Applied Mathematics and Computational Intelligence
Volume 10, No.1, Dec 2021 [166-175]

Regression Analysis on Crime Rates in Kaduna State from 2002-2016

Rilwanu Yakubu Mohammed[1], Shehu Sunusi Lalin[2], Ishaya Adamu Dansadiq[3]

[1]Department of Maths/Statistics, Isa Mustapha Agwai I Polytechnic, Lafia
[2]Department of Mathematics, Nasarawa State University, Keffi
[3]Department of Maths/Statistics, Isa Mustapha Agwai I Polytechnic, Lafia

* Corresponding author: shehuslalin@nsuk.edu.ng

Received: 30 April 2021; Accepted: 11 August 2021; Available online: 20 September 2021

ABSTRACT

The study determines the effect of unemployment on robbery, murder, rape, stealing, and drug abuse. Secondary method of data collection was used to collect the data from the Nigeria Police State Command, Kaduna and other law enforcement agencies in the state. Descriptive statistics and regression analysis were used to analyze the data obtained. The result showed that in 2003 reported cases were high in robbery with 132 persons, murder 69 persons, rape 32 persons, stealing 250 persons and drug abuse was 176 persons. While the year 2016 recorded low cases with robbery having 58 persons, murder 96 persons, rape 84 persons, stealing 94 persons and drug abuse 143 persons. Therefore, we concluded that there is no significant difference between unemployment rate and robbery, murder, stealing and drug abuse and the case of rape has p-value 0.0104 < 0.05 which was statistically significant at 5% level of significant. The Nigerian police and other law enforcement agencies are therefore advised to improve on security measures in the state.

Keywords: Regression, Crime rate, Descriptive statistics, Multiple regression, Analysis.

1 INTRODUCTION

Crime is one of the problems that continually bedevil the existence of mankind. Since early days, crime had been a disturbing threat to human personality, property and lawful authority. This is because it is associated with fear, loss of properties, injuries to victims and in some cases loss of human lives. The relationship between crime and evolution of mankind is considered a historical one as Cain (first son of Adam and Eve) committed the first violent crime when he murdered his brother Abel because of jealousy. Today, in the modern complex world, the situation is highly disturbing. Crime started in the primitive days as a simple and less organized issue, and ended today as very complex and organized. Therefore, the existence of crime and its problems have spanned the history of mankind. Due to the complex nature of the subject of crime, its causes and consequences, various academic disciplines such as criminology, sociology, geography, psychology and demography conducted a lot of studies on it.

According to [1], a considerable variability in the forecasting performance across models, cities, crimes, and forecast horizons. He uses a common panel data set of city level crime rates from 1980-2004, and illustrated ability of several regression models to forecast crime rates. He noted that there

is evidence of heterogeneity across cities and that heterogeneous models do not perform notably better than the homogeneous alternatives.

Crime data which consists of eighteen major crimes reported to the police between the periods of 1996 - 2014 were analyzed by [2]. Crimes consist of robbery, kidnapping, house and store breakings, theft/stealing, grievous hurt and wounding, murder, rape, assault etc. The report showed that there were average of 56 cases of murder, 152 cases of armed robbery, 12 cases of suicide, 541 cases of grievous harms/wounding, 1196 cases of assaults, 392 cases of burglary/stealing, 263 cases of store breaking, 544 cases of house breaking, 1385 cases of stealing/theft, 99 cases of rape/indecent assaults, 12 cases of kidnapping, 442 cases of false pretense/cheating, 110 cases of unlawful possession, 16 cases of arson, 327 cases of breach of peace, 23 cases of forgery, 8 cases of child stealing and 29 cases of unnatural offences. Using correlation analysis and principal component analysis (PCA) were employed to explain the correlation between the crimes and to determine the distribution of the crimes in the state. The result obtained showed a significant correlation between store breaking and stealing, Assaults have a strong positive relationship with false pretense and breach of peace, kidnapping has a strong positive relationship with false pretense and breach of peace, false pretense has a strong positive relationship with breach of peace, there is weak positive relationship between murder case and armed robbery, suicide, grievous harm/wounding, assaults, store breaking, house breaking, stealing, false pretense, unlawful possession, breach of peace, forgery, child stealing and unusual crimes. Murder has a positive relationship with rape and arson, and a weak negative relationship with burglary. Armed robbery has a positive relationship with kidnapping and weak positive relationship with other forms of crimes, a weak negative relationship with burglary, store breaking, stealing and unlawful possession, suicide has a positive relationship with rape, grievous harm/wounding has a positive relationship with assaults, stealing/theft is more prevalent in Oyo State, besides stealing/theft, assaults is prominent among other crimes. The PCA has suggested retaining six components that explain about 83.79 percent of the total variability of the data set. See the summary in Table 4.1 below in data presentation.

[3] discovered drug usage, as proxies by drug deaths, increased from mid-1980's to about 1988-1989 and that, felony and drug arrests also increases substantially. It also shows an increase in murders, assaults, and motor vehicle thefts but robberies increased in the later 1980s and burglaries declined throughout the 1980s. It was further shown in the report that arrest rates and total arrests for non-drug crimes did not decline during this period of increased drug arrests.

[3] used multivariate analysis and found that robberies, burglaries and motor vehicle thefts increased when there were unexpected increases in drug usage. They however did not find any a relationship between drug use and murders or assaults, as well constant arrest rates by police. Additionally, they found that evidence of police deterrence, either directly, or through arrests, of property-related and assault offenses, but not for murders. The time series approach, established a causal relationship between drug usage and property- related felonies.

Study by [4] shows that although crime rates tend to increase with population density, the correlation is not significant. Hence, study suggests that violent and property crime rates are positively associated with percentage of commercially zoned areas, percentage of one-person household and unemployment rate [5].

Related Studies have also attributed crime to other factors such as unemployment ([6], [7], [8] and [9]); absence/presence of tree canopy [10]; presence of young populations Andresen[7],

combination of industrial/commercial land use with household dwellings [11]; urbanization, weak criminal justice system and negative role models [9].

However, since China implemented economic reform and an open-door policy in the late 1970s, crime reporting has increased approximately 50 times the rates observed in the 1950s and 1960s..The Nigerian National Bureau of Statistics in its report [12] released that the major crimes in Nigeria include rape, kidnapping, murder, burglary, fraud, terrorism, robbery, cyber-crimes, bribery and corruption, money laundering and so on. State like Lagos, Abuja, Delta, Kano, Plateau, Ondo, Oyo, Bauchi, Adamawa and Gombe States made the top ten list of states with high number of crimes.

The fourth involves sanction risk perceptions. Establishing the link between risk perceptions and sanction regimes is imperative; unless perceptions adjust, however crudely, to changes in the sanction regime, desired deterrent effects will not be achieved [13].

Nagin explain that over the past four decades, much has been learned about the foundations of deterrence that were laid out more than two centuries ago by Cesare Beccaria and Jeremy Bentham. We now know that deterrence is ubiquitous but that the effects are heterogeneous, ranging in size from seemingly null to very large. There is little evidence that increasing already long prison sentences has a material deterrence effect. Evidence on the deterrent effect of the certainty of punishmentis more consistent, but the source of the effect is less clear. In this essay I have argued that the certainty effect stems primarily from police functioning in their official guardian role rather than in their apprehension agent role. These conclusions have important policy implications that are developed in detail in Durlauf and Nagin [13].

An individual committing crime faces costs from law-enforcement activities. The severity of the punishment including fines and jail time is one part of the total cost, and the other part is the probability of getting caught. Therefore, the costs will equal the probability of punishment (p) times the cost of punishment (c). Thus, the net expected returns from crime equal the probability of punishment (p) times the cost of punishment (c). Crime reduction can occur through reducing the benefits of crime, raising the probability of being caught, or increasing the costs of punishment conditional upon being caught [14].

2 MATERIAL AND METHODS

The data used for this study were numbers of persons who committed various crimes such as robbery, murder, rape, stealing, unemployment and drug abuse reported to Kaduna state police state command, Kaduna, Nigeria. Crime data were also collected through law enforcement report which reflect crimes that were reported, recorded and not subsequently cancelled from 2002-2016. The data are reliable that you can obtain from Kaduna police state command. Multiple Regression Analysis was used to determine the effect of unemployment on crime rates (robbery, murder, rape, stealing and drug abuse) using R-Statistical Packages.

2.1 Model and Data

In Statistics, linear regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables denoted by x. The case of one

explanatory variable is called simple linear regression. For more than one explanatory variable, the process is called multiple linear regressions [14].

In linear regression, the relationships are modeled using linear predictor functions whose unknown model parameters are estimated from the data.

In general, the dependent variable or response Y may be related to k independent or regressor variables. The model;

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon_i \quad (1)$$

is called a multiple linear regression model with k regressor variables. The parameters $\beta_j, j = 1, \dots, k$, are called the regression coefficients. This model describes a hyper-plane in the k dimensional space of the regressor variables $\{x_j\}$. The parameter β_j represents the expected change in response Y per unit change in x_j when all the remaining regressors x_j $(i \neq j)$ are held constant.

Multiple linear regression models are often used as approximating functions. That is, the true functional relationship between Y and $x_1, x_2, \dots, x_k$ is unknown, but over certain ranges of the independent variables the linear regression model is an adequate approximation.

2.1.1 Model Calibration

Matrix Approach to Multiple Regressions.

Suppose that there are k regressor variables and n observations, $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$, $i = 1,2, \dots, n$ and that the model relating the regressors to the response is

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \; i = 1,2, \dots, n \quad (2)$$

This model is a system of n equations that can be expressed in matrix notation as

$$Y = X\beta + \varepsilon \quad (3)$$

$$\text{Where } Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \qquad X = \begin{pmatrix} x_{11}, & x_{12}, & \dots & x_{1n} \\ x_{21}, & x_{22}, & \cdots & x_{11} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1}, & x_{n2}, & \cdots & x_{nn} \end{pmatrix}, \; \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} \text{ and } \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (4)$$

In general, Y is an $(n \times 1)$ vector of the observations, X is an $(n \times p)$ matrix of the levels of the independent variables, β is a $(p \times 1)$ vector of the regression coefficients, and is an $(n \times 1)$ vector of random errors. Where $p = k + 1$.

To find the vector of least squares estimators, $\hat{\beta}$ that minimizes the sum of square of deviation between observed and predicted Y (error), we use the equation

$$\hat{\beta} = X^{-1}Y \quad (5)$$

Which is the least squares estimator for β in the regression equation [14].

In matrix notation, the estimate is obtained as

$$\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \vdots\\ \hat{\beta}_k\end{pmatrix}=\begin{pmatrix} n & \sum_{i=1}^{n}x_{i1} & \sum_{i=1}^{n}x_{i2} & \cdots & \sum_{i=1}^{n}x_{ik}\\ \sum_{i=1}^{n}x_{i1} & \sum_{i=1}^{n}x_{i1}^2 & \sum_{i=1}^{n}x_{i1}x_{i2} & \cdots & \sum_{i=1}^{n}x_{i1}x_{ik}\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ \sum_{i=1}^{n}x_{ik} & \sum_{i=1}^{n}x_{ik}x_{i1} & \sum_{i=1}^{n}x_{ik}x_{i2} & \cdots & \sum_{i=1}^{n}x_{ik}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum_{i=1}^{n}y_i\\ \sum_{i=1}^{n}x_{i1}y_i\\ \vdots\\ \sum_{i=1}^{n}x_{ik}y_i\end{pmatrix} \quad (6)$$

The matrix notation of the fitted model is $\hat{y}=X\hat{\beta}$ and the fitted regression model is

$$\hat{y}=\hat{\beta}_0+\sum_{j=1}^{k}\hat{\beta}_j x_i \quad i=1,2,\dots,n \quad (7)$$

2.1.2 Assumptions of the Linear Model

(i)Linear relationship

(ii) Homoscedasticity

(iii) Multicollinearity

(iv) Normality Tests

2.1.3 Analysis of Variance

The model was tested using analysis of variance (ANOVA) and summarized in Table 1 below.

Table 1: Analysis of Variance for Testing Significance of Regression Model

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_{cal}
Regression	$SS_R=\hat{\beta}'(X'y)-n\bar{y}^2$	k	$MS_R={SS_R}/{k}$	${SS_R}/{MS_E}$
Residual (error)	$SS_E=SS_T-SS_R$	$n-k-1$	$MS_E={SS_E}/{n-k-1}$	
Total	$SS_T=y'y-n\bar{y}^2$	$n-1$		

Where *k* is number of parameters, *n* is number of observations.

The null hypothesis that the model is insignificant and is rejected if F_{cal} F_{cal} is greater than $F_{\alpha(k),(n-p)}$.

2.1.4 Coefficient of Determination (R^2)

The proportion of variation of the dependent variable explained by the auxilliary variables in a regression analysis is given by

$$R^2=\frac{SS_R}{SS_T}=1-\frac{SS_E}{SS_T} \qquad R^2_{adj}=1-\frac{SS_E/(n-k-1)}{SS_T/(n-1)} \quad (8)$$

2.1.5 Tests on Individual Regression Coefficients

To test the significance of each coefficient in regression analysis, t-test was used and gives the equation below

$$t = \frac{\widehat{\beta}_j}{\sqrt{\widehat{\sigma}^2 c_{jj}}} = \frac{\widehat{\beta}_j}{se(\widehat{\beta}_j)} \text{ where } \{c_{jj}\} = (X'X)^{-1} \quad (9)$$

Hypothesis tested H_0: $\beta_j = 0$ vs H_1: $\beta_j \neq 0$. Null hypothesis is rejected if $t_{cal} > t_{\alpha/2(n-k-1)}$.

3 RESULTS AND DISCUSSION

Table 3.1: Various numbers of crimes reported from 2002-2016 in Kaduna state

Time	Robbery	Murder	Rape	Stealing	Unemployment	Drug Abuse
1	2002	121	48	21	246	16	172
2	2003	132	69	32	250	20	176
3	2004	62	65	36	241	16	169
4	2005	59	51	28	251	12	131
5	2006	55	46	20	212	14	122
6	2007	40	52	31	234	9	149
7	2008	50	70	41	207	13	176
8	2009	57	73	50	282	12	159
9	2010	55	69	35	148	13	124
10	2011	38	95	41	279	9	156
11	2012	64	55	41	102	6	120
12	2013	72	84	48	157	10	164
13	2014	65	84	63	320	8	188
14	2015	56	92	92	417	6	189
15	2016	58	96	84	94	5	143

Source; Kaduna State Police Head Quarter, 2016.

Table 3.2: 18 major crimes reported to police in Oyo State (1996-2014) in researched by Olufolabo et al, [2]

Crime	**Number**	**Crime**	**Number**	**Crime**	**Number**
Murder	56	Store breaking	263	Unlawful possession	110
Armed robbery	152	House breaking	544	Arson	16
Suicide	12	Stealing/theft	1385	Breach of peace	327
Grievous/wounding	541	Rape/Ind assault	99	Forgery	23
Assaults	1196	Kidnapping	12	Child stealing	08
Burglary/stealing	392	False /cheating	442	Unnatural offences	29

Table 3.3: Estimates, Std. Error, T-valve and P-value Computed

	Estimate	**Std. Error**	**t value**	**Pr(>|t|)**
(Intercept)	6.486322	5.221721	1.242	0.2456
x_1	0.061026	0.035993	1.696	0.1242
x_2	-0.002083	0.066207	-0.031	0.9756
x_3	-0.126505	0.039193	-3.228	0.0104 *
x_4	0.001140	0.012464	0.091	0.9291
x_5	0.040732	0.056500	0.721	0.4893

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

y = unemployment, x_1 = robbery, x_2 = murder, x_3 = rape, x_4 = stealing, x_5 = drug abuse.

Table 3.1 above present the number of crimes reported to Police Head Quarters of Kaduna state police command from 2002-2016. In 2002 reported robbery was 121, murder was 48, rape 21, stealing 246, unemployment 16 and drug abuse was 172, while the year 2003 had 132, 69, 32, 250, 20, and 176 for rubbery, murder, rape, stealing, unemployment, and drug abuse respectively. In 2004, reported cases stood at 62 for rubbery, 65 for murder, 36 for rape, 241 for stealing, 16 for unemployment and 169 for drug abuse. The year 2005 had 59, 51, 28, 251, 12 and 131 as reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse respectively. In 2006, reported cases of rubbery were 55, murder was 46, rape 20, stealing 212, unemployment 14 and drug abuse was 122. The reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse for the year 2007 are 40, 52, 31, 234, 9 and 149 respectively. In 2008, reported cases stood at 50 for rubbery, 70 for murder, 41 for rape, 207 for stealing, 13 for unemployment and 176 for drug abuse. The year 2009 had 57, 73, 50, 282, 12 and 159 as reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse respectively. In 2010, reported cases of rubbery were 55, murder was 69, rape 35, stealing 148, unemployment 13 and drug abuse was 124. The reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse for the year 2011 stood at 38, 95, 41, 279, 9 and 156 respectively. The year 2012 had 64, 55, 41, 102, 6 and 120 as reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse respectively. The year 2013 had 72, 84, 48, 157, 10 and 164 as reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse respectively. In 2014 reported robbery was 65, murder was 84, rape 63, stealing 320, unemployment 8 and drug abuse was 188; while the year 2015 had 56, 92, 92, 417, 6, and 189 for rubbery, murder, rape, stealing, unemployment, and drug abuse respectively. Finally, the year 2016 had 58, 96, 84, 94, 5 and 143 as reported cases of robbery, murder, rape, stealing, unemployment, and drug abuse respectively.

Results of Table 3.3 shows the intercept was 6.486322 with parameters estimates as 0.061026, -0.002083, -0.126505, 0.0011407 and 0.040732 for robbery, murder, rape, stealing, and drug abuse respectively. The standard error is 5.22172, 0.0355993, 0.066207, 0.039193, 0.012464, 0.056500 with p-value $x_1 = 0.1245$, $x_2 = 0.9756$, $x_4 = 0.9291$, $x_5 = 0.4893 > 0.05$ that there is no sufficient reason

to reject null hypothesis at 5% level of significant. Therefore we concluded that there is no significant different between unemployment rate and robbery, murder, stealing and drug abuse. While the case of rape has p-value 0.0104 < 0.05 which was statistically significant at 5% level of significant. The positive signs of the coefficients indicate that there is increase in the number of crimes in x_1, x_4, x_5 and negative sign indicate decreases in the numbers of crime x_2, x_3, during the period under study.

4 CONCLUSION

Based on the analysis carried out it was observed that there is no significant difference between unemployment rate and robbery, murder, stealing and drug abuse. While the case of rape has *p*-value 0.0104 < 0.05 which was statistically significant at 5% level of significance and recommended that the model for prediction of future crimes of robbery, murder, rape, stealing and drug abuse in Kaduna state is unemployment (y) = $6.486322 + 0.061026x_1 - 0.002083x_2 - 0.126505x_3 + 0.001140x_4 + 0.040732x_5$.

The researchers recommends the model for prediction of crime rate by the Nigeria police and other law enforcement agencies in order to improve on security measures in the country.

ACKNOWLEDGEMENT

We acknowledged Kaduna state police Head Quarter and JIYAH Amina Muhammad an undergraduate student of Ahmadu Bello University Zaria, Nigeria for supporting us with data for this research. We also acknowledged the editorial of AMCI UNIMAP for giving us the opportunity to publish in their journal.

REFERENCES

[1] V. P. John, *Forecasting Crime: A City Level Analysis*, Charlottesville: Department of Economics University of Virginia, 2007.

[2] O. O. Olufolabo, O. J. Akintande and M. I. Ekum, "Analyzing the Distribution of Crimes in Oyo State (Nigeria) using Principal Component Analysis (PCA)," *IOSR Journal of Mathematics*, XI(3), pp. 90–96, June 2015.

[3] H. Corman, and H. N. Mocan, "A Time Series Analysis of Crime and Drug Use in New York City," *NBER Working Paper*, pp. 1–4, 1996.

[4] B. H. Baltagi, "Estimating an economic model of crime using panel data from North Carolina," *Journal of Applied Econometrics*, vol. 21, pp. 543–547, 2006.

[5] N. J. Kepple and B. Freisthler, "Exploring the ecological association between crime and medical marijuana dispensaries," *Journal of Studies on Alcohol and Drugs*, vol. 73, no. 4, pp. 523–530, 2012.

[6] K. Edmark, "Unemployment and crime: Is there a connection?," *The Scandinavian Journal of Economics*, vol. 107, no. 2, pp. 353–373, 2005.

[7] M. A. Andresen, "A spatial analysis of crime in Vancouver, British Columbia: a synthesis of social disorganization and routine activity theory," *The Canadian Geographer*, vol. 50, no.4, pp. 487–502, 2006.

[8] D. G. Omotor, "Socio-economic determinants of crime in Nigeria," *Pakistan Journal of Social Sciences*, vol. 6, no. 2, pp. 54–59, 2009.

[9] A. B. Hassan, F. D. Lass and J. Makinde, "Cybercrime in Nigeria: Casuses, effects and the way out," *ARPN Journal of Science and Technology*, vol. 2 no. 7, pp. 626–631, 2012.

[10] A. Troy, J. M. Grove and J. O'Neil-Dunne. "The relationship between tree canopy and crime rates across an urban-rural gradient in the greater Baltimore region," *Landscape and Urban Planning*, vol. 106, pp. 262-270, 2012.

[11] J. O. Shopeju, "Urbanization and crime in Nigeria," *International Journal of Agricultural Sciences, Sciences, Environment and Technology*, vol. 2, no. 1, pp. 154–163, 2007.

[12] S. D. Nagin, "Deterrence in the Twenty-First Century," *Crime and Justice*, vol. 42, no. 1, pp. 199-263, Aug. 2013.

[13] O. Alison, "The Economics of Crime: An Analysis of Crime Rates in America," *The Park Place Economist*, vol. 10, no.1, p. 30–35, 2002.

[14] J. Cohen, P. Cohen, S. G. West and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences,* 2nd ed., Hillsdale, NJ: Laurence Erlbaum Associates, 2003.

[15] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 3rd ed., W. Anderson and J. Welter, Eds., New York: John Wiley & Sons, Inc., 2003.

[16] Nigeria's National Bureau of Statistics, "Crime Statistics: Reported Offences by Type and States," 2016. [Online]. Available: www.nigerianstat.gov.ng