Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Determining Spatial Distribution of the BSP Supervised Banks in the Philippines: A Multivariate Cluster Analysis

Vivian Sesgundo[1*], Peter John Aranas[2]

[1,2]Graduate School, Polytechnic University of the Philippines.*

Email: [2]peterjohn5416@gmail.com
Corresponding Email: [1]bebs.sesgundo@gmail.com*

Received: 29 November 2023 **Accepted:** 16 February 2024 **Published:** 01 April 2024

Abstract: ***The goal of this study was to determine the significant variables that contribute to the number of Bangko Sentral ng Pilipinas (BSP) Supervised banks in the Philippine regions. A model that can predict the number of banks needed in the region was also presented. Spatial distribution of the banks was also analyzed. These analyses determined if the current number of banks in the region is sufficient to provide the financial services needed by the people. The ArcGIS Pro was used to perform Ordinary Least Square Regression, Global Moran's I and Multivariate Clustering Analyses to the Regional Distribution of BSP Supervised banks in the Philippines and the categorized economic, demographic, labor market and potential market variables from the Philippine Statistics Authority (PSA) in 2020. Results of this study show that the population density, economically active population, functional literacy rate, and families' ownership of personal computer are the significant factors, which represent every category of the independent variable, that contributed to the number of banks in the region. Using the chosen passing model, the difference between the actual and estimated number of banks was used as the indicator by which regions need more physical banks to promote financial inclusion. Regions X, CAR, and BARMM are the best locations for the future physical bank to be established. Furthermore, the results from Global Moran's I analysis showed that there is a clustering of high number of banks in the Philippines. There were 7 clusters formed for the number of banks based on population density, economically active population, functional literacy rate, and families' ownership of personal computer. Among these explanatory variables, population density is the greatest contributor in forming the clusters.***

Keywords: ***Financial Institutions (FIs), Spatial Distribution, Ordinary Least Squares (OLS), Cluster Analysis, Applied Statistics.***

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

1. INTRODUCTION

The Philippines has a total of 29, 275 financial institutions (FIs) as of July 2022 (Bangko Sentral ng Pilipinas). This consists of 13,191 banks, 15,994 non-banks, 89 insurance companies, and 1 offshore banking unit. The researcher wanted to determine if the number of available FIs in each region is enough to serve the financial services needed by the residents in the area. This research covers only BSP-supervised banks in the Philippines, specifically head office and branches with distribution per region from the Geographical Distribution of Banking Offices Philippine Banking System reported by the BSP as of 2020.

There were programs and policies relative to financial inclusion which the BSP initiatively launched to increase Filipinos' access to financial services. Moreover, to increase the financial literacy of different target audiences, various activities were conducted by BSP. To build an inclusive financial system was the overall vision where a wide range of appropriate financial products and services were effectively accessible. The BSP believes that it is a shared responsibility among financial consumers, financial institutions, and the BSP, to promote consumer protection and financial education (BSP FIS Topline Report, 2019). This research determines if there is a region which needs more banks to promote financial inclusion.

According to the BSP's 2019 Financial Inclusion Survey, account ownership increased from 22.6% in 2017 to 28.6% in 2019. The number of adult Filipinos with accounts climbed from 15.8 million in 2017 to 20.9 million in 2019, or 5.1 million adults who entered the financial system during that time. In terms of demographic groups, North/Central Luzon and the Visayas, people in socioeconomic class E who are 40–49 years old, less educated (elementary graduates), working adults, rural areas, and men are the top gainers. According to data showing saving as the primary use case for an account and savers more likely to hold an account, a rise in account ownership also corresponds with an increase in the number of persons who save (BSP, 2019). This research shows if the mentioned demographic characteristics are factors to the number of banks available in the Region using demographic data from the 2020 Census conducted by PSA and other latest available variables.

2. RELATED WORKS

The physical distribution of stores and offices is greatly influenced by digitalization of processes, services, and changes in customers' needs [8]. A bank's branch location is important to be strategically selected. More profitable customers and an increase in probable sales can be guaranteed with a good branch location. Transactions are performed online and through mobile channels which leaves the strengthening of the relationship with customers as a primary role of branches. Branches have evolved from being a transaction channel to a physical point of communication which can have a great effect on sales. People still like to see the physical presence of the bank's branches located nearby for available convenient access. When having trouble, the presence of a branch acts as a contingency device, when people opt to talk to someone. Branches should be treated as interconnected outlets through which customers engage and develop a relationship with the bank [7]. Since a poor site choice would be

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

expensive and difficult to undo, choosing the right location is seen as a crucial decision. Therefore, significant discrepancies in the availability of financial services or abrupt changes or disparities in the operating conditions to which bank branches are exposed serve as the primary drivers of study on the spatial distribution of bank branches [6].

Bank branch location approaches are formulated as linear programming problems, problems of multi-criteria analysis, problems of neural networks, or problems of data envelopment analysis (Boufounou, 1995; Cornu'ejols et al., 1983; Hopmans, 1986) [9]. Unit operating expenses, the level of competition in the prospective location, and the region's accessibility were listed by Min as the three main considerations when determining where to put bank branches. W'ojcik et. al. [2017] found a connection between socioeconomic factors and the placement of bank branches. Bank branches are distributed differently in terms of geography as well. Urban locations tend to have a concentration of bank branches.

A GIS-based spatial interaction model was given to address the bank-branch problem, and it was used in New Zealand (Morrison and O'Brien, 2001). A study conducted by Ansong, D., Chowa, G., and Adjabeng, B. used spatial analyst tools, geographically weighted Poisson regression, and Ghana's banking sector data to show the inequality in availability of branch-based services, highlighting the differences of branch allocation determinants between district and regional level. The study shows that population size, percentage of urban residents, workforce size, and literacy level are associated with bank allocation varying by district [1]. According to a study, banks should choose branch locations based on the geographic regions' demographic, economic, and social aspects in order to offer the finest financial services in line with client demand. Using panel regression models and the tests for normality of the error term, heteroscedasticity, multicollinearity, and model specification, Gorfu and Mamo looked at the influence of economic, social-cultural, and demographic variables on the decision of where to locate a bank branch over time. They discovered the importance of government capital expenditure, the number of employed people, the number of businesses in all sectors, the number of vocational training schools, and the number of undergraduate students [2].

3. METHODOLOGY

To determine the significant factors contributing to the number of banks in a region and to have a statistical model that can predict the number of banks needed in the area, the Ordinary Least Squares (OLS) Linear regression was applied to quantitative variables. Spatial distribution of BSP-supervised Banks in the Philippines was determined using the Multivariate Cluster Analysis and Spatial autocorrelation using Moran's I statistics. Secondary data from the banking directory by the BSP and demographic profile by the PSA were used for this study. To show the characteristics of the Region in the Philippines, the following variables were presented given the data from the BSP and PSA: a. Population, b. Population Growth Rate, c. Population Density, d. Economically-Active Population Rate, e. No. of Establishments in Operation, f. No. of Establishments in Education Sector, g. No. of Transportation and Storage Establishments, h. Certified Alienable and Disposable Land, i. Average Family Income, j. Average Family Savings, k. Unemployment rate, l. Higher Education graduates, m. Functional

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Literacy rate, n. Share of Families Ownership-Personal computer, o. Share of Families Ownership-Own Cellular Phone, and p. Proportion of Population 10 To 64 Years Old Using the Internet for Email/Research. To determine if there is significant relationship between each of the characteristics of the region and the number of banks in the area, Ordinary Least Squares (OLS) Linear Regression was utilized. After finding out the significant factors in the number of banks in the region, these significant factors were used to determine the model that could predict the number of banks needed by the region.

To define the spatial distribution of the BSP Supervised Banks in the Philippines, a multivariate cluster analysis- Spatial autocorrelation using Moran's I statistics was used. After the spatial distribution was presented, Region that needs more physical banks to promote were also determined. Furthermore, the best locations for future banks were suggested. Because it is multidirectional and multidimensional, spatial autocorrelation is beneficial for identifying patterns in challenging data sets. Its value ranges from -1 to 1, and it resembles correlation coefficients. Moran's coefficient, on the other hand, measures perfect correlation to no correlation in a slightly different way (as a result of the more intricate, spatial calculations). Moran's I index is an inferential statistic that you have to determine statistical significance before you can read the result. This is done with a simple hypothesis test, calculating a z-score and its associated p-value.

Some of the most popular machine learning approaches are clustering, grouping, and classification procedures. To identify natural clusters in your data, the Multivariate Clustering tool uses unsupervised machine learning techniques. These classification techniques are regarded as unsupervised since they can locate clusters in your data without the use of a collection of preclassified characteristics [4].

4. RESULTS AND DISCUSSION

4.1 Characteristics of the Regions in the Philippines

The results tell us that BARMM and NCR are the extreme values in Number of Banks, Number of establishments in operation, Number of establishments in education sector, Number of transportation and storage establishments, Average family Income, Higher education graduates, Functional literacy rate, share of families' ownership-personal computer and Proportion of population 10 to 64 years old using the internet for email/research. CAR is the region, which is lowest in terms of population, population density, and economically active population but highest in Average family savings. Region IV-A has the highest population, and economically active population rate. The region with the lowest population growth rate and share of families' ownership-cellular phone is Region VIII. Region VI has the highest certified alienable and disposable land while NCR has the lowest. BARMM has the highest unemployment rate while Region X has the lowest. (Refer to Table 1)

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Table I Summary of Highest and Lowest Values for Each Variable

Variable	Highest	Lowest
Number of Banks	NCR	BARMM
Population	IV-A	CAR
Population Growth Rate	BARMM	VIII
Population Density	NCR	CAR
Economically -Active Population Rate	IV-A	CAR
No. of Establishments in Operation	NCR	BARMM
No. of Establishments in Education Sector	NCR	BARMM
No. of Transportation and Storage Establishments	NCR	BARMM
Certified Alienable and Disposable Land	VI	NCR
Average Family Income	NCR	BARMM
Average Family Savings	CAR	BARMM
Unemployment rate	BARMM	X
Higher Education graduates	NCR	BARMM
Functional Literacy rate	NCR	BARMM
Share of Families Ownership-Personal computer	NCR	BARMM
Share of Families Ownership-Own Cellular Phone	NCR	VIII
Proportion of Population 10 To 64 Years Old Using the Internet for Email/Research	NCR	BARMM

1. Relationship between the Characteristics of the Regions and the Number of Banks

For this study, the researcher wanted to find out if the variables mentioned above are significantly related to the dependent variable, number of BSP supervised banks in the Philippines. To determine the significant factors contributing to the number of banks in a region and to have a statistical model that can predict the number of banks needed in the area, OLS Linear Regression was applied to quantitative variables. As one of the assumptions of OLS is that the independent variables should be significantly related to the dependent variable, the researcher utilized the correlation statistical method through Microsoft Excel.

The correlation matrix shows variables have a positive moderate to strong correlation with the number of banks in the region except the following four (4) which have weak negative/positive correlation: Population Growth Rate, Certified Alienable and Disposable Land, Average Savings of Families at Current Prices (Thousands), and Unemployment rate. We can also notice the variables which are negatively correlated with each other. This means that as certified alienable and disposable land increases, population density and number of

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

transportation and storage establishments decreases. As well as when average savings of families increases, certified alienable and disposable land decreases. When functional literacy rate increases, the population growth rate and unemployment rate decreases. By sorting the correlation coefficients from highest to lowest, we can see the variables with the strongest correlation with the number of banks in the region. The number of higher education graduates, number of establishments in operation, number of transportation and storage establishments, number of establishments in education sector, and population density are the top five variables that has strongest correlation with the number of banks in the region. However, certified alienable and disposable land, unemployment rate, population growth rate, and average saving of families have the weakest correlation with the number of banks in the region. The results shows that most of the variables have significant relationship with the number of banks in the region except the following: Population growth rate, Certified Alienable and Disposable Land, Average Savings of Families At Current Prices (Thousands), and unemployment rate.

Prior to determining the best model needed for this study, we must check first if the assumptions for OLS are met. Assumptions: a) The model must be linear. The scatter plots of each independent variable in x-axis and the number of BSP supervised banks as dependent variable in y-axis are linear. It is noticeable that the following has a negative relationship with the number of BSP Supervised banks in the region: population growth rate, certified alienable and disposable land, and unemployment rate. This means that as these variables decrease, the number of banks in the region increases. The rest of the variables have a positive correlation with the number of banks. Thus, the first assumption is met. b) The data must be randomly sampled. Our data demographic data is from the 2020 Philippine Census, some variables are from actual numbers while other variables are product of regular surveys of the PSA that used random samples. Thus, we can say that this second assumption is met. c) The explanatory variables must not be collinear. The researcher performed the multicollinearity test on Microsoft Excel, and the result of multicollinearity shows that there is a multicollinearity among the following variables: Population, Economically-active Population, Number of Establishments in Operation, and Number of Establishments in Education Sector; Population Density has strong collinearity with Number of Transportation and Storage Establishments; Higher Education Graduates has a collinearity with Number of Establishments in Operation and in Education Sector. The researcher considered the multicollinearity results in determining which among the correlated explanatory variables should be chosen as the representative of the variables. d) The residuals have an expected sum of zero. By plotting the distribution of the residuals generated after the individual OLS Regression analysis, we can see if the mean is zero. Based on the results from OLS Regression Analysis of each independent variable with the number of banks in the region, histogram of the residuals showed that means are zero (0). Thus, we can say that this assumption was met. e) The residuals have homogeneous variance. If the Koenker Statistic is significant, the variances of the residuals are not homogenous. This means that the model is inconsistent. Based on the results, if individual variables will form a model for the number of banks, the following variables have homogenous variances of residuals: Population Growth rate, Population density, Number of transportation and storage establishments, Average savings of families, unemployment rate, Functional literacy rate, and Proportion of population 10 to 64 years old using the internet for email/research. f) The

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

residuals are normally distributed. From the results of the individual OLS regression analysis, Jarque-Bera Statistics determined the variables which have normally distributed residuals, namely: Number of Establishments in Operation, Higher-Education-Graduates AY-2018-2019, Families Ownership - Personal Computer, Families Ownership - Own Cellular Phone, And Proportion of Population 10 to 64 Years Old Using the Internet for Email/Research. We will consider these results in determining the best model for the number of banks in the region. g) Adjacent residuals must not show autocorrelation. After running OLS in ArcGIS for every relationship of the number of banks in the region with the independent variables, the researcher executed the Global Moran's I Analysis for the residuals to see if there is autocorrelation. The Global Moran's I spatial autocorrelation report shows that the following variables do not show autocorrelation of residuals: Population, Population Density, Economically-active population, Number of Establishments in Operation, Number of Establishments in Education Sector, Average Income of Families, Average Savings of Families, Functional Literacy rate, Families ownership-Personal Computer, Families Ownership-Cellular phone, and Proportion Of Population 10 to 64 Years Old Using the Internet for Email/Research.

2. Best Model to Determine the Needed Number of Banks in the Region

The researcher utilized the Exploratory Regression Analysis tool of ArcGIS to determine the best model for this study. Considering the individual relationship of the independent variables with the number of banks in the region, the following are the results of the Exploratory Regression Analysis. Please see Appendix A for the list of passing models per indicated number of variables to include from 1 to 16. Arranging the passing models in descending order of AICc to see the best fit model that can explain the number of banks in the region, for this study, the researcher decided to look for a model that has a representative variable from each of the four categories of explanatory variables. A model with the highest adjusted R-squared and lowest AICc consists of variables A+C+M+N. It explains 99.62% of the number of banks in the region. However, it does not represent the four categories of the explanatory variables. The second model, with variables C+D+M+N, represents the four categories of explanatory variables. This model can explain 99.59% of the number of banks in the region. Thus, for this study, the researcher used population density, economically-active population, functional literacy rate, and families' ownership-personal computer to determine the number of banks needed in the region. Below are the results of OLS Regression Analysis for the selected model.

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Table II Results of OLS Regression Analysis for the Selected Model

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	-1260.033604	261.515115	-4.818206	0.000421*	232.089551	-5.429084	0.000151*	-------
C	0.101673	0.003854	26.379156	0.000000*	0.001510	67.332777	0.000000*	1.892390
D	0.000159	0.000007	24.016860	0.000000*	0.000004	35.735546	0.000000*	1.524656
M	11.286635	3.419686	3.300489	0.006340*	2.935781	3.844509	0.002339*	2.154130
N	9.699732	3.702229	2.619971	0.022384*	3.109518	3.119368	0.008868*	3.058293

OLS Diagnostics

Input Features	Regions_ExportFeatures6	Dependent Variable	BANKS
Number of Observations	17	Akaike's Information Criterion (AICc)['d']	201.001174
Multiple R-Squared['d']	0.996956	Adjusted R-Squared['d']	0.995941
Joint F-Statistic['e']	982.526722	Prob(>F), (4,12) degrees of freedom	0.000000*
Joint Wald Statistic['e']	69804.815813	Prob(>chi-squared), (4) degrees of freedom	0.000000*
Koenker (BP) Statistic['f']	0.852130	Prob(>chi-squared), (4) degrees of freedom	0.931326
Jarque-Bera Statistic['g']	1.743687	Prob(>chi-squared), (2) degrees of freedom	0.418180

Therefore, the equation to estimate the number of banks in the Region is: Number of Banks = -1260.033604+(0.101673*C) +(0.000159*D) +(11.286635*M) +(9.699732*N) where: C = Population Density, D = Economically-Active Population, M = Functional Literacy Rate, and N = Families Ownership - Personal Computer

3. Characteristic of the Spatial Distribution of BSP Supervised Banks in the Philippines

The actual and estimated number of banks maps show that the highest number of banks is in NCR followed by its neighbors. The least number of banks is in Mindanao Area and CAR. MIMAROPA is the only Region that changes color from actual to estimated map. Note that the research used the model Number of Banks = -1260.033604+(0.101673*C) +(0.000159*D) +(11.286635*M) +(9.699732*N) for this analysis.

Below are the results of Moran's I Analysis. It shows that both the actual and estimated number of banks in the regions are clustered.

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Figure 1 Spatial Autocorrelation Report for Actual Number of Banks

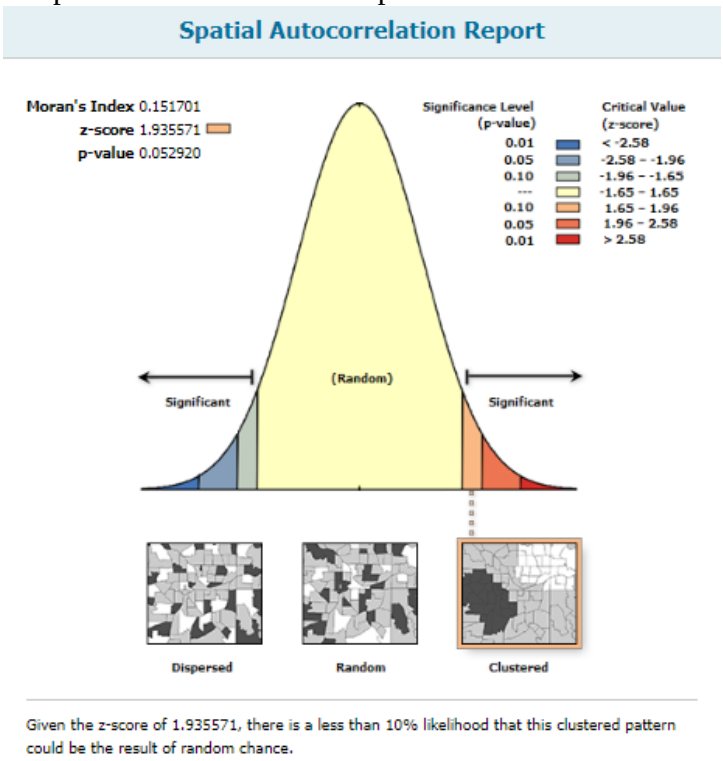
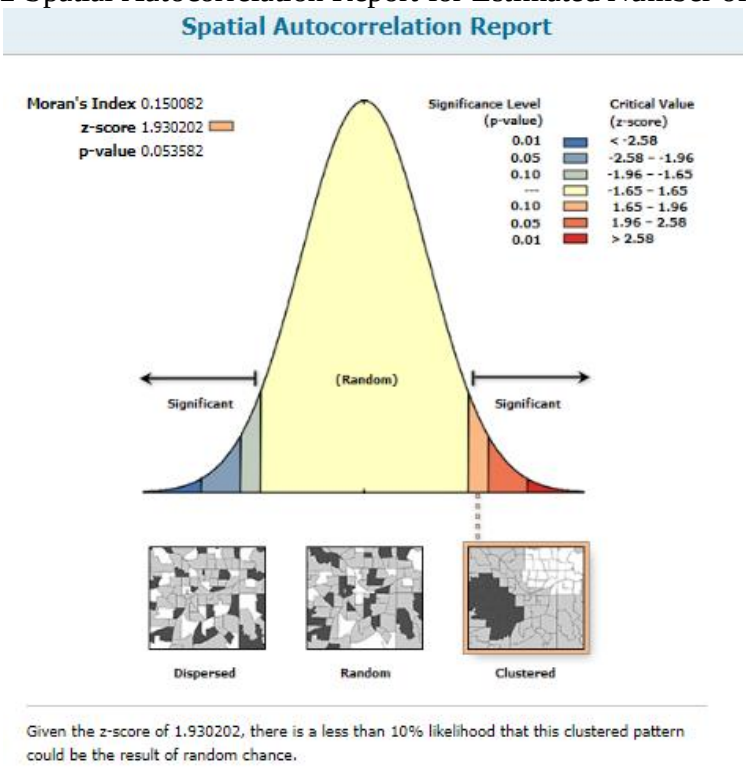


Figure 2 Spatial Autocorrelation Report for Estimated Number of Banks



Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

It was established that the distribution of banks in the Philippines is clustered. Using the chosen model for this study, results of the multivariate clustering tool of ArcGIS show that cluster of seven (7) has the highest statistics.

Table III Variable-Wise Summary of the Selected Model

Variable-Wise Summary

Variable	Mean	Std. Dev.	Min	Max	R2
C	1568.498824	5053.809114	90.710000	21765.280000	0.999579
BANKS	756.588235	889.040399	18.000000	3835.000000	0.976320
N	19.704118	6.688574	7.080000	35.660000	0.945899
D	4082197.000000	2632795.762240	1178144.000000	10673858.000000	0.928796
M	89.924118	6.077249	71.560000	96.530000	0.926567

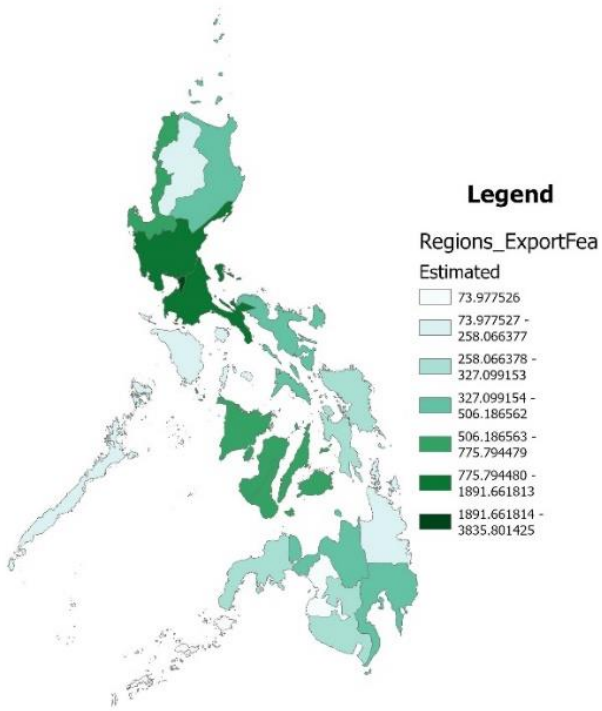
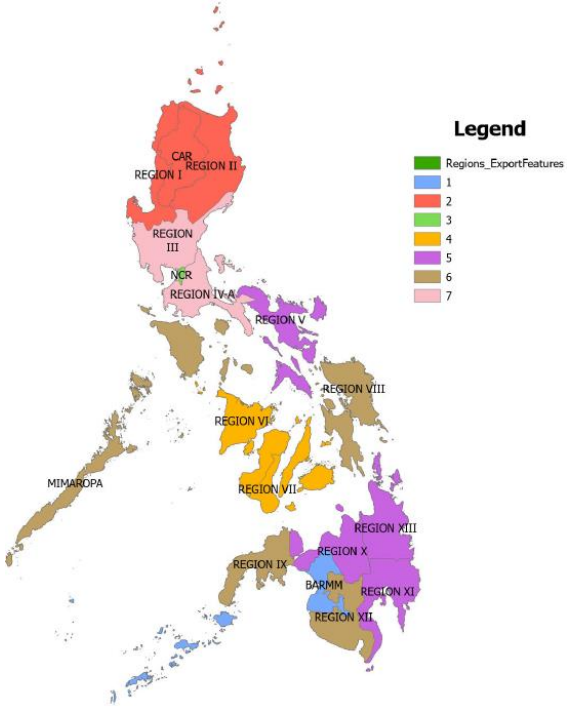


Figure 3 Map of Estimated Number of Banks

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Figure 4 Clustering of Banks Using the Selected Model



Cluster six (6) and five (5) have the highest number of regions composed of 4 regions each, MIMAROPA, VIII, IX, and XII and V, X, XI, and XIII, respectively. Moreover, clusters one (1), and three (3) have only one Region each, NCR and BARMM. Cluster seven (7) consists of Regions III and IV-A. Clusters four (4) has two (2) Regions (VI and VII). Lastly, cluster 2 is composed of Regions I, CAR and II. Table above shows that variable C, which is the Population Density, is the highest contributor in determining distinct clusters.

Boxplot presents the characteristics of the clusters formed. Cluster 3 consisting of NCR, has the highest number of banks, maximum population density, functional literacy rate and families' ownership of personal computer. All values relative to NCR are above averages. At the same time, values of cluster 7 is above average. This cluster has an average of 1,611 banks, 772 persons per sq.km. of land, 9.38 million economically-active population, 94.94% functional literacy rate and 25.32% families' ownership of personal computer. Moreover, cluster 1 has the lowest average number of banks. This cluster has an average of 120 persons per sq.km., 2.79 million economically-active population, (71.56%) lowest functional literacy rate and only 7.08% families' ownership of personal computer.

Since the population density is the greatest contributor in defining the clusters for the number of banks, the table below from PSA shows the distribution of Population Density per region highlighted by corresponding color of the cluster. We can see that descending order of population density per region could be grouped as identified clusters of the number of banks.

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Table IV Distribution of Population Density per Region

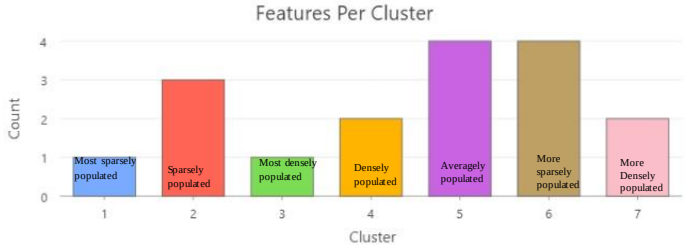
Table 3. Total Population, Land Area, and Population Density by Region: 2020

Rank	Region	Total Population	Land Area (Square Kilometer)	Population Density (Persons Per Square Kilometer of land)
	Philippines	109,033,245 [a]	300,000.00 [b]	363
1	National Capital Region (NCR)	13,484,462	619.54	21,765
2	IV-A - CALABARZON	16,195,042	16,576.26	977
3	III - Central Luzon	12,422,172	21,906.19	567
4	VII - Central Visayas	8,081,988	15,872.58	509
5	I - Ilocos	5,301,139	12,964.62	409
6	VI - Western Visayas	7,954,723	20,778.29	383
7	V - Bicol	6,082,165	18,114.47	336
8	XI - Davao	5,243,536	20,433.38	257
9	X - Northern Mindanao	5,022,768	20,458.51	246
10	IX - Zamboanga Peninsula	3,875,576	16,904.03	229
11	XII - SOCCSKSARGEN	4,901,486	22,786.08	215
12	II - Cagayan Valley	3,685,744	29,836.88	124
13	VIII - Eastern Visayas	4,547,150	23,234.78	196
14	XIII - Caraga	2,804,788	21,120.56	133
15	Bangsamoro Autonomous Region in Muslim Mindanao (BARMM)	4,404,288	36,650.95	120
16	MIMAROPA Region	3,228,558	29,606.25	109
17	Cordillera Administrative Region (CAR)	1,797,660	19,818.12	91

[a] *Excludes 2,098 Filipinos in Philippine embassies, consulates, and missions abroad.*
[b] *Estimated total land area of the Philippines. Regional details do not add up to national total due unfinished cadastral surveys.*

Sources: Philippine Statistics Authority, *2020 Census of Population and Housing*
Land Management Bureau, *2013 Masterlist of Land Areas of Cities and Municipalities*

Figure 5 Distribution of Regions Per Cluster



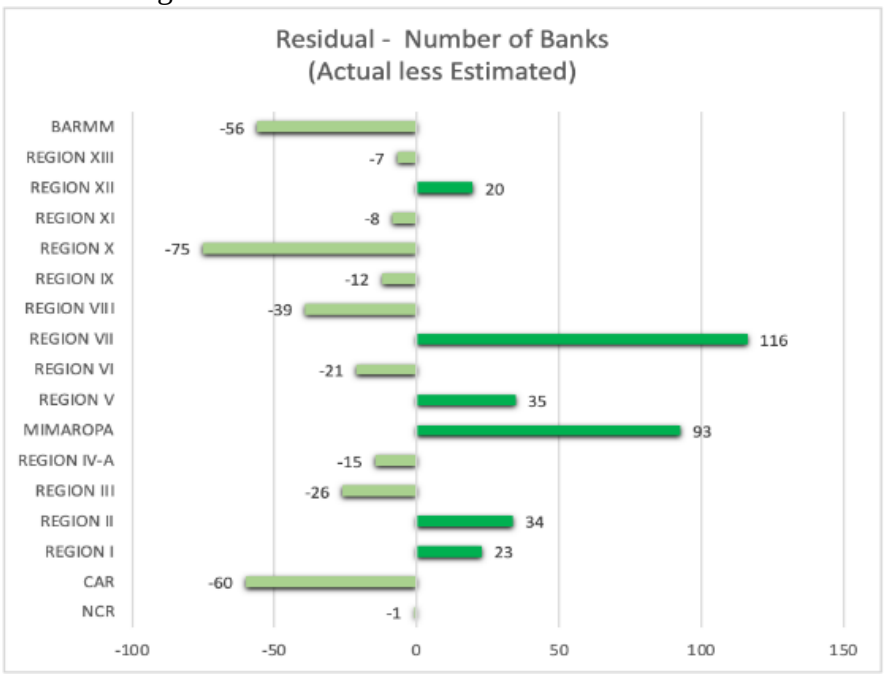
Hence, the clusters of the number of banks were named based on the population density. Note that population density is the representative explanatory variable from demographic characteristics of the Regions.

4. Region in the Philippines that Needs More Banks to Promote Financial Inclusion

The graph of actual number of banks in the Region less the estimated number of banks using the model shows that Regions X, CAR, and BARMM are the top three regions with the lowest residual value. Other regions with negative residuals are: XIII, XI, IX, VIII, IV-A, and III. It can be said that these regions with negative residuals are the regions which need more banks in the area.

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

Figure 6 Residual Chart of the Number of Banks



5. Best Locations for the future Banks/Branches

Based on the results of the residual chart of the number of banks, future banks are best be located or put up in the Regions X, CAR, and BARMM. These are the regions with the highest negative difference between the actual and estimated number of banks in the region.

This study was designed to determine the number of banks needed in the Philippine Regions. With the available regional data for categories in demographic, economic, labor market and market potential, the researcher utilized the OLS Regression Analysis, Global Moran's I Analysis, and Cluster and Outlier Analysis of ArcGIS. The charts showing the characteristics of each Regions per explanatory variable tells us that BARMM and NCR are the extreme values in Number of Banks, Number of establishments in operation, Number of establishments in education sector, Number of transportation and storage establishments, Average family Income, Higher education graduates, Functional literacy rate, Share of families ownership-personal computer and Proportion of population 10 to 64 years old using the internet for email/research. CAR is the region, which is lowest in terms of population, population density, and economically active population but highest in Average family savings. Region IV-A has the highest population, and economically active population rate. The region with the lowest population growth rate and share of families' ownership-cellular phone is Region VIII. Region VI has the highest certified alienable and disposable land while NCR has the lowest. BARMM has the highest unemployment rate while Region X has the lowest.

The research examined which among these variables has a significant relationship with the number of banks in the region using the OLS regression Analysis that tests the significance of the correlation coefficients. Every explanatory variable in this study has a linear relationship with the number of banks. However, population growth rate, certified alienable and disposable land, average savings, and unemployment rate have a weak relationship with the number of

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

banks in the region and these relationships are tested to be not significant. Higher-Education-Graduates AY 2018-2019, Number of Establishments in Operation, Number of Transportation and Storage Establishments, Number of Establishments in Education Sector, and Population Density are the top five variables that have the strongest significant relationship with the number of banks in the region.

After the researcher determined the significant variables, the Exploratory Regression Analysis was run to find the best model in estimating the number of banks in the Philippine Regions. The researcher wanted to have a model which represents every category of the exploratory variables. Therefore, among the passing models, the researcher choses the second model with the highest significance because the first model with the highest significance do not represent every category of the exploratory variables. This chosen model consists of the following exploratory variables: Population density, economically-active population, functional literacy rate, and families-ownership-personal computer.

5. CONCLUSION

The results of the OLS and Exploratory Regression Analyses in ArcGIS show significant variables as well as the passing models for this study. The researcher opted to the model which represents the four categories of variables. The researcher also utilized Global Moran's I analysis in ArcGIS. After the analyses were done, the researcher concluded that significant variables for the number of banks are used in determining the best model for estimation. The results from Global Moran's I analysis showed that there is a clustering of high number of banks in the Philippines. Regions III and IV-A are clustered based on the results of Cluster and Outlier Analysis. Clustering factors are population and establishments as these are the primary clients of the banks.

Since the number of banks is positively correlated mainly to Higher education graduates, number of establishments in operation, number of transportation and storage establishments and the most significant exploratory variable is the population density, we may say that as these exploratory variables increases there is a need to increase the number of banks in the area.

The best model for this study in estimating the number of banks in the region is Number of Banks = -1260.033604+(0.101673*C) +(0.000159*D) +(11.286635*M) +(9.699732*N). The result of this model is the estimated number of banks needed in the region. The difference between the actual and estimated number of banks is the indicator of which regions are lacking physical banks. Based on the results of this study, Regions X, CAR, and BARMM are the best locations for the future physical bank to be established.

6. REFERENCES

1. D. Ansong, G. Chowa, and B. Adjabeng, Spatial analysis of the distribution and determinants of bank branch presence in Ghana, 2015.
2. A. Gorfu, and A. Mamo, Determinants of the Geographical Distribution of Bank Branches: A Case in Ethiopian Commercial Banking Industry, 2013.

Journal of Corporate Finance Management and Banking System
ISSN: 2799-1059
Vol: 04, No. 03, April-May 2024
http://journal.hmjournals.com/index.php/JCFMBS
DOI: https://doi.org/10.55529/jcfmbs.43.1.15

3. N.M. El-Meouch, R. Tésits, and L. Alpek, Measuring Spatial Distribution in the Banking System in Hungary, 2022.
4. (2023) ArcGIS Pro, ESRI website. [Online]. Available: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-multivariate-clustering-works.htm
5. M. Rysman, R. M. Townsend, and C. Walsh, Branch Location Strategies and Financial Service Access in Thai Banking, 2022.
6. M. Boďa and K. Čunderlíková, Determinants of bank branch density: a case study of Slovakia, 2020.
7. Perceptive Analytics, Location Strategy to Improve Effectiveness of a Branch Network, 2013.
8. S. Bellintani, A. Celani and G. Ciaramella, Methods and models for building bank branches. A Multicriteria approach for a decision-making tool, 2020.
9. S. Avetisyan, Criteria Used in the Location Analysis in Banking Literature, 2018.
10. Z. Namaziana, and E. Roghanianb, A decision problem for bank branch site selection: A GIS Mapping perspective with Maximal Covering Location Problem: A case study of Isfahan, Iran, 2019.