May
Cluster 3: MonthCom=August, MonthCom=April, MonthCom=November, MonthCom=January

Table 9.0 : Cluster Characteristics of Municipality 3 by Crime Category Year 2014 - 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
CrimeCategory	PHYSICALINJURIES	31%	PHYSICALINJURIES	26%	MURDER	29%
CrimeCategory	THEFT	24%	MURDER	21%	RAPE	23%
CrimeCategory	RAPE	16%	THEFT	18%	THEFT	18%
CrimeCategory	MURDER	14%	ROBBERY	11%	PHYSICALINJURIES	14%
CrimeCategory	ROBBERY	12%	RAPE	10%	ROBBERY	9%
CrimeCategory	CARNAPPING	2%	MOTORNAPPING	7%	MOTORNAPPING	4%
CrimeCategory	HOMICIDE	1%	CARNAPPING	5%	CARNAPPING	1%
CrimeCategory			HOMICIDE	2%	HOMICIDE	1%

Table 9.0 shows the cluster characteristics of Municipality 3 by Crime Category 2014 – 2016. The summary shows that *Physical Injuries* ranked the highest in the year 2014 and 2015 with 31%, and 26% probability, and *Murder* in the year 2016 with 29% probability.

Rule Set for Year 2015

Cluster 3: CrimeCategory=THEFT, CrimeCategory=ROBBERY, CrimeCategory=RAPE, CrimeCategory=HOMICIDE, CrimeCategory=CARNAPPING
Cluster 2: CrimeCategory=MURDER, CrimeCategory=MOTORNAPPING, CrimeCategory=CARNAPPING
Cluster 1: CrimeCategory=PHYSICALINJURIES

Rule Set for Year 2016

Cluster 2: CrimeCategory=RAPE, CrimeCategory=PHYSICALINJURIES, CrimeCategory=MOTORNAPPING
Cluster 1: CrimeCategory=MURDER, CrimeCategory=CARNAPPING, CrimeCategory=MOTORNAPPING
Cluster 3: CrimeCategory=THEFT, CrimeCategory=ROBBERY, CrimeCategory=HOMICIDE, CrimeCategory=CARNAPPING

Table 10.0 : Cluster Characteristics of Municipality 3 by Barangay Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
Barangay	Halayhay	17%	Halayhay	13%	Halayhay	25%
Barangay	Bagtas	14%	Bagtas	13%	Calibuyo	11%
Barangay	Capipisa	8%	Biwas	9%	Capipisa	11%
Barangay	Bucal	8%	Capipisa	8%	Bagtas	8%
Barangay	Amaya II	7%	Bunga	7%	Biwas	8%
Barangay	Biwas	6%	Amaya I	7%	Amaya II	7%
Barangay	Calibuyo	6%	Calibuyo	6%	Biga	5%
Barangay	Amaya I	5%	Bucal	6%	Bucal	5%
Barangay	Daang Amaya II	4%	Amaya VII	4%	Daang Amaya I	3%
Barangay	Bunga	4%	Daang Amaya I	4%	Daang Amaya III	3%
Barangay	Amaya V	4%	Amaya II	4%	Daang Amaya II	3%
Barangay	Amaya III	3%	Amaya III	4%	Bunga	3%
Barangay	Biga	3%	Amaya V	3%	Amaya I	2%
Barangay	Amaya VI	3%	Biga	3%	Amaya V	2%
Barangay	Daang Amaya III	2%	Daang Amaya II	3%	Amaya III	1%
Barangay	Amaya VII	2%	Amaya VI	3%	Amaya VI	1%
Barangay	Daang Amaya I	2%	Amaya IV	2%	Amaya VII	1%
Barangay	Amaya IV	2%	Daang Amaya III	2%		

Table 9.0 shows the cluster characteristics of Municipality 3 by Barangay 2014 – 2016. The summary shows that *Halayhay* ranked the highest in the year 2014 - 2016 with 17%, 13%, and 25% probability.

Rule Set for Year 2014

Cluster 5: Barangay=Amaya I, Barangay=Amaya II, Barangay=Amaya VI, Barangay=Daang Amaya II, Barangay=Calibuyo, Barangay=Amaya V, Barangay=Daang Amaya III, Barangay=Daang Amaya I, Barangay=Amaya VII, Barangay=Biga
Cluster 6: Barangay=Amaya III, Barangay=Bunga, Barangay=Biga, Barangay=Daang Amaya II, Barangay=Daang Amaya III, Barangay=Calibuyo, Barangay=Amaya IV, Barangay=Daang Amaya I, Barangay=Amaya VI, Barangay=Amaya II, Barangay=Amaya VII
Cluster 1: Barangay=Halayhay
Cluster 2: Barangay=Bagtas
Cluster 3: Barangay=Biwas, Barangay=Amaya I, Barangay=Amaya IV, Barangay=Calibuyo, Barangay=Amaya II, Barangay=Bunga, Barangay=Amaya VII, Barangay=Amaya III
Cluster 7: Barangay=Bucal, Barangay=Amaya V, Barangay=Calibuyo, Barangay=Daang Amaya I, Barangay=Amaya II
Cluster 4: Barangay=Capipisa, Barangay=Amaya III, Barangay=Amaya V, Barangay=Biga, Barangay=Amaya IV, Barangay=Amaya VII, Barangay=Amaya I

Rule Set for Year 2015

Cluster 7: Barangay=Amaya III, Barangay=Bunga, Barangay=Amaya VII, Barangay=Amaya VI, Barangay=Amaya V, Barangay=Amaya IV, Barangay=Biga, Barangay=Calibuyo, Barangay=Daang Amaya II, Barangay=Amaya II
Cluster 8: Barangay=Bucal, Barangay=Amaya V, Barangay=Amaya II, Barangay=Daang Amaya II, Barangay=Biga, Barangay=Bunga, Barangay=Daang Amaya III, Barangay=Amaya VII, Barangay=Amaya VI
Cluster 1: Barangay=Capipisa, Barangay=Amaya I
Cluster 3: Barangay=Halayhay
Cluster 2: Barangay=Bagtas, Barangay=Bunga
Cluster 4: Barangay=Biwas, Barangay=Amaya VI, Barangay=Amaya VII
Cluster 5: Barangay=Daang Amaya I, Barangay=Amaya IV, Barangay=Amaya II, Barangay=Calibuyo, Barangay=Bucal, Barangay=Daang Amaya II, Barangay=Biga, Barangay=Daang Amaya III, Barangay=Amaya III, Barangay=Amaya VII, Barangay=Amaya V
Cluster 6: Barangay=Calibuyo, Barangay=Bunga, Barangay=Daang Amaya I, Barangay=Amaya VI, Barangay=Amaya II, Barangay=Bucal, Barangay=Biga, Barangay=Amaya VII, Barangay=Amaya III, Barangay=Daang Amaya III, Barangay=Daang Amaya II, Barangay=Amaya IV

Rule Set for Year 2016

Cluster 6: Barangay=Amaya II, Barangay=Bagtas, Barangay=Biga, Barangay=Daang Amaya I, Barangay=Amaya I, Barangay=Bucal, Barangay=Bunga, Barangay=Amaya V, Barangay=Amaya III, Barangay=Amaya VI, Barangay=Amaya VII
Cluster 1: Barangay=Halayhay
Cluster 4: Barangay=Biwas,Barangay=Bagtas, Barangay=Daang Amaya III, Barangay=Daang Amaya II, Barangay=Amaya III, Barangay=Amaya V, Barangay=Amaya VII, Barangay=Amaya I, Barangay=Daang Amaya I,

Barangay=Biga, Barangay=Bunga
Cluster 3: Barangay=Calibuyo, Barangay=Biga, Barangay=Daang Amaya II
Cluster 2: Barangay=Capipisa, Barangay=Amaya II, Barangay=Bucal,
Barangay=Amaya VI, Barangay=Amaya VII,
Barangay=Daang Amaya III, Barangay=Amaya I
Cluster 5: Barangay=Bunga, Barangay=Bucal, Barangay=Bagtas, Barangay=Biga,
Barangay=Amaya II, Barangay=Daang Amaya II, Barangay=Daang Amaya I,
Barangay=Amaya I, Barangay=Amaya VI, Barangay=Daang Amaya III,
Barangay=Amaya VII, Barangay=Amaya V, Barangay=Amaya III

Table 11.0 : Cluster Characteristics of Municipality 3 by Seasonal (Month) Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
MonthComttd	December	17%	September	12%	September	23%
MonthComttd	April	13%	April	12%	August	12%
MonthComttd	June	9%	November	11%	May	11%
MonthComttd	January	9%	August	10%	March	9%
MonthComttd	February	9%	July	9%	July	8%
MonthComttd	March	8%	February	8%	April	8%
MonthComttd	August	6%	January	8%	February	7%
MonthComttd	November	6%	March	8%	October	7%
MonthComttd	July	6%	December	7%	December	5%
MonthComttd	October	6%	October	6%	January	4%
MonthComttd	September	6%	May	5%	June	4%
MonthComttd	May	5%	June	4%	November	3%

Table 11.0 shows the cluster characteristics of Municipality 3 by Seasonal (Month) 2014 – 2016. The summary shows that *December* ranked the highest in the year 2014 with 17% probability, and *September* in the year 2015 and 2016 with 12% and 23% probability.

Rule Set for Year 2014

Cluster 6: MonthComttd=August, MonthComttd=July, MonthComttd=September,
MonthComttd=May, MonthComttd=October
Cluster 5: MonthComttd=February, MonthComttd=November, MonthComttd=October,
MonthComttd=July, MonthComttd=August, MonthComttd=May
Cluster 1: MonthComttd=June, MonthComttd=January
Cluster 2: MonthComttd=December
Cluster 3: MonthComttd=April, MonthComttd=September
Cluster 4: MonthComttd=March, MonthComttd=May, MonthComttd=September,
MonthComttd=February, MonthComttd=November

Rule Set for Year 2015

Cluster 4: MonthComtd=June, MonthComtd=July, MonthComtd=February,
MonthComtd=December, MonthComtd=October, MonthComtd=January
Cluster 6: MonthComtd=January, MonthComtd=July, MonthComtd=May,
MonthComtd=February, MonthComtd=October
Cluster 2: MonthComtd=August, MonthComtd=March
Cluster 5: MonthComtd=April
Cluster 3: MonthComtd=September
Cluster 1: MonthComtd=November, MonthComtd=December,
MonthComtd=May, MonthComtd=October

Rule Set for Year 2016

Cluster 5: MonthComtd=December, MonthComtd=June, MonthComtd=April,
MonthComtd=February, MonthComtd=October, MonthComtd=July,

MonthComtd=November
Cluster 1: MonthComtd=September
Cluster 2: MonthComtd=March, MonthComtd=August
Cluster 4: MonthComtd=January, MonthComtd=February, MonthComtd=October, MonthComtd=July, MonthComtd=November, MonthComtd=December, MonthComtd=April, MonthComtd=June
Cluster 3: MonthComtd=May, MonthComtd=April, MonthComtd=June, MonthComtd=December

Table 12.0 : Cluster Characteristics of Municipality 4 by Crime Category Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
CrimeCategory	THEFT	35%	THEFT	27%	THEFT	40%
CrimeCategory	PHYSICALINJURIES	31%	ROBBERY	22%	ROBBERY	22%
CrimeCategory	ROBBERY	15%	PHYSICALINJURIES	22%	PHYSICALINJURIES	18%
CrimeCategory	MURDER	13%	MURDER	13%	MOTORNAPPING	10%
CrimeCategory	RAPE	2%	RAPE	8%	MURDER	6%
CrimeCategory	CARNAPPING	2%	MOTORNAPPING	6%	RAPE	2%
CrimeCategory	HOMICIDE	2%	HOMICIDE	2%	HOMICIDE	2%

Table 12.0 shows the cluster characteristics of Municipality 4 by Crime Category 2014 – 2016. The summary shows that *Theft* ranked the highest in the year 2014 - 2016 with 35%, 27%, and 40% probability.

Rule Set for Year 2014

Cluster 2: CrimeCategory=THEFT
Cluster 3: CrimeCategory=ROBBERY, CrimeCategory=MURDER, CrimeCategory=HOMICIDE,CrimeCategory=CARNAPPING, CrimeCategory=RAPE
Cluster 1: CrimeCategory=PHYSICALINJURIES

Rule Set for Year 2015

Cluster 3: CrimeCategory=ROBBERY,CrimeCategory=MOTORNAPPING, CrimeCategory=RAPE,CrimeCategory=HOMICIDE
Cluster 1: CrimeCategory=THEFT
Cluster 2: CrimeCategory=PHYSICALINJURIES
Cluster 5: CrimeCategory=MURDER, CrimeCategory=ROBBERY, CrimeCategory=MOTORNAPPING
Cluster 4: CrimeCategory=RAPE, CrimeCategory=MURDER, CrimeCategory=MOTORNAPPING, CrimeCategory=HOMICIDE

Rule Set for Year 2016

Cluster 1: CrimeCategory=THEFT
Cluster 2: CrimeCategory=ROBBERY
Cluster 4: CrimeCategory=HOMICIDE, CrimeCategory=RAPE, CrimeCategory=MURDER, CrimeCategory=MOTORNAPPING, CrimeCategory=PHYSICALINJURIES
Cluster 3: CrimeCategory=PHYSICALINJURIES, CrimeCategory=MOTORNAPPING, CrimeCategory=MURDER, CrimeCategory=RAPE

Table 13.0 : Cluster Characteristics of Municipality 4 by Barangay Year 2014 – 2016

Population (All)						
Values	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
Barangay	San Agustin	20%	San Agustin	29%	San Agustin	36%
Barangay	Perez	18%	Perez	16%	Perez	19%
Barangay	De Ocampo	11%	Cabuco	10%	Luciano	12%
Barangay	Inocencio	9%	Conchu	9%	De Ocampo	6%
Barangay	Luciano	9%	Osorio	6%	Inocencio	5%
Barangay	Cabuco	8%	Inocencio	6%	Cabuco	5%
Barangay	Lapidario	6%	Luciano	6%	Conchu	4%
Barangay	Aguado	6%	Aguado	5%	Lapidario	3%
Barangay	Cabezas	3%	Lapidario	3%	Osorio	3%
Barangay	Osorio	3%	Cabezas	3%	Gregorio	3%
Barangay	Lallana	3%	De Ocampo	2%	Cabezas	2%
Barangay	Conchu	2%	Gregorio	2%	Aguado	1%
Barangay	Gregorio	1%	Lallana	1%	Lallana	1%

Table 13.0 shows the cluster characteristics of Municipality 4 by Barangay 2014 – 2016. The summary shows that *San Agustin* ranked the highest in the year 2014 - 2016 with 20%, 29%, and 36% probability.

Rule Set for Year 2014

Cluster 3: Barangay=De Ocampo, Barangay=Lapidario, Barangay=Lallana, Barangay=Conchu, Barangay=Osorio

Cluster 4: Barangay=Cabuco, Barangay=Aguado, Barangay=Cabezas, Barangay=Osorio, Barangay=De Ocampo, Barangay=Lallana, Barangay=Conchu

Cluster 2: Barangay=San Agustin

Cluster 5: Barangay=Inocencio, Barangay=Luciano, Barangay=Aguado,Barangay=Cabezas

Cluster 1: Barangay=Perez

Rule Set for Year 2015

Cluster 2: Barangay=Perez, Barangay=Cabuco, Barangay=Osorio, Barangay=Cabezas, Barangay=Lapidario, Barangay=Gregorio, Barangay=De Ocampo, Barangay=Aguado, Barangay=Lallana

Cluster 3: Barangay=Conchu, Barangay=Inocencio, Barangay=Luciano, Barangay=Osorio, Barangay=Cabuco, Barangay=De Ocampo, Barangay=Lallana, Barangay=Cabezas, Barangay=Gregorio, Barangay=Lapidario

Cluster 1: Barangay=San Agustin, Barangay=Aguado

Rule Set for Year 2016

Cluster 1: Barangay=San Agustin

Cluster 2: Barangay=Perez

Cluster 3: Barangay=Luciano, Barangay=De Ocampo, Barangay=Conchu, Barangay=Lapidario, Barangay=Cabezas, Barangay=Aguado, Barangay=Gregorio

Cluster 4: Barangay=Cabuco, Barangay=Gregorio, Barangay=Cabezas, Barangay=Osorio, Barangay=Inocencio, Barangay=Aguado, Barangay=De Ocampo,Barangay=Lallana, Barangay=Lapidario

Cluster 5: Barangay=Inocencio, Barangay=Conchu, Barangay=Lapidario, Barangay=De Ocampo, Barangay=Osorio, Barangay=Lallana, Barangay=Aguado, Barangay=Cabuco, Barangay=Gregorio, Barangay=Cabezas

Table 14.0 : Cluster Characteristics of Municipality 4 by Seasonal (Month) Year 2014 – 2016

Population (All)						
Variables	2014		2015		2016	
	Values	Probability	Values	Probability	Values	Probability
MonthComttd	July	13%	July	16%	July	18%
MonthComttd	December	10%	April	13%	September	12%
MonthComttd	January	10%	October	13%	May	10%
MonthComttd	September	9%	August	13%	August	10%
MonthComttd	May	9%	December	8%	November	9%
MonthComttd	October	8%	June	8%	January	9%
MonthComttd	June	8%	November	8%	December	8%
MonthComttd	April	8%	September	5%	April	7%
MonthComttd	November	7%	January	5%	March	6%
MonthComttd	February	7%	February	4%	October	6%
MonthComttd	August	6%	May	4%	February	5%
MonthComttd	March	6%	March	3%		

Table 14.0 shows the cluster characteristics of Municipality 4 by Seasonal (Month) 2014 – 2016. The summary shows that *July* ranked the highest in the year 2014 - 2016 with 13%, 16%, and 18% probability.

Rule Set for Year 2014

Cluster 1: MonthComttd=July, MonthComttd=January
Cluster 5: MonthComttd=February, MonthComttd=March, MonthComttd=June, MonthComttd=October, MonthComttd=November, MonthComttd=December
Cluster 6: MonthComttd=August, MonthComttd=October, MonthComttd=November, MonthComttd=February, MonthComttd=December, MonthComttd=March
Cluster 2: MonthComttd=May, MonthComttd=April
Cluster 4: MonthComttd=December, MonthComttd=March, MonthComttd=October, MonthComttd=November, MonthComttd=June
Cluster 3: MonthComttd=September, MonthComttd=June, MonthComttd=December

Rule Set for Year 2015

Cluster 7: MonthComttd=January, MonthComttd=November, MonthComttd=May, MonthComttd=March, MonthComttd=September,
Cluster 1: MonthComttd=July
Cluster 5: MonthComttd=December, MonthComttd=September, MonthComttd=June, MonthComttd=May, MonthComttd=March, MonthComttd=February, MonthComttd=November
Cluster 6: MonthComttd=June, MonthComttd=December, MonthComttd=February, MonthComttd=January,MonthComttd=September, MonthComttd=November, MonthComttd=March
Cluster 4: MonthComttd=April
Cluster 2: MonthComttd=October
Cluster 3: MonthComttd=August, MonthComttd=November

Rule Set for Year 2016

Cluster 3: MonthComttd=November, MonthComttd=December, MonthComttd=April, MonthComttd=February, MonthComttd=March
Cluster4: MonthComttd=October, MonthComttd=January, MonthComttd=May, MonthComttd=April, MonthComttd=December
Cluster 1: MonthComttd=September, MonthComttd=August, MonthComttd=January
Cluster 2: MonthComttd=July, MonthComttd=February, MonthComttd=March

4.1 APPLICATION SOFTWARE FOR CRIME MANAGEMENT SYSTEM (CRIMS)

The clustering algorithm has been implemented in the application software for Crime Management System (CriMS) for Philippine National Police – District VI in the province of Cavite. The following sample screen shots of the system are shown below.

Figure 2.0 Login Page

Figure 2.0 shows the Login Page of the system. The user is asked to input the username and password in able to access the system.

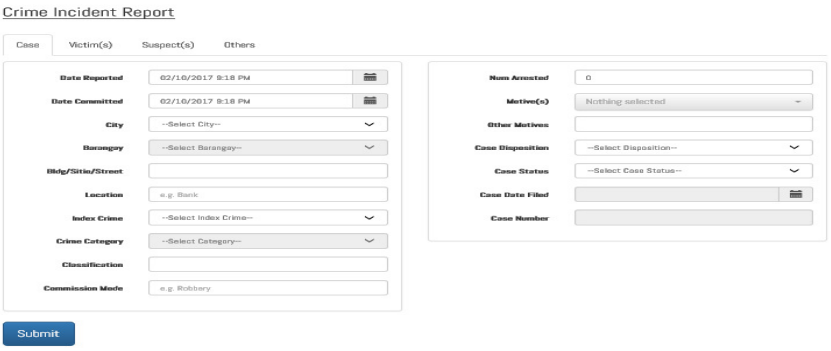

Figure 3.0 Data Entry

Figure 3.0 shows the data entry of the system. The user inputs the detailed information on the crime committed such as the date, time, city, barangay, crime category, and etc.

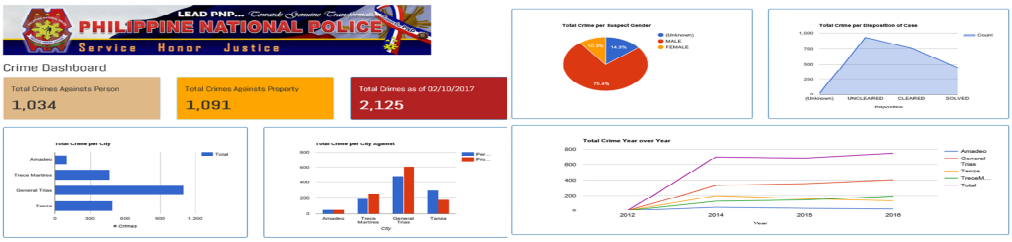

Figure 4.0 Dashboard

Figure 4.0 shows the dashboard of the system which presents the summary of the crime committed that is total crime per city, total crime per city against person and against property, total crime per suspect gender, total crime per disposition of the case and total crime over the year.

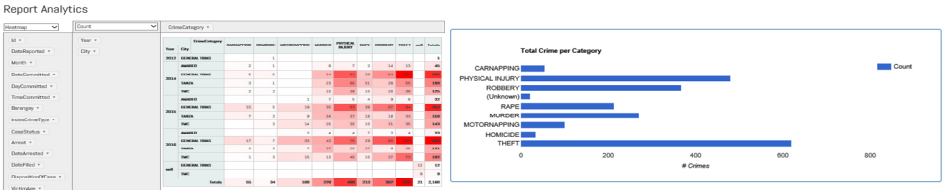

Figure 5.0 Report Analytics

Figure 5.0 shows the report analytics. This module presents the totality of the crime committed per year, city, crime category, city and barangay, and etc. The data can also be viewed in graphs.

4.2 . STATISTICAL TREATMENT

The study adapted the ISO 25010:2011 model for system and software engineering quality requirements and evaluation. The result is interpreted using weighted mean.

$$\bar{x} = \frac{\sum xw}{\sum w}$$

where $\bar{x}$ = mean

x = measurement or value

w = number of measurements

Eq. (5)

Table 15.0 : Software Evaluation (ISO 25010:2011 Quality Model)

No	Criterion		Mean	
	Characteristics	Subcharacteristics		
1	Functional Suitability			4.4
		Functional Completeness	4.4	
		Functional Correctness	4.4	
		Functional Appropriateness	4.4	
2	Performance Efficiency			4.4
		Time Behaviour	4.4	
		Resource Utilization	4.5	
		Capacity	4.4	
3	Compatibility			4.4
		Co-existence	4.3	
		Interoperabilty	4.5	
		Interoperabilty	4.3	
4	Usability			4.5
		Learnability	4.5	
		User Interface Aesthetics	4.3	
		Operability	4.6	
5	Reliability			4.4
		Maturity	4.5	
		Availabilty	4.3	
		Fault Tolerance	4.3	
6	Security			4.5
		Confidentiality	4.4	
		Integrity	4.5	
		Accountability	4.7	
7	Maintainability			4.4
		Modularity	4.3	
		Modifiability	4.5	
		Testability	4.4	
8	Portability			4.4
		Adaptabilitty	4.3	
		Adaptabilitty	4.5	
		Instability	4.4	
			Average Mean	4.4

The computed weighted mean for each functional suitability criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very functional, highly functional, functional, poorly functional and not functional at all. As presented in the table, the computed functional suitability level in all criterions got a mean of 4.4 which can be interpreted that the functional suitability of the system is highly functional. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall functional suitability level of highly functional as evaluated by the respondents of the study.

The computed weighted mean for each performance efficiency criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very efficient, highly efficient, efficient, poorly efficient and not efficient at all. As presented in the table, the computed performance efficiency level in all criterion ranges from 4.4 to 4.5 which can be interpreted that the performance efficiency of the system is highly efficient. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall performance efficiency level of highly efficient as evaluated by the respondents of the study.

The computed weighted mean for each compatibility criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very compatible, highly compatible, compatible, poorly compatible and not compatible at all. As presented in the table, the computed compatibility level in all criterion ranges from 4.3 to 4.5 which can be interpreted that the compatibility of the system is highly compatible. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall compatibility level of highly compatible as evaluated by the respondents of the study.

The computed weighted mean for each usability criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very usable, highly usable, usable, poorly usable and not usable at all. As presented in the table, the computed usability level in all criterion ranges from 4.3 to 4.5 which can be interpreted that the usability of the system is highly usable. The overall computed weighted mean for all criterions is 4.5 which can be interpreted an overall usability level of very usable as evaluated by the respondents of the study.

The computed weighted mean for each reliability criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very reliable, highly reliable, reliable, poorly reliable and not reliable at all. As presented in the table, the computed reliability level in all criterion ranges from 4.3 to 4.5 which can be interpreted that the reliability of the system is highly reliable. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall reliability level of highly reliable as evaluated by the respondents of the study.

The computed weighted mean for each security criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very secured, highly secured, secured, poorly secured and not secure at all. As presented in the table, the computed security level in all criterion ranges from 4.4 to 4.7 which can be interpreted that the security of the system is highly secured The overall computed weighted mean for all criterions is 4.5 which can be interpreted an overall security level of very secured as evaluated by the respondents of the study.

The computed weighted mean for each maintainability criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very maintainable, highly maintainable, maintainable, poorly maintainable and not maintainable at all. As presented in the table, the computed maintainability level in all criterion ranges from 4.3 to 4.5 which can be interpreted that the maintainability of the system is highly maintainable. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall maintainability level of highly maintainable as evaluated by the respondents of the study.

The computed weighted mean for each portability criteria as shown in the table 15.0 have been interpreted using 5 point Likert scale as follows: very portable, highly portable, portable, poorly portable and not portable at all. As presented in the table, the computed portability level in all

criterion ranges from 4.3 to 4.5 which can be interpreted that the portability of the system is highly maintainable. The overall computed weighted mean for all criterions is 4.4 which can be interpreted an overall portability level of highly portable as evaluated by the respondents of the study.

The Table 15.0 presented the average computed weighted mean for all criterions is 4.4 which can be interpreted that the system has an overall rating of high quality.

5. CONCLUSION

The study aspire to measure the performance efficiency of the Philippine National Police District VI in the Province of Cavite using Data Envelopment Analysis and conduct data mining using a clustering algorithm to identify the crime pattern such as hotspots, hot place, seasonal and frequency. The results of the performance efficiency were obtained by Municipality 3 as efficient Decision-Making Unit (DMU). Therefore, inefficient DMUs can benchmark the best practices of the efficient DMU in its strategic, tactical and operational undertakings. The Clustering Algorithm was implemented in the web-based software application *Crime Management System* (CriMS) developed by the researchers. The researchers recommend that the software application must be applied in the Philippine National Police District VI in the Province of Cavite. Auxiliary studies must be conducted to determine the usefulness and effectiveness of the application software, hence institutionalized its implementation. Supplementary studies should be conducted by using the results of this study's data set for a detailed numerical benchmarking, empirical analysis and develop a new algorithm which are functional solution mitigation initiatives to crime prevention and other methods.

REFERENCES

[1] Mojares, J. G. (2013). Urbanization and its Effect in the Calabarzon Area, Philippines. Journal of Global Intelligence & Policy,6(10)

[2] Guanghua Wan & Iva Sebastian (2011). Poverty in Asia and the Pacific: An Update. Asian Development Bank Economics Working Paper No 267

[3] Boussofiane, A., Dyson, R. G., & Thanassoulis, E. (1991). Applied data envelopment analysis. European Journal of Operational Research, 52(1), 1-15

[4] Norman, M., & Stoker, B. (1991). Data envelopment analysis:the assessment of performance. John Wiley & Sons, Inc.

[5] Lertworasirikul, S., Fang, S. C., Joines, J. A., & Nuttle, H. L. (2003). Fuzzy data envelopment analysis (DEA): a possibility approach. Fuzzy sets and systems, 139(2), 379-394

[6] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European journal of operational research, 2(6), 429-444

[7] Berry, M. J., & Linoff, G. (1997). Data mining techniques: for marketing, sales, and customer support. John Wiley & Sons, Inc.

[8] Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.

[9] Agrawal, R., & Srikant, R.(2000, May). Privacy-preserving data mining. In ACM Sigmod Record (Vol. 29, No. 2, pp. 439-450). ACM

[10] L Rokach, O Maimon (2010). Data mining and knowledge discovery handbook. Springer US.

[11] Sankar Rajagopal (2011). Customer Data Clustering using data mining technique. International Journal of Database Management Systems (IJDMS) Vol.3, No.4, November 2011

[12] Chan, S., & Karim, M. A. (2016). Financial market regulation, country governance, and bank efficiency: evidence from East Asian countries. Contemporary Economics, (1), 39. doi:10.5709/ce.1897-9254.197

[13] Kinachi, H., Najjari, V., & Alp, İ. (2016). Using Data Envelopment Analysis and Stochastic Frontier Analysis Methods to Evaluate Efficiency of Hydroelectricity Centers. Gazi University Journal of Science, 29(1), 167-176

[14] Osman, I. H., Berbary, L. N., Sidani, Y., Al-Ayoubi, B., & Emrouznejad, A. (2011). Data Envelopment Analysis Model for the Appraisal and Relative Performance Evaluation of Nurses at an Intensive Care Unit. Journal Of Medical Systems, (5), 1039

[15] Ulucan, A. (2011). Measuring the Efficiency of Turkish Universities Using Measure-Specific Data Envelopment Analysis. Sosyoekonomi, 14(1), 181

[16] Estrada S., Song H., Kim Y., Namn S., Kang S. (2009). A Method of Stepwise Benchmarking for Inefficient DMUs based on the Proximity-based Target Selection. Expert Systems with Applications 36 (2009) 11595-11604

[17] Grubesic, T. H., Wei, R., & Murray, A. T. (2014). Spatial Clustering Overview and Comparison: Accuracy, Sensitivity, and Computational Expense. Annals of the Association of American Geographers, 104(6), 1134-1156

[18] Jin-ho, J., & Seung-Ryul, J.(2016). Designing a Crime-Prevention System by Converging Big Data and IoT. Journal of Korean Society for Internet Information, 17(3),115-128.doi: 10.7472/jksii.2016.17.3.115

[19] Ceccato, V., & Uittenbogaard, A. C. (2014). Space–Time Dynamics of Crime in Transport Nodes. Annals of the Association of American Geographers, 104(1), 131-150

[20] Mohammad A.Tayebi, Martin Ester, Uwe Glasser, Patricia L.Brantingham (2014). CRIMETRACER: Activity space based crime location prediction", ASONAM, 2014, 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 2014, pp. 472-480, doi:10.1109/ASONAM.2014.6921628

[21] Sankar Rajagopal (2011).Customer Data Clustering using data mining technique. International Journal of Database Management Systems (IJDMS) Vol.3, No.4, November 2011

[22] Phillips, P., & Lee, I. (2011). Crime analysis through spatial areal aggregated density patterns. Geoinformatica, 15(1), 49-74. doi:10.1007/s10707-010-0116-1

[23] Semwal, V.B., Singha, J., Sharma, P. et al., Multimedia Tools Appl (2016). doi: 10.1007/s11042-016-4110-y

[24] Devi & Rajagopalan S.P. (2012). Anaylsis of Customer Behavior using Clustering and Assocation Rule. International Journal of Computer Applications (IJCA) Vol.43, No.23, April 2012

[25] Seddawy, A. E., Khedr, A. & Sultan, T. (2012). Enhanced K-mean Algorithm to Improve Decision Support System under Uncertain Situation. International Journal of Modern Engineering Research (IJMER) Vol.2, No.3, Aug 2012

AUTHORS

MS. MENGVI P. GATPANDAN is on her Dissertation of Doctor in Information Technology at AMA Computer University. She finished her Bachelor of Science in Computer Science in 1996 at AMA Computer University, Makati City where she also received the degree of Master of Science in Computer Science in 2003. She is a full-time faculty of Jose Rizal University for 11 years, and also an active member of Philippine Society of Information Technology Educators. She is in the academe in teaching profession for 18 years to present. With a stint in database management, she is an IBM DB2 and Cognos BI Certified and also holder of NC II Tesda Certification Program.

DR. SHANETH C. AMBAT received her Doctor of Philosophy in Engineering major in Information Technology in Hannam University, South Korea in 2009. She received her Bachelor of Science in Computer Science in 1995 and Master of Science in Computer Science in 2004 respectively at AMA Computer University, Makati City. She is currently the Dean of the School of Graduate Studies at AMA Computer University, Quezon City. Her research interest includes data envelopment analysis, data mining, SOM, reinforcement learning algorithm, and fuzzy logic.