A Thesis On

Analysis of Crime of Bangladesh Using Multiple Linear Regression

Md Imdadul Hoque

Roll: ASH1501052M

Session: 2014-2015

DEPT. OF COMPUTER SCIENCE AND TELECOMMUNICATION ENGINEERING

NOAKHALI SCIENCE AND TECHNOLOGY UNIVERSITY

NOAKHALI-3814, BANGLADESH,

CERTIFICATION

This thesis titled as "Analysis of Crime Of Bangladesh Using Multiple Linear Regression", submitted by Md Imdadul Hoque, Roll: ASH 1501052M, Session: 2014-2015, have been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Computer Science and Telecommunication Engineering as B.Sc. Engineering (CSTE) to be awarded by the Noakhali Science And Technology University.

Candidate

Md Imdadul Hoque
Roll: ASH 1501052M
Session: 2014-2015

Supervisor:

Abul Kalam Azad

Assistant Professor
Dept. of Computer science and Telecommunication Engineering
Noakhali Science and Technology University
Sonapur, Noakhali-3814, Bangladesh

Abstract

For globalization, we are connected to each other very easily. We are getting busy all day long. In this busy life it is very awful that crime increasing day by day. Hence the necessity of taking proper steps to predict or prevent the crimes is very significant. In this case, machine learning and data mining techniques can play an important role to discover future trends and patterns of crime, so that, police and detective agencies can be alert about that crime and take necessary steps to prevent it. In this paper, Multiple linear regression model is used to predict future crime trends of Bangladesh. The dataset collected form Bangladesh police, World Bank and Bangladesh Bank. Multiple linear regression (MLR) model is trained on this dataset. After training the model, crime prediction is done for Murder, Robbery, Narcotics, Women and Child repression for Metropolitan area of Bangladesh. It was found that crime seems to be increased in next two years in accordance with unemployment rate and GDP.

Keyword*: Multiple Linear Regression, GDP, Unemployment rate*

CONTENTS

LIST OF FIGURES

LIST OF TABLE

Chapter 1: Introduction

Crime is an illicit act punishable by a state.There are Several types of crime in Bangladesh such as Murder, Kidnapping, Drug Trafficking,Women and Child Repression, Narcotics, Money Laundering, Extortion, Contract killing, Fraud, Human Trafficking, Robbery, Corruption, Black Marketeering, Political Violence, Terrorism and Abduction etc. are very common. The main objective of this paper is to predict the occurrence of crime in next two years.

1.1 Introduction

Crime and violence are constantly evolving phenomena across the world that is closely linked with the patterns of Urbanization. In developing countries like Bangladesh, there are poorest and deprived inhabitants afflicted by urban crime violence. The world civilization, urbanization is not a new phenomenon and there is evidence that urbanization is closely related to Industrialization. Urbanization is considered as a catalyst that boosts industrialization-led economic growth and facilitate transition from agricultural-based income and employment to non-agriculture based livelihood opportunities.

Urbanization in Bangladesh has been deported to the background of historical antiquities. Still, then the country lacks the huge urban settlement that is identified as towns and cities in the modern sense of the term. Rather, until recently, the country was largely rural. In 1961, more than 5% of the population lived in the urban areas. Since the independence in 1971, Bangladesh has experienced enormously high rate of urbanization which has exceeded more than three times higher than that of the national population growth rate.In fact, cities are growing rapidly so as to accommodate 50% of the world population by 2025. But urban crime is growing even faster, according to the United Nation Crime Prevention and Criminal Justice Balance. A study conducted by the UN International crime and justice research institute (UNICRI) indicates that urban crime has occurred very frequently and it is also hypothesized that more than half of the urban population worldwide have been victimized at least once in the study period 1988-1993.Crime and violence affect all members of society, regardless of sex, age, and income but are more evident in urban areas, especially poor and marginalized neighborhoods. It intrudes into homes, schools, commercial establishments, public transport, and sports and other public venues [1]

In previous, crime had been solved by the right of the criminal justice and law enforcement specialists. At this time, the use of computerized systems to track crimes and trace criminals has increased, computer data analysts helping the law enforcement officers, also helping the detectives to improvise their thinking and capacity of prediction. The process of criminology is used to identify crime and criminal characteristics. Crime occurrence possibility can be determined with the analysis of criminology techniques.

The criminal characteristics are identified by the help of police department, detective agencies and crime branches. Criminology department has been worked with the proceeding of crime tracking ever since 1800 [2].

1.2 Motivation

Bangladesh is a developing country which is run by the government of People's Republic of Bangladesh. Government changes after five years with the political figure. The world is connected with the help of globalization. The world becomes globalized for networking with the effects of internet. For the internet revolution there are many developing countries like Bangladesh are familiar with technology. Because of the pros and cons of technology we have easily connected and developed our communication. Being wrapped in crime easily is one of the important cons of technology. Basically it depends on geographical area, so six areas of Bangladesh known as Metropolitan areas are identified for analysis. In this paper, Multiple Linear Regression (MLR) model was used to analyse crime data that is collected from the website of Bangladesh police. After analysing crime data a prediction will show that may help the Bangladesh police, detective agencies and many other crime branches.

1.3 Objectives

The study of this paper will focus on followings:

1. Concern about Crime.
2. Helping the police and detective agencies.
3. Assuming the number of crimes in next two years.
4. Mark Metropolitan based crime area.

1.4 Expected Output

Predicting crime data of 2019 and 2020 in six metropolitan areas.

Chapter 2: Literature review

Data mining and Machine learning is a powerful technique to analyze the crime data and predict data with different methods. These techniques are helpful to investigate criminal data. Some examples of data mining technique usage to analyze crime data are classification and machine learning algorithms, based on existing research.

2.1 Related works with Crime

Machine learning and data mining techniques can play a significant role to discover future trends and patterns of crime. Linear regression model is used to forecast future crime trends of Bangladesh. The linear regression model was used before to train the data set collected from Bangladesh police. After training the model, crime forecasting was done for dacoit, robbery, murder, women & child repression, kidnapping, burglary, theft and others for different region of Bangladesh [3]**.**

By considering the geographical approach, crime data was analyzed with different types of models such as linear regression model, additive regression model, decision stump model. Linear regression is the best model for crime data analysis. The data cooperation was done with open source data mining software weka**.**

Use of clustering algorithm for a data mining approach to help detect the crimes patterns. K-means clustering with some enhancements to aid in the process of identification of crime patterns [4].

By hotspot mapping technique the identification of crime reduction resource and prediction of crime was done. There are various types of mapping techniques such as point mapping, thematic mapping of geographic areas, spatial ellipses, grid thematic mapping [5].

Four important factors play a role in the analysis of criminal careers: crime nature, frequency, duration and severity. A visual clustering of these criminal careers enable the identification of classes of criminals [6].

A comprehensive survey identified the efficient and effective methods or techniques on data mining for crime data analysis. They pursued the illegal activities of professional fraudsters based on knowledge discovered from their own analysis [7].

Employ an ensemble of data mining classification techniques to perform the crime forecasting. A variety of classification methods such as: One Nearest Neighbor (1NN), Decision Tree (J48), Support Vector Machine (SVM), Neural Network (Neural) with 2 layer networks, and Naïve Bayesian

(Bayes) were used to predict the crime "hotspot". Finally, the best forecasting approach was proposed to achieve the most stable outcomes [8]**.**

Crime is classically unforeseeable and a social nuisance. In this paper, linear regression model is used to forecast future crime trends of Bangladesh. The real dataset of crime is collected from the website of Bangladesh police. The linear regression model is trained on this dataset. After training the model, crime forecasting was done**.**

Data Mining is the procedure which includes evaluating and examining large pre-existing databases in order to generate new information which may be essential to the organization. Aim of this work is to perform a survey on the supervised learning and unsupervised learning techniques that has been applied towards criminal identification [9].

Chapter 3: Methodology

3.1 MLR Model

Multiple linear regression (MLR), also known simply as multiple regression, is a statistical technique that uses several explanatory variables to predict the outcome of a response variable. The goal of multiple linear regression (MLR) is to model the linear relationship between the explanatory (independent) variables and response (dependent) variable [10]

The Multiple Regression Equation with three indepent variables has the form
$y = a + b_1x_1 + b_2x_2 + b_3x_3$
where **a** is the intercept

- b_1,b_2,b_3 are regression co-efficient
- y is the dependent variable
- x_1,x_2,x_3 are indepent variable

3.1.1 Calculation of Regression Coefficient

The normal equation for this multiple regression are:
$\sum yx_1 = b_1\sum x_1^2 + b_2\sum x_1x_2 + b_3\sum x_1x_3$
$\sum yx_2 = b_1\sum x_1x_2 + b_2\sum x_2^2 + b_3\sum x_2x_3$
$\sum yx_3 = b_1\sum x_1x_3 + b_2\sum x_2x_3 + b_3\sum x_3^2$

The following matrix are used to solve this set of equation:

$$A = \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 \\ \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 \\ \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 \end{bmatrix}$$

$$B = \begin{vmatrix} b_1 \\ b_2 \\ b_3 \end{vmatrix} \qquad C = \begin{vmatrix} \sum y\,x_1 \\ \sum y\,x_2 \\ \sum y\,x_3 \end{vmatrix}$$

where
A · B = C
To solve for B, multiply both sides of the equation by the inverse
of A, A- 1.
A-l · A · B = C · A-l
Since A- 1 · A= I, the identity matrix, then
I · B = C · A- l or B = C · A- 1

$$A^{-1}C = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \begin{vmatrix} \sum y\, x_1 \\ \sum y\, x_2 \\ \sum y\, x_3 \end{vmatrix} = B = \begin{vmatrix} b_1 \\ b_2 \\ b_3 \end{vmatrix}$$

The intercept $a = y - b_1x_1 - b_2x_2 - b_3x_3$

$y \quad \sum y \div n$

$x_i' = \sum x_i \div n$
where n is the observation number.

A simple linear regression is a function that allows an analyst or statistician to make predictions about one variable based on the information that is known about another variable. Linear regression can only be used when one has two continuous variables—an independent variable and a dependent variable. The independent variable is the parameter that is used to calculate the dependent variable or outcome. A multiple regression model extends to several explanatory variables

The multiple regression model is based on the following assumptions:

- There is a linear relationship between the dependent variables and the independent variables.
- The independent variables are not too highly correlated with each other.
- $\mathbf{y_i}$ observations are selected independently and randomly from the population.
- Residuals should be normally distributed with a mean of zero and variance.

The coefficient of determination (R-squared) is a statistical metric that is used to measure how much of the variation in outcome can be explained by the variation in the independent variables. R^2 always increases as more predictors are added to the MLR model even though the predictors may not be related to the outcome variable.

R^2 by itself can't thus be used to identify which predictors should be included in a model and which should be excluded. R^2 can only be between 0 and 1, where 0 indicates that the

outcome cannot be predicted by any of the independent variables and 1 indicates that the outcome can be predicted without error from the independent variables.
The Formula for Correlation is [11]

$$r = \frac{n * (sum(x, y) - (sum(x) * (sum(y)))}{\sqrt{((\mathrm{n} * \mathrm{sum}(\mathrm{x})^2) * (\mathrm{n} * (\mathrm{sum}(\mathrm{y}^2) - \mathrm{sum}(\mathrm{y})^2))}}$$

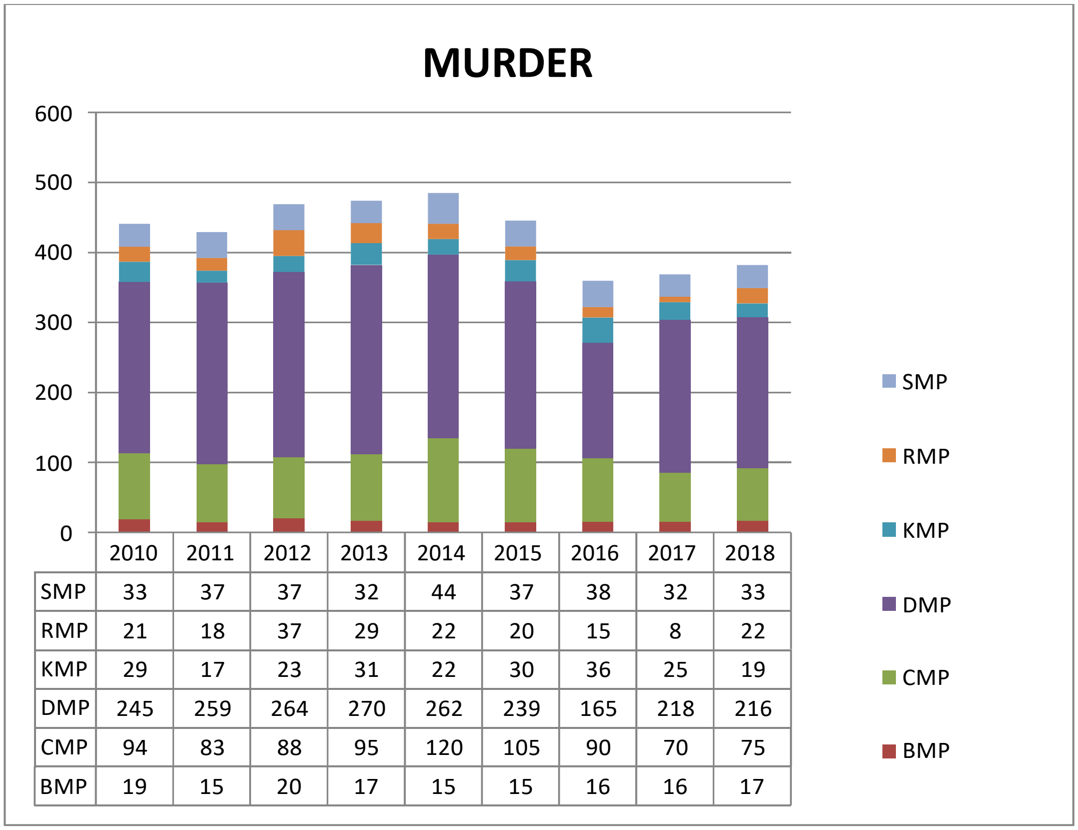
When interpreting the results of a multiple regression, beta coefficients are valid while holding all other variables constant ("all else equal"). The output from a multiple regression can be displayed horizontally as an equation, or vertically in table form.

From the resultant equation we found the value of Intercept, Year, Ur, Gdp coefficient. We used the equation of Multiple Linear Regression such as:
$y = a + b_1x_1 + b_2x_2 + b_3x_3$

So predicted equation:
For Murder, Murder = $a + b_1$*Year $+b_2$*Ur$+b_3$*Gdp
For Robbery, Robbery = $a + b_1$*Year $+b_2$*Ur$+b_3$*Gdp
For Narcotics, Narcotics = $a + b_1$*Year $+b_2$*Ur$+b_3$*Gdp
For women & Child Repression,
Women & Child = $a + b_1$*Year $+b_2$*Ur$+b_3$*Gdp
Where X_1= Dependent Variable (Year)
X_2 = Dependent Variable (Ur)
X_3 = Dependent Variable (Gdp)

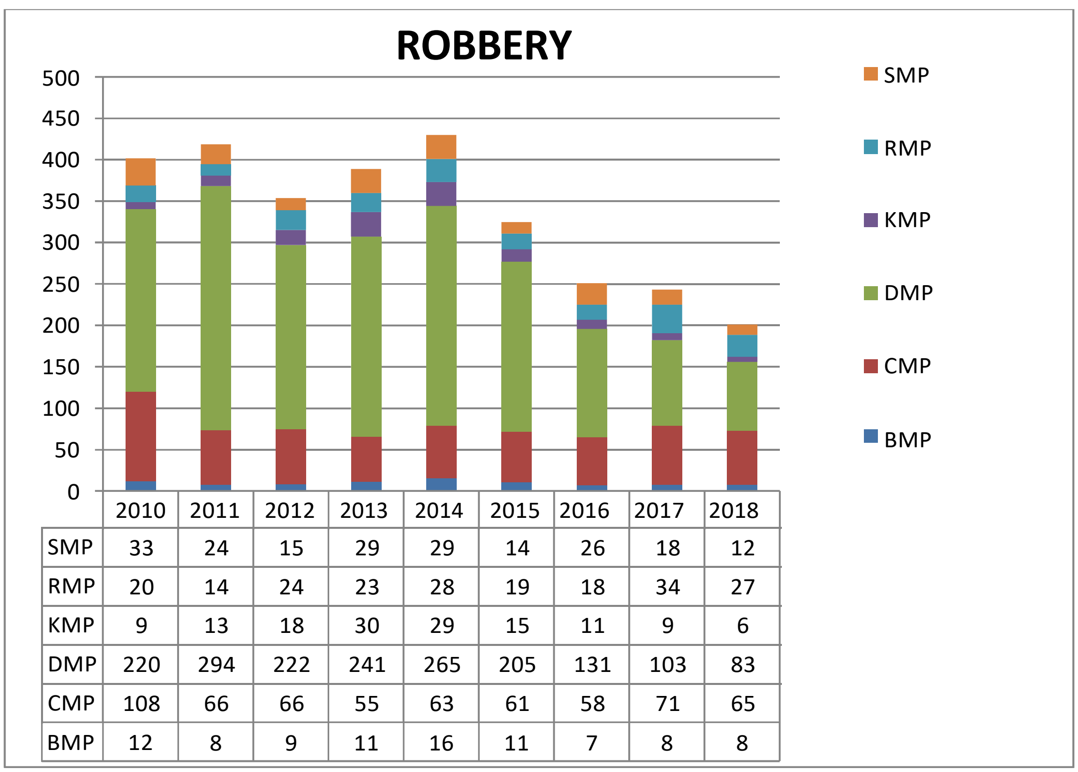
3.2 Data collection

Data are collected from the website of Bangladesh police(14) . From year 2010 to 2018 Murder, Robbery, Woman & Child Repression, Narcotics data was collected [12] [13] [14] [15] [16] [17] [18] [19] [20].
There are six region of metropolitan such as:
DMP = Dhaka Metropolitan Police
RMP = Rajshahi Metropolitan Police
CMP = Chittagong Metropolitan Police
KMP = Khulna Metropolitan Police
SMP = Sylhet Metropolitan Police
BMP = Barisal Metropolitan Police



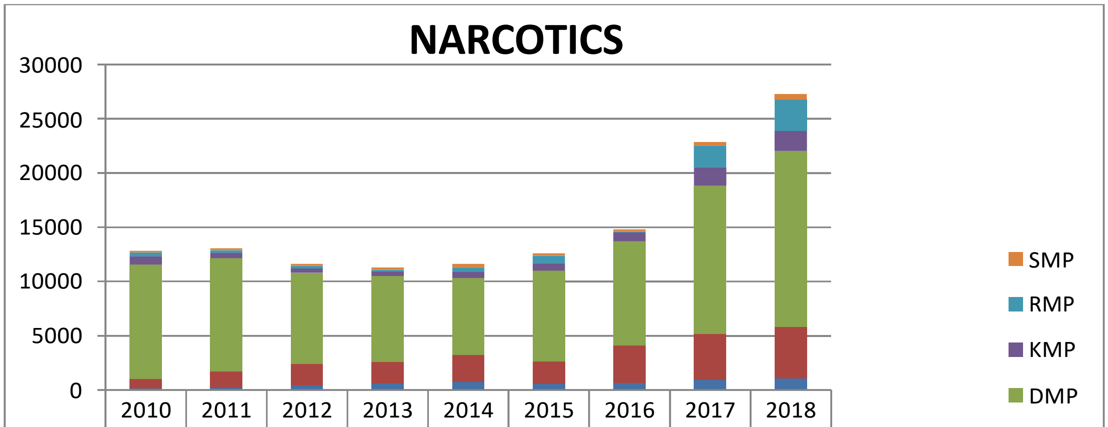
	2010	2011	2012	2013	2014	2015	2016	2017	2018
SMP	33	37	37	32	44	37	38	32	33
RMP	21	18	37	29	22	20	15	8	22
KMP	29	17	23	31	22	30	36	25	19
DMP	245	259	264	270	262	239	165	218	216
CMP	94	83	88	95	120	105	90	70	75
BMP	19	15	20	17	15	15	16	16	17

Figure 3. 1 Murder data of Metropolitan Area from 2010 to 2018



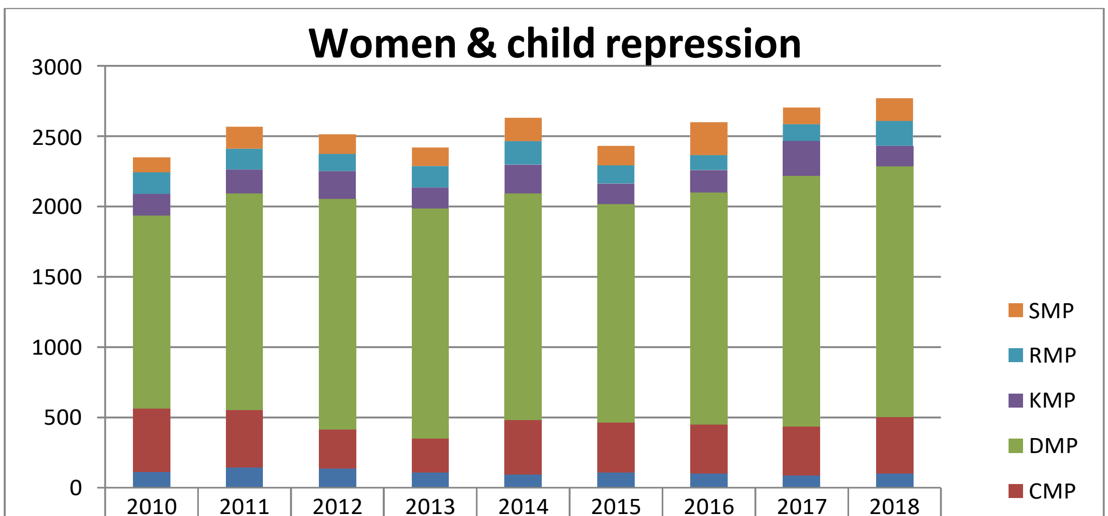
	2010	2011	2012	2013	2014	2015	2016	2017	2018
SMP	33	24	15	29	29	14	26	18	12
RMP	20	14	24	23	28	19	18	34	27
KMP	9	13	18	30	29	15	11	9	6
DMP	220	294	222	241	265	205	131	103	83
CMP	108	66	66	55	63	61	58	71	65
BMP	12	8	9	11	16	11	7	8	8

Figure 3. 2 Robbery data of Metropolitan Area from 2010 to 2018



	2010	2011	2012	2013	2014	2015	2016	2017	2018
SMP	154	178	186	194	311	183	164	301	520
RMP	332	253	229	144	395	733	114	2021	2847
KMP	792	495	437	442	599	659	801	1706	1858
DMP	10535	10405	8345	7927	7063	8365	9627	13638	16215
CMP	866	1541	1976	1977	2488	2090	3457	4260	4747
BMP	155	196	459	607	759	557	655	921	1088

Figure 3. 3 Narcotics data of Metropolitan Area from 2010 to 2018



	2010	2011	2012	2013	2014	2015	2016	2017	2018
SMP	104	158	140	131	165	139	235	120	159
RMP	157	146	122	152	169	131	108	120	177
KMP	153	171	199	153	206	146	158	248	145
DMP	1370	1538	1637	1631	1611	1550	1651	1779	1782
CMP	455	408	281	244	387	357	344	350	405
BMP	112	147	135	110	94	110	105	89	101

Figure 3. 4 Women & Child Repression data of Metropolitan Area from 2010 to 2018.

Table 3. 1 Gross domestic product (GDP) of Bangladesh from 2010 to 2018 [21]

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018
Gdp	5.57	6.46	6.52	6.01	6.06	6.55	7.1	7.3	7.9

Table 3. 2 Unemployment Rate of Bangladesh from 2010 to 2018 [22]

Year	Unemployment Rate
2010	3.379
2011	3.718
2012	4.048
2013	4.426
2014	4.411
2015	4.416
2016	4.350
2017	4.372
2018	4.308

3.3 Analysis Data

For data analysis using MATLab R2016a.

Chapter 4: Result and Discussion

4.1 Result and Discussion

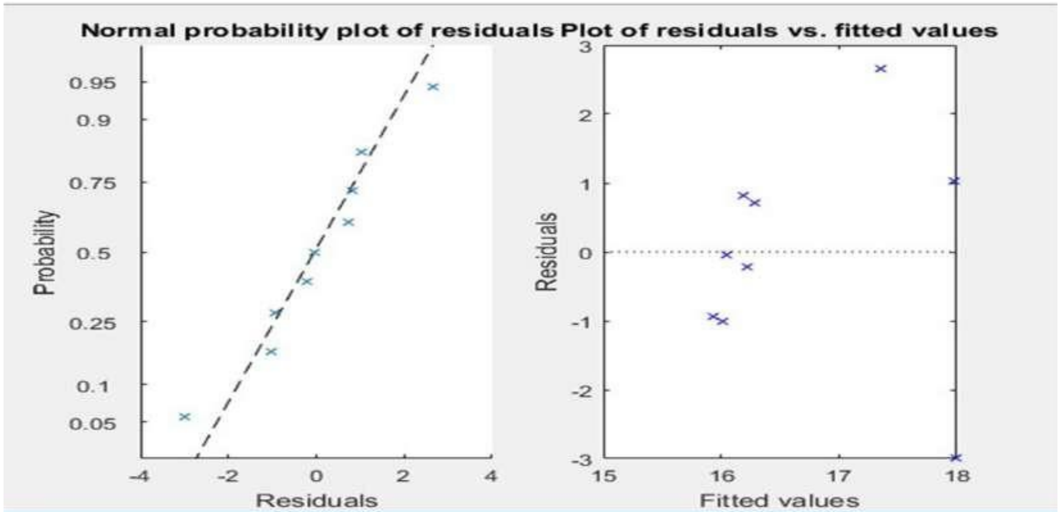
All of the results from the implementation of Multiple Linear Regression (MLR) model that is provided in this section. In this paper, we tried to find the predicted data from analysis. We got some prediction value which is closely related to the actual value. After analyzing the data with MATLab we have shown some screenshot of crime result given below:

4.1.1 Murder Result

Screenshot No-1: BMP Murder Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    _______    ________    _______

    (Intercept)      648.77       1672     0.38803    0.71396
    Year           -0.31417    0.84283    -0.37275    0.72461
    Ur              -1.1459     3.5908    -0.31911    0.76254
    Gdp             0.81512     2.3301     0.34983    0.74072


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 2.01
R-squared: 0.225,  Adjusted R-Squared -0.24
F-statistic vs. constant model: 0.485, p-value = 0.707

Correlation :

    1.0000   -0.3798   -0.4510   -0.1939
   -0.3798    1.0000    0.7651    0.8572
   -0.4510    0.7651    1.0000    0.4661
   -0.1939    0.8572    0.4661    1.0000
```

Screenshot No-2: BMP Murder Figure



Screenshot No-3: CMP Murder Coefficient

Estimated Coefficients:

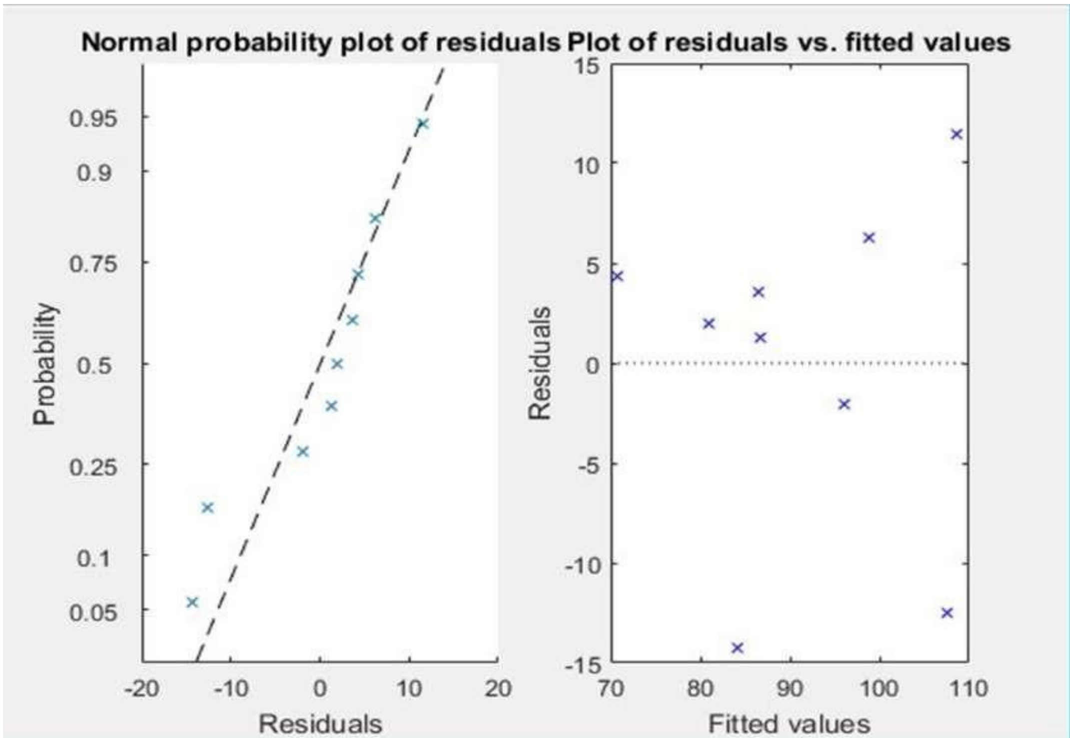
	Estimate	SE	tStat	pValue
(Intercept)	-4813	8897.1	-0.54096	0.61175
Year	2.4878	4.485	0.55469	0.603
Ur	14.425	19.108	0.75494	0.4843
Gdp	-25.166	12.399	-2.0296	0.098158

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 10.7
R-squared: 0.689, Adjusted R-Squared 0.502
F-statistic vs. constant model: 3.69, p-value = 0.097

Correlation :

1.0000	-0.3046	0.1423	-0.6467
-0.3046	1.0000	0.7651	0.8572
0.1423	0.7651	1.0000	0.4661
-0.6467	0.8572	0.4661	1.0000

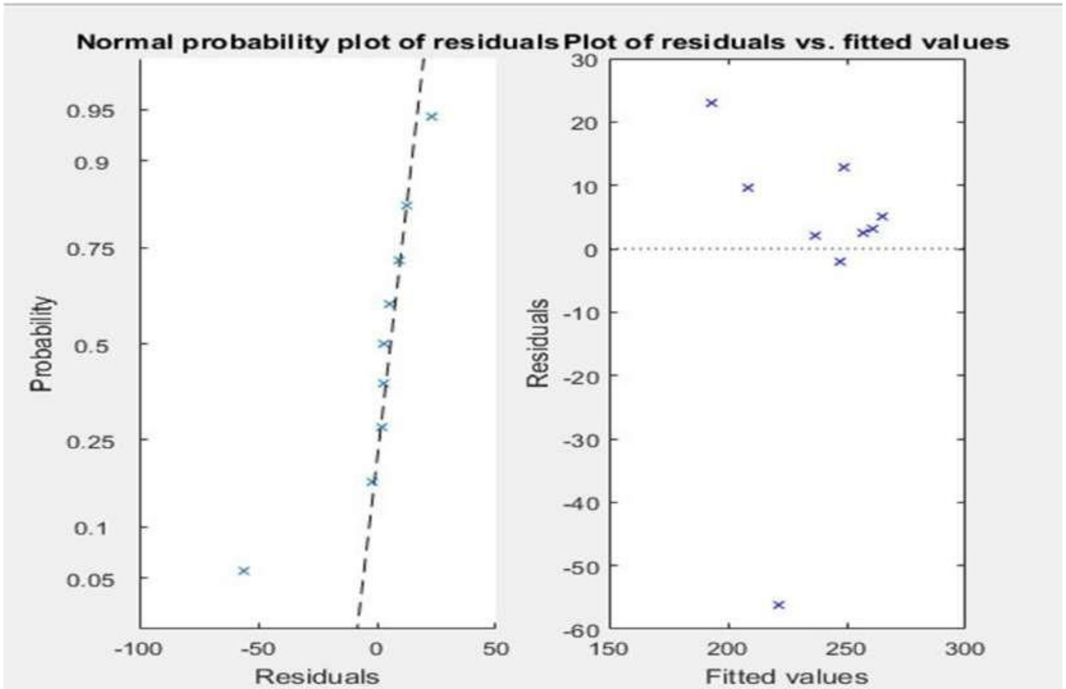
Screenshot No-4: CMP Murder Figure



Screenshot No-5: DMP Murder Coefficient

```
Estimated Coefficients:
                    Estimate      SE        tStat       pValue
                    ________    ______    ________    _______

    (Intercept)       30465      23540      1.2942    0.25216
    Year            -15.147     11.866     -1.2765    0.25786
    Ur               58.111     50.555      1.1494    0.30237
    Gdp              5.6384     32.806     0.17187    0.87028


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 28.3
R-squared: 0.555,  Adjusted R-Squared 0.288
F-statistic vs. constant model: 2.08, p-value = 0.221

Correlation :

    1.0000   -0.6376   -0.2417   -0.6375
   -0.6376    1.0000    0.7651    0.8572
   -0.2417    0.7651    1.0000    0.4661
   -0.6375    0.8572    0.4661    1.0000
```

Screenshot No-6: DMP Murder Figure



Screenshot No-7: KMP Murder Coefficient

Estimated Coefficients:

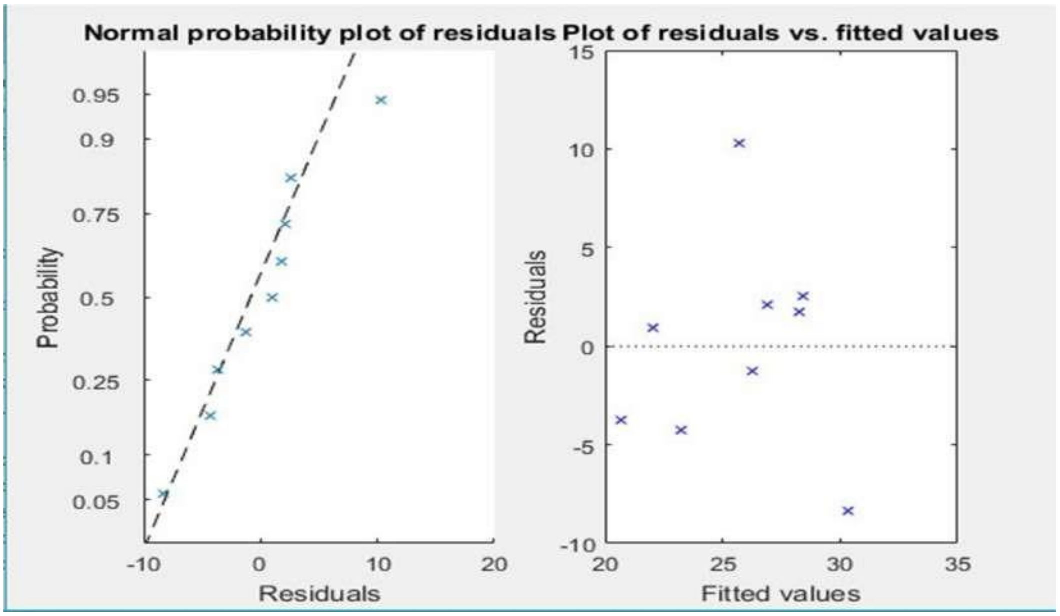
	Estimate	SE	tStat	pValue
(Intercept)	-4638.2	5569.5	-0.83279	0.44291
Year	2.3486	2.8075	0.83651	0.441
Ur	-1.4675	11.961	-0.12269	0.90713
Gdp	-9.0627	7.7617	-1.1676	0.29561

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 6.69
R-squared: 0.269, Adjusted R-Squared -0.17
F-statistic vs. constant model: 0.612, p-value = 0.636

Correlation :

1.0000	0.0665	0.2146	-0.2066
0.0665	1.0000	0.7651	0.8572
0.2146	0.7651	1.0000	0.4661
-0.2066	0.8572	0.4661	1.0000

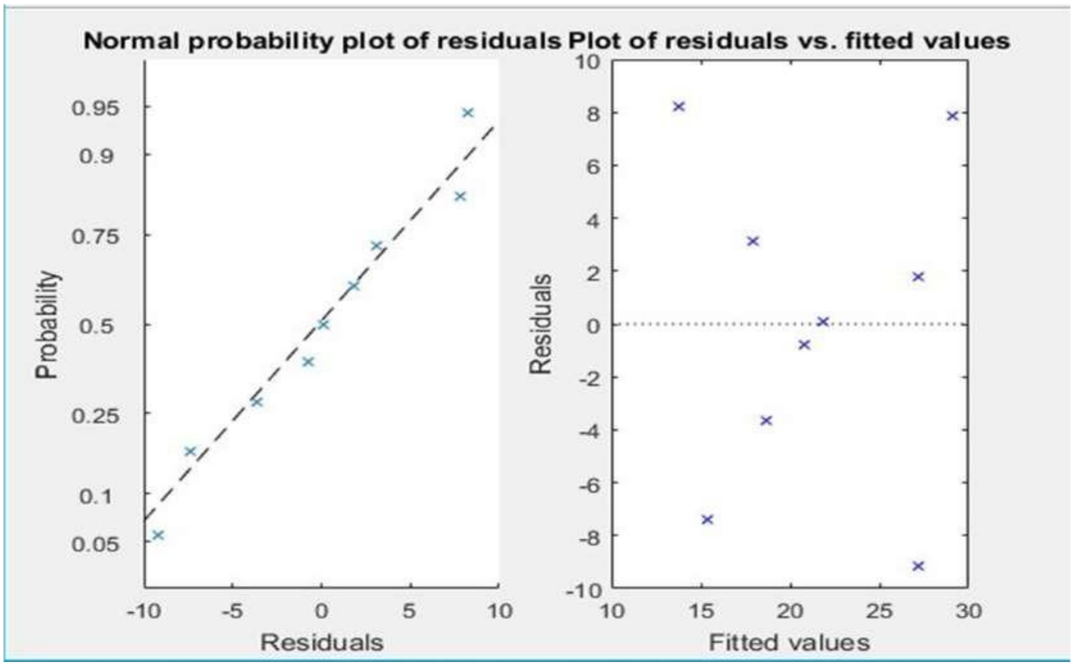
Screenshot No-8: KMP Murder Figure



Screenshot No-9: RMP Murder Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    ______    ________    ________

    (Intercept)      10827     6400.9      1.6914     0.15154
    Year           -5.4364     3.2267     -1.6848     0.15284
    Ur              20.862     13.747      1.5176     0.18957
    Gdp              8.592     8.9204     0.96319     0.37969


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 7.68
R-squared: 0.449,  Adjusted R-Squared 0.119
F-statistic vs. constant model: 1.36, p-value = 0.356

Correlation :

    1.0000   -0.4405   -0.0852   -0.3578
   -0.4405    1.0000    0.7651    0.8572
   -0.0852    0.7651    1.0000    0.4661
   -0.3578    0.8572    0.4661    1.0000
```

Screenshot No-10: RMP Murder Figure



Screenshot No-11: SMP Murder Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    ______    ________    ________

    (Intercept)      1208.2    3749.2     0.32225      0.7603
    Year           -0.59186    1.8899    -0.31316      0.7668
    Ur               5.7234    8.0519     0.71082     0.50897
    Gdp            -0.61934    5.2249    -0.11854     0.91026


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 4.5
R-squared: 0.162,  Adjusted R-Squared -0.34
F-statistic vs. constant model: 0.323, p-value = 0.81

Correlation :

    1.0000   -0.0939    0.1784   -0.2155
   -0.0939    1.0000    0.7651    0.8572
    0.1784    0.7651    1.0000    0.4661
   -0.2155    0.8572    0.4661    1.0000
```

Screenshot No-12: SMP Murder Figure

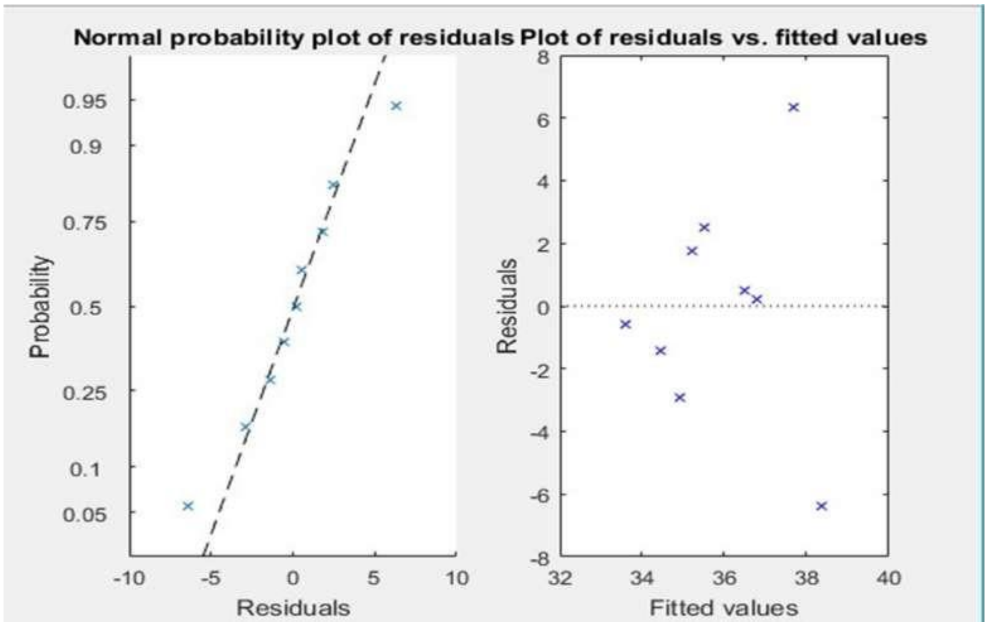


Table 4. 1 Murder Prediction of Metropolitan Area

Actual Value				Actual Value		
year	BMP	Predicted value		year	RMP	Predicted value
2010	19	17		2010	21	18
2011	15	18		2011	18	27
2012	20	17		2012	37	29
2013	17	16		2013	29	27
2014	15	16		2014	22	22
2015	15	16		2015	20	21
2016	16	16		2016	15	19
2017	16	16		2017	8	15
2018	17	16		2018	22	14
2019		15		2019		15
2020		15		2020		13
Actual Value				Actual Value		
year	KMP	Predicted Value		year	SMP	Predicted Value
2010	29	27		2010	33	34
2011	17	21		2011	37	35
2012	23	22		2012	37	37
2013	31	28		2013	32	38
2014	22	30		2014	44	38
2015	30	28		2015	37	37
2016	36	26		2016	38	36
2017	25	26		2017	32	35
2018	19	23		2018	33	34
2019		27		2019		35
2020		27		2020		35
Actual Value				Actual Value		
year	DMP	Predicted Value		year	CMP	Predicted Value
2010	245	247		2010	94	96
2011	259	257		2011	83	81
2012	264	261		2012	88	87
2013	270	265		2013	95	108
2014	262	249		2014	120	109
2015	239	237		2015	105	99
2016	165	221		2016	90	86
2017	218	208		2017	70	84
2018	216	193		2018	75	71
2019		199		2019		83
2020		191		2020		81

4.1.2 Robbery Result

Screenshot No-13: BMP Robbery Coefficient

Estimated Coefficients:

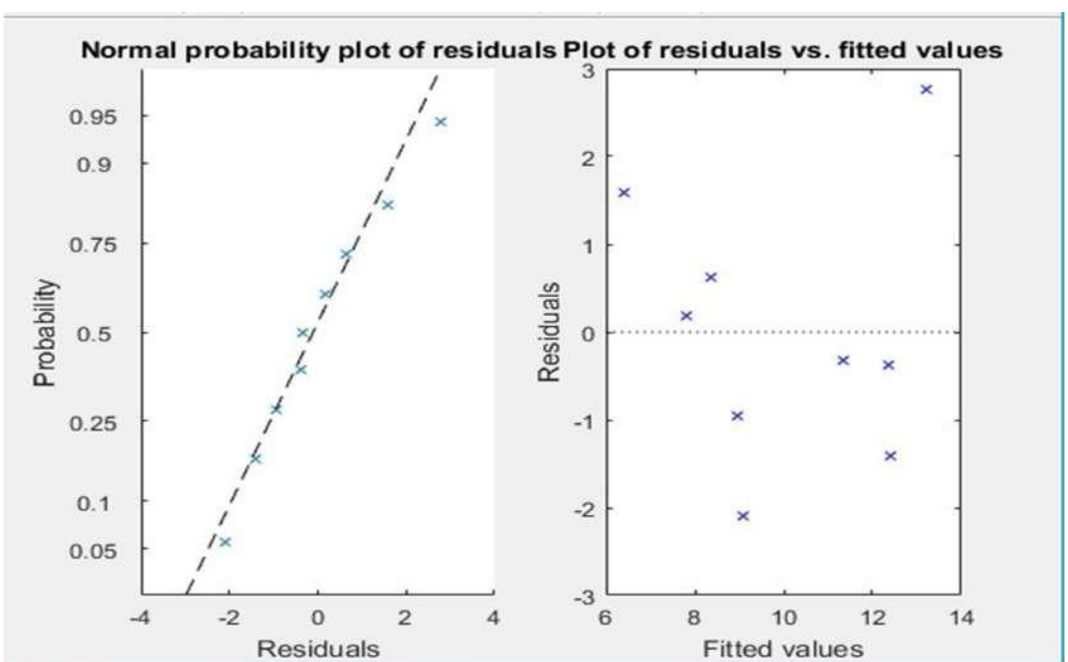
	Estimate	SE	tStat	pValue
(Intercept)	-2194	1586.8	-1.3827	0.22532
Year	1.1157	0.79988	1.3949	0.22185
Ur	-0.56075	3.4078	-0.16455	0.87574
Gdp	-6.1615	2.2113	-2.7863	0.038613

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 1.9
R-squared: 0.717, Adjusted R-Squared 0.546
F-statistic vs. constant model: 4.21, p-value = 0.0778

Correlation :

1.0000	-0.3227	0.0203	-0.6792
-0.3227	1.0000	0.7651	0.8572
0.0203	0.7651	1.0000	0.4661
-0.6792	0.8572	0.4661	1.0000

Screenshot No-14: BMP Robbery Figure



Screenshot No-15: CMP Robbery Coefficient

Estimated Coefficients:

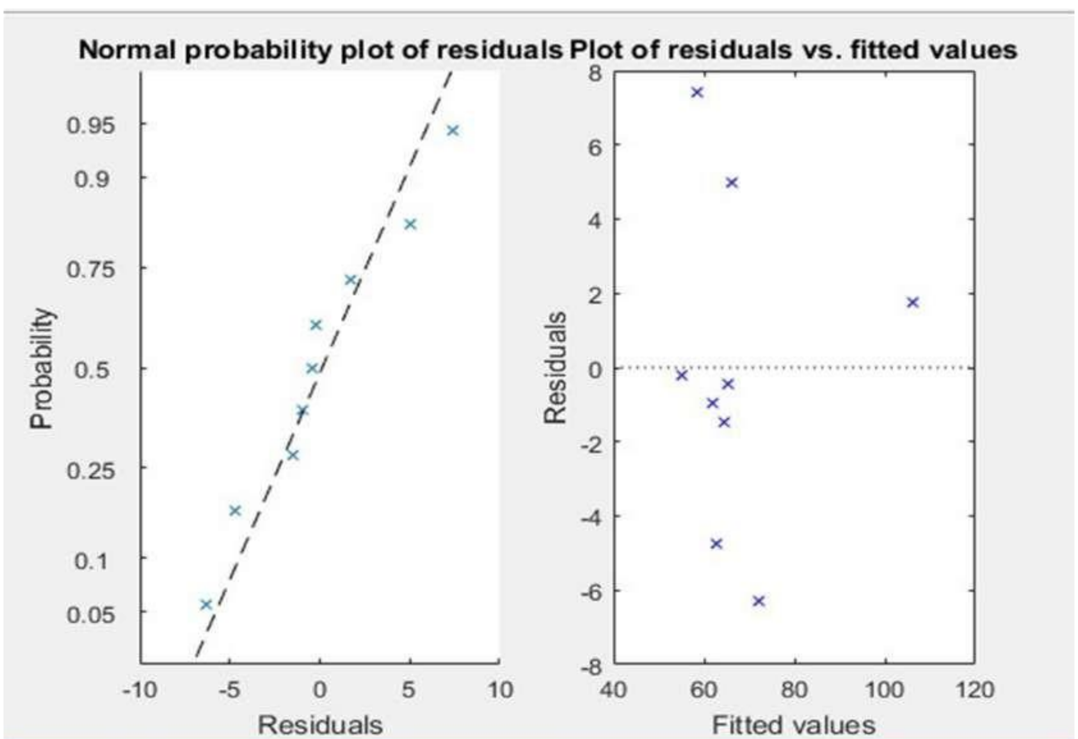
	Estimate	SE	tStat	pValue
(Intercept)	-18526	4539.2	-4.0814	0.0095264
Year	9.4463	2.2882	4.1283	0.0090998
Ur	-65.855	9.7485	-6.7554	0.0010793
Gdp	-23.688	6.3258	-3.7447	0.013367

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 5.45
R-squared: 0.925, Adjusted R-Squared 0.879
F-statistic vs. constant model: 20.4, p-value = 0.00309

Correlation :

1.0000	-0.4859	-0.8165	-0.4072
-0.4859	1.0000	0.7651	0.8572
-0.8165	0.7651	1.0000	0.4661
-0.4072	0.8572	0.4661	1.0000

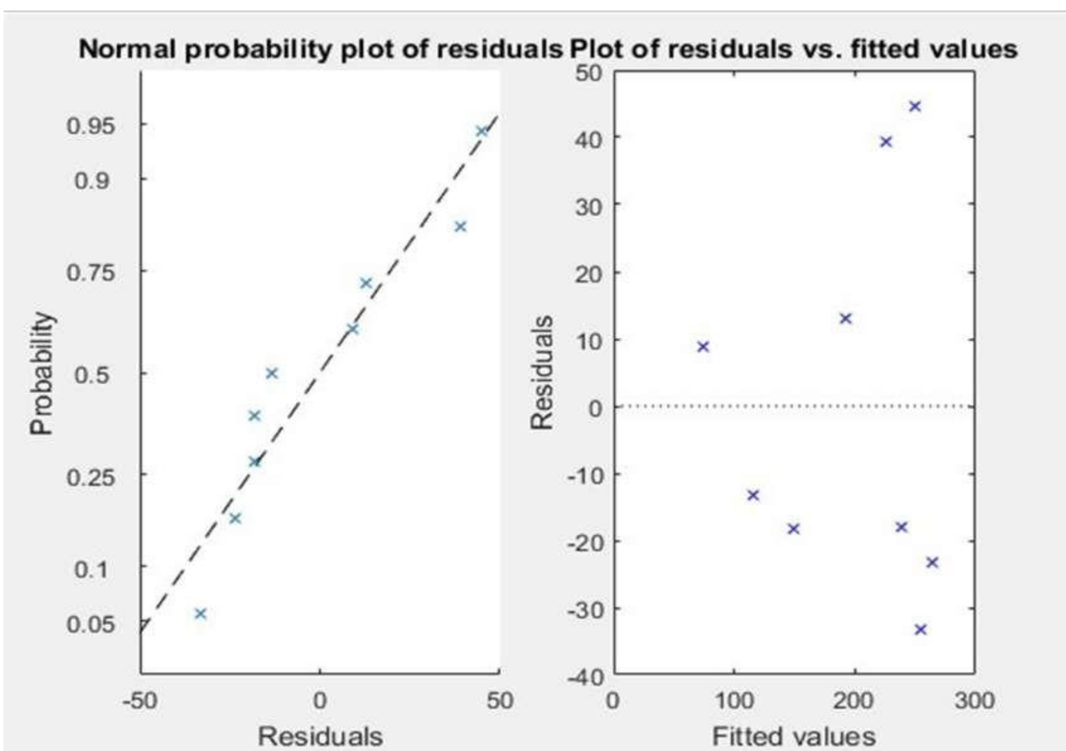
Screenshot No-16: CMP Robbery Figure



Screenshot No-17: DMP Robbery Coefficient

```
Estimated Coefficients:
                    Estimate      SE       tStat       pValue
                    ________    ______    ________    ________

    (Intercept)       74091      29501      2.5115     0.05373
    Year            -36.974     14.871     -2.4863    0.055412
    Ur               129.02     63.357      2.0364    0.097313
    Gdp              5.1871     41.113     0.12617     0.90452


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 35.4
R-squared: 0.855,  Adjusted R-Squared 0.769
F-statistic vs. constant model: 9.86, p-value = 0.0153

Correlation :

    1.0000   -0.8301   -0.3728   -0.8225
   -0.8301    1.0000    0.7651    0.8572
   -0.3728    0.7651    1.0000    0.4661
   -0.8225    0.8572    0.4661    1.0000
```

Screenshot No-18: DMP Robbery Figure



Screenshot No-19: KMP Robbery Coefficient

Estimated Coefficients:

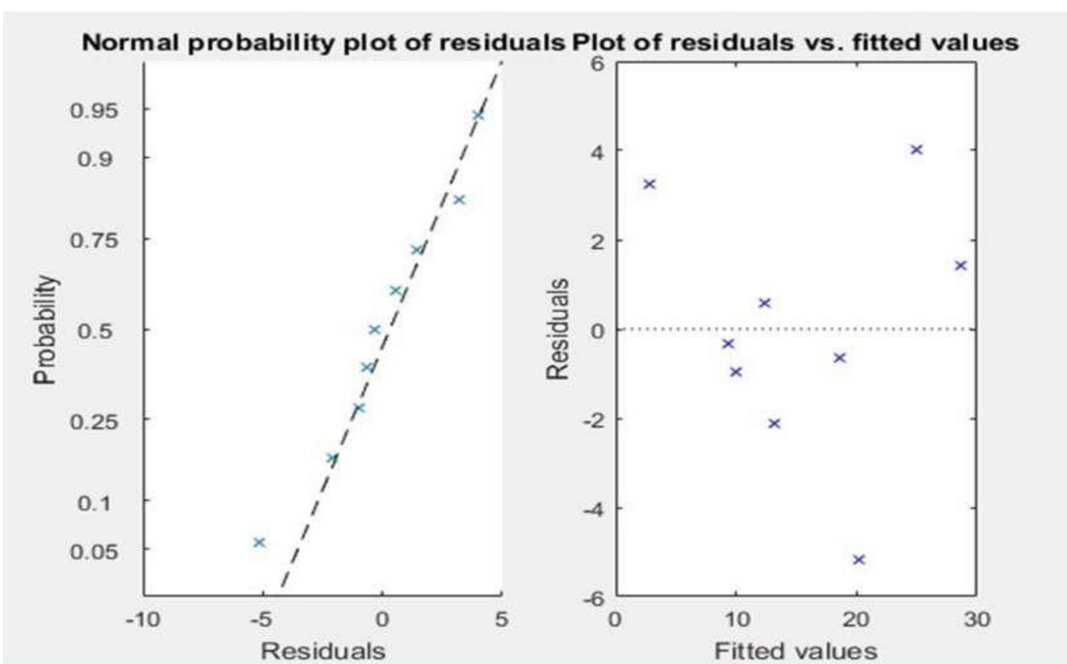
	Estimate	SE	tStat	pValue
(Intercept)	5866.5	2929.3	2.0027	0.10159
Year	-2.9507	1.4767	-1.9983	0.10217
Ur	28.545	6.2912	4.5373	0.0061841
Gdp	-4.0611	4.0824	-0.9948	0.36551

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 3.52
R-squared: 0.897, Adjusted R-Squared 0.835
F-statistic vs. constant model: 14.5, p-value = 0.00668

Correlation :

1.0000	-0.2793	0.3618	-0.5630
-0.2793	1.0000	0.7651	0.8572
0.3618	0.7651	1.0000	0.4661
-0.5630	0.8572	0.4661	1.0000

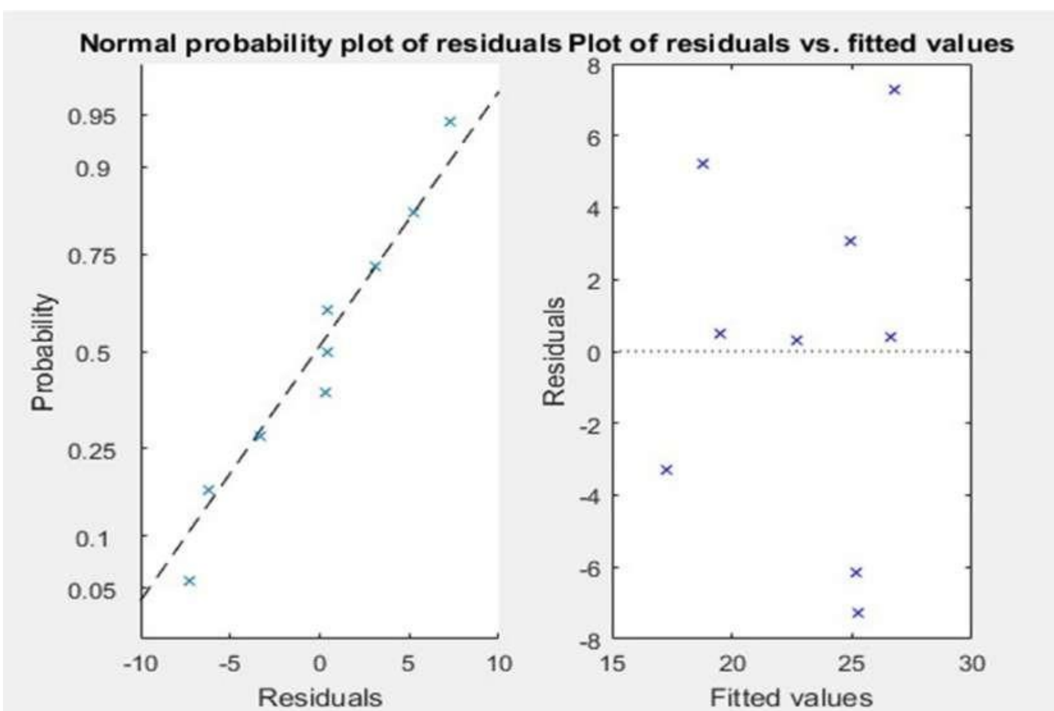
Screenshot No-20: KMP Robbery Figure



Screenshot No-21: RMP Robbery Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    ______    ________    ________

    (Intercept)     -4796.7    5152.2      -0.931     0.39458
    Year             2.4118    2.5972     0.92863      0.3957
    Ur              -2.0047    11.065    -0.18118     0.86334
    Gdp             -4.4414    7.1802    -0.61855      0.5633


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 6.18
R-squared: 0.349,  Adjusted R-Squared -0.0409
F-statistic vs. constant model: 0.895, p-value = 0.505

Correlation :

    1.0000    0.5421    0.4637    0.3481
    0.5421    1.0000    0.7651    0.8572
    0.4637    0.7651    1.0000    0.4661
    0.3481    0.8572    0.4661    1.0000
```

Screenshot No-22: RMP Robbery Figure



Screenshot No-23: SMP Robbery Coefficient

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	-4587.6	5044.1	-0.9095	0.40479
Year	2.351	2.5427	0.92461	0.3976
Ur	-9.0102	10.833	-0.83174	0.44345
Gdp	-13.261	7.0295	-1.8864	0.1179

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 6.06
R-squared: 0.608, Adjusted R-Squared 0.373
F-statistic vs. constant model: 2.58, p-value = 0.166

Correlation :

1.0000	-0.5672	-0.3795	-0.7342
-0.5672	1.0000	0.7651	0.8572
-0.3795	0.7651	1.0000	0.4661
-0.7342	0.8572	0.4661	1.0000

Screenshot No-24: SMP Robbery Figure

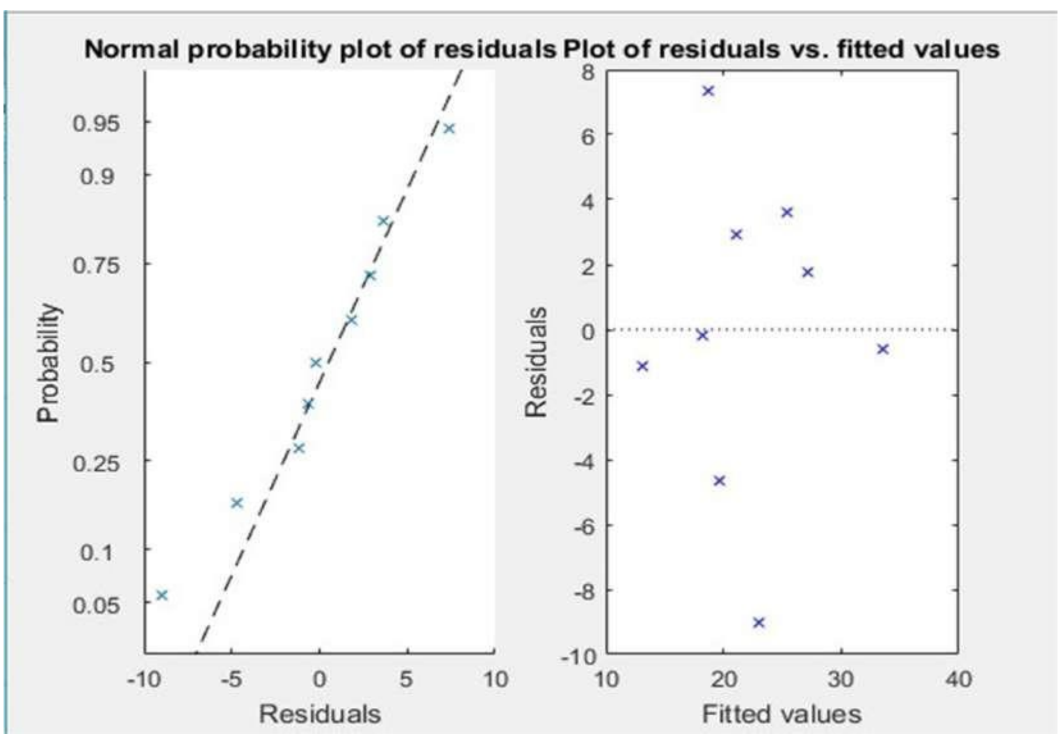


Table 4. 2 Robbery Prediction of Metropolitan Area

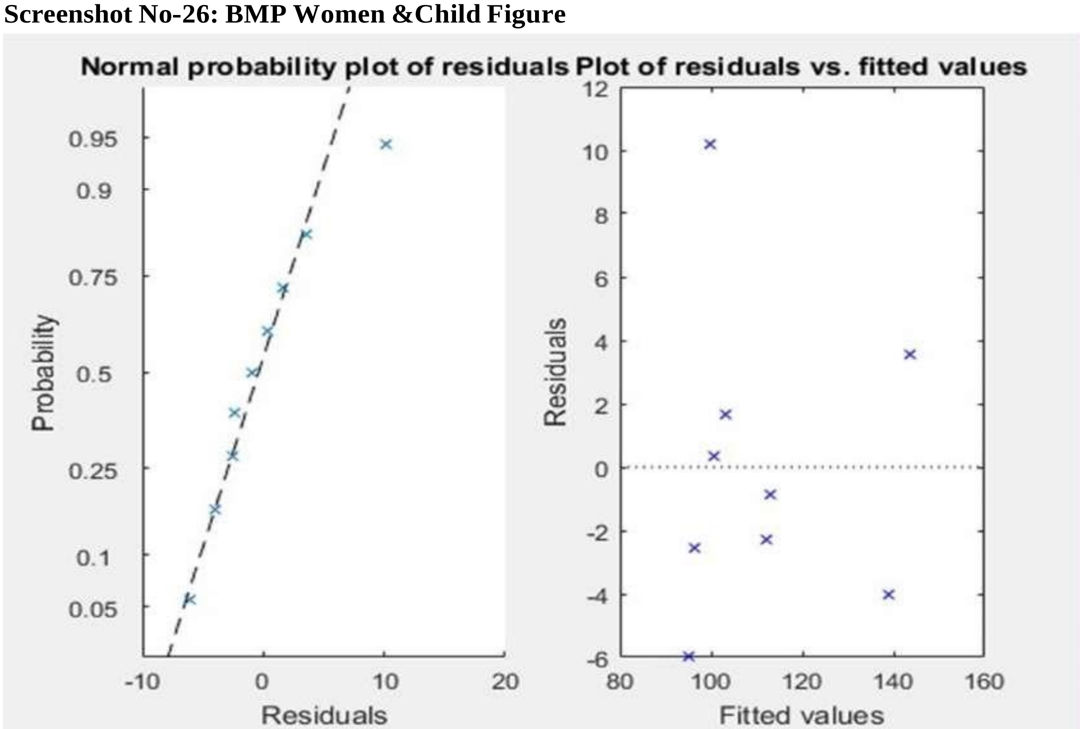
	Actual Value				Actual Value	
YEAR	BMP	Predicted Value		YEAR	CMP	Predicted Value
2010	12	12		2010	108	106
2011	8	8		2011	66	72
2012	9	8		2012	66	59
2013	11	12		2013	55	55
2014	16	13		2014	63	64
2015	11	11		2015	61	62
2016	7	9		2016	58	63
2017	8	9		2017	71	66
2018	8	6		2018	65	65
2019		8		2019		54
2020		8		2020		51
	Actual Value				Actual Value	
YEAR	DMP	Predicted Value		YEAR	KMP	Predicted Value
2010	220	238		2010	9	9
2011	294	249		2011	13	12
2012	222	255		2012	18	19
2013	241	264		2013	30	29
2014	265	226		2014	29	25
2015	205	192		2015	15	20
2016	131	149		2016	11	13
2017	103	116		2017	9	10
2018	83	74		2018	6	3
2019		84		2019		11
2020		62		2020		10
	Actual Value				Actual Value	
YEAR	RMP	Predicted Value		YEAR	SMP	Predicted Value
2010	20	20		2010	33	34
2011	14	17		2011	24	21
2012	24	19		2012	15	20
2013	23	23		2013	29	26
2014	28	25		2014	29	27
2015	19	25		2015	14	23
2016	18	25		2016	26	19
2017	34	27		2017	18	18
2018	27	27		2018	12	13
2019		29		2019		14
2020		30		2020		12

4.1.3 Women and Child Repression Result

Screenshot No-25: BMP Women & Child Coefficient

```
Estimated Coefficients:
                    Estimate      SE       tStat       pValue
                    ________    ______    _______    _________

    (Intercept)       34720     5057.9     6.8645    0.0010029
    Year            -17.387     2.5496    -6.8193    0.0010338
    Ur               31.727     10.863     2.9208     0.032984
    Gdp              41.818     7.0488     5.9327    0.0019417


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 6.07
R-squared: 0.934,  Adjusted R-Squared 0.894
F-statistic vs. constant model: 23.5, p-value = 0.00225

Correlation :

    1.0000   -0.6804   -0.5634   -0.2752
   -0.6804    1.0000    0.7651    0.8572
   -0.5634    0.7651    1.0000    0.4661
   -0.2752    0.8572    0.4661    1.0000
```

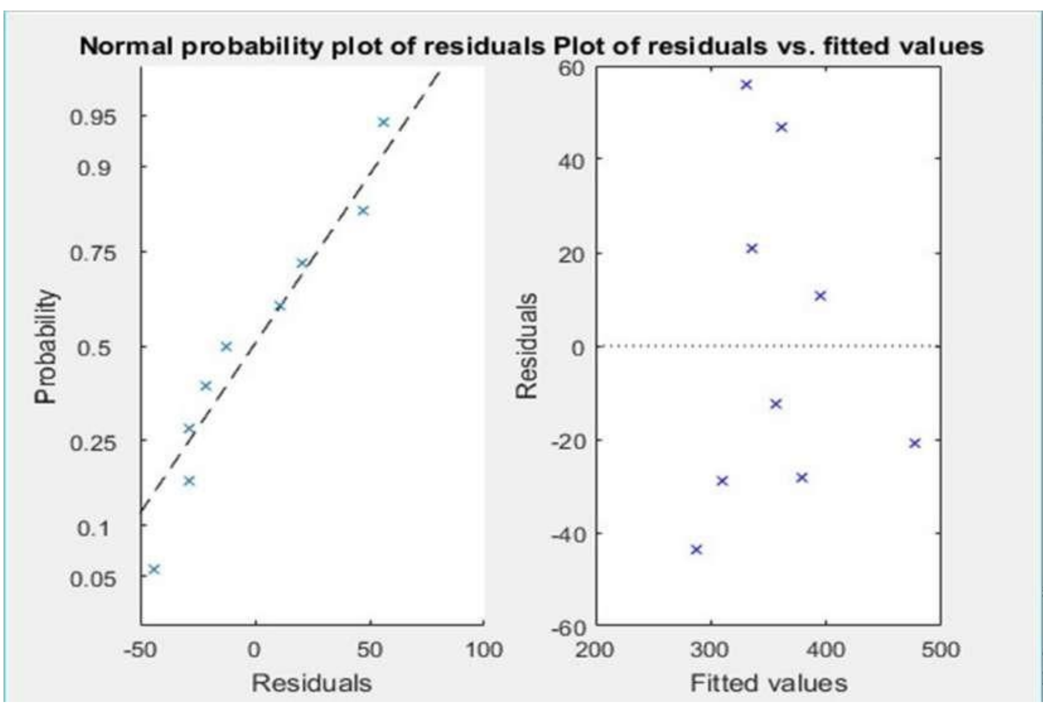

Screenshot No-26: BMP Women &Child Figure



Screenshot No-27: CMP Women & Child Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    _______    _______    ________

    (Intercept)     -84093      37327     -2.2529    0.074011
    Year            42.734     18.816      2.2711    0.072337
    Ur             -271.25     80.165     -3.3837    0.019596
    Gdp             -73.69     52.019     -1.4166     0.21578


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 44.8
R-squared: 0.706,  Adjusted R-Squared 0.529
F-statistic vs. constant model: 4, p-value = 0.0849

Correlation :

    1.0000   -0.0943   -0.5621   -0.0018
   -0.0943    1.0000    0.7651    0.8572
   -0.5621    0.7651    1.0000    0.4661
   -0.0018    0.8572    0.4661    1.0000
```

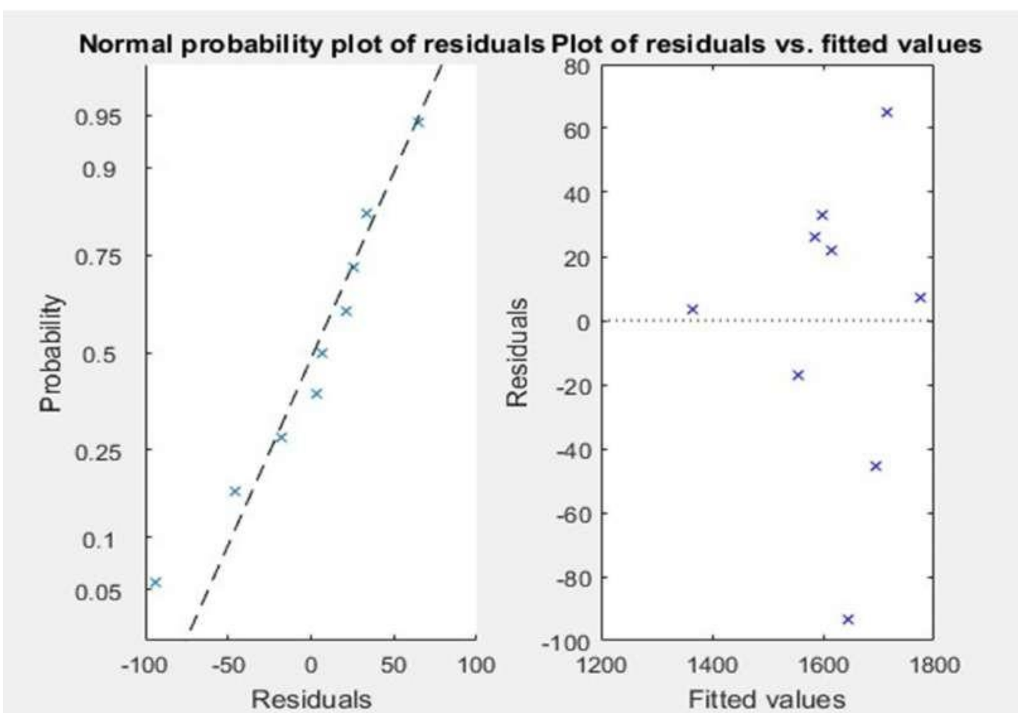
Screenshot No-28: CMP Women & Child Figure



Screenshot No-29: DMP Women & Child Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat        pValue
                   ________    ______    ________    ________

    (Intercept)      35172      49447     0.71131     0.50869
    Year           -17.591     24.926    -0.70575     0.51186
    Ur              207.27     106.19      1.9518     0.10841
    Gdp             153.14      68.91      2.2223    0.076908


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 59.4
R-squared: 0.861,  Adjusted R-Squared 0.778
F-statistic vs. constant model: 10.3, p-value = 0.0139

Correlation :

    1.0000    0.8397    0.7315    0.8354
    0.8397    1.0000    0.7651    0.8572
    0.7315    0.7651    1.0000    0.4661
    0.8354    0.8572    0.4661    1.0000
```

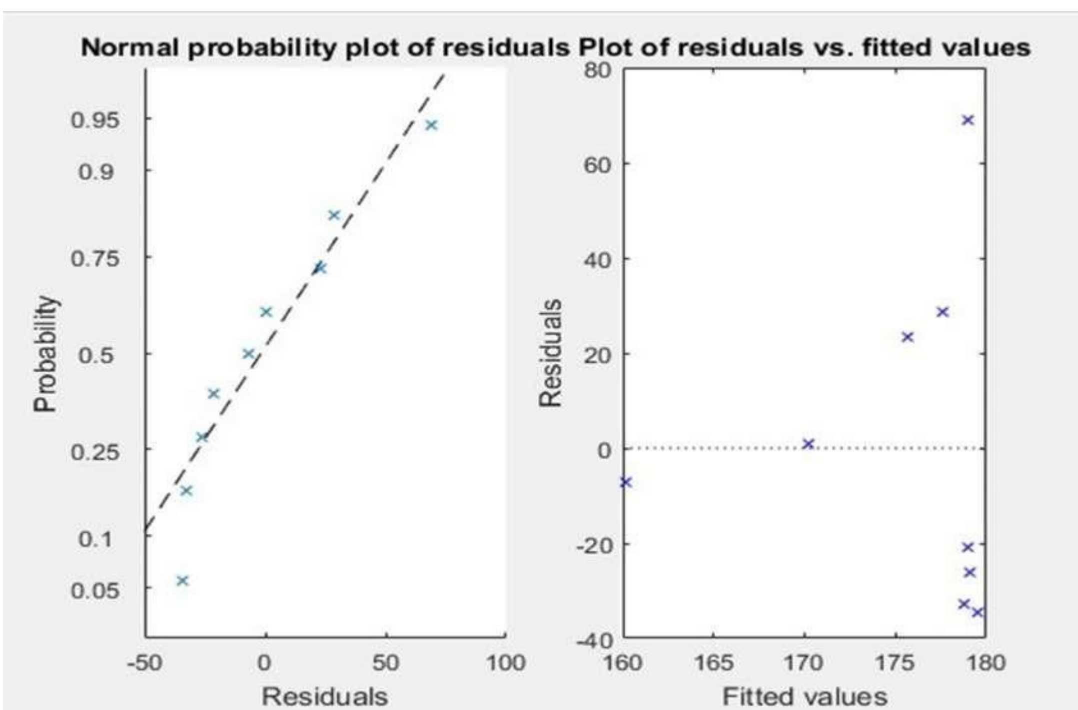
Screenshot No-30: DMP Women & Child Figure



Screenshot No-31: KMP Women & Child Coefficient

```
Estimated Coefficients:
                   Estimate      SE         tStat        pValue
                   ________    ______    __________    _______

    (Intercept)       3005      36401      0.082553    0.93741
    Year           -1.4638      18.35     -0.079771    0.93951
    Ur              19.969     78.176       0.25543    0.80856
    Gdp             5.3605     50.729       0.10567    0.91995


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 43.7
R-squared: 0.0335,  Adjusted R-Squared -0.546
F-statistic vs. constant model: 0.0578, p-value = 0.98

Correlation :

    1.0000    0.1429    0.1766    0.1113
    0.1429    1.0000    0.7651    0.8572
    0.1766    0.7651    1.0000    0.4661
    0.1113    0.8572    0.4661    1.0000
```

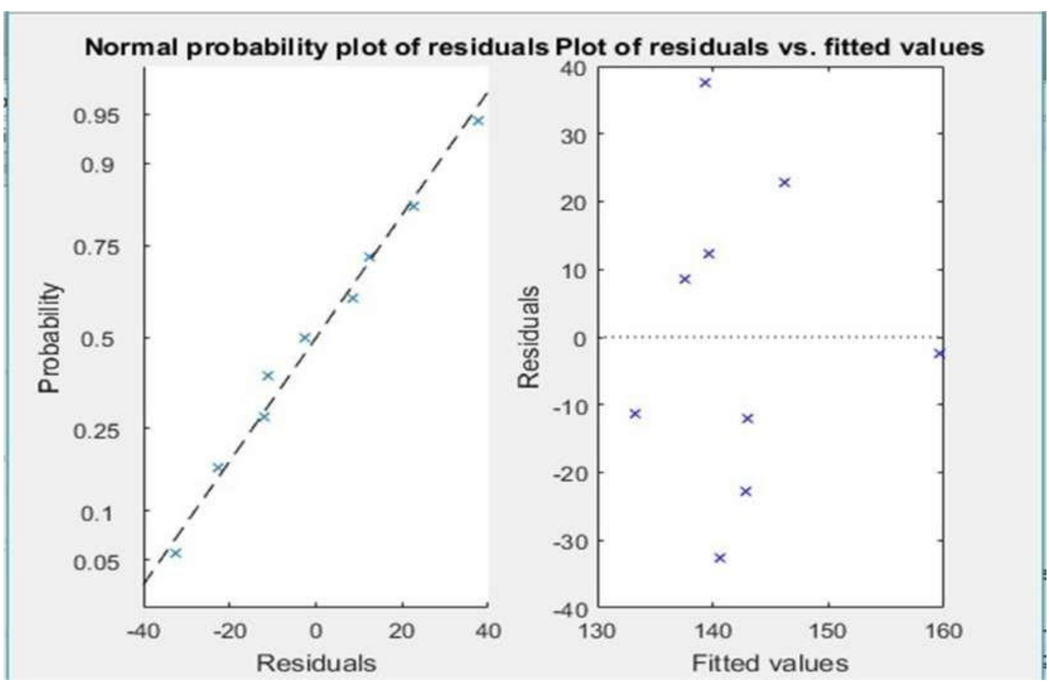
Screenshot No-32: KMP Women & Child Figure



Screenshot No-33: RMP Women & Child Coefficient

```
Estimated Coefficients:
                   Estimate      SE        tStat       pValue
                   ________    ______    ________    _______

    (Intercept)     -14069      23617     -0.5957    0.57731
    Year            7.1888     11.905     0.60384     0.5723
    Ur             -30.748     50.721    -0.60623    0.57083
    Gdp            -21.084     32.913    -0.64059    0.54999


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 28.4
R-squared: 0.0977,  Adjusted R-Squared -0.444
F-statistic vs. constant model: 0.181, p-value = 0.905

Correlation :

    1.0000   -0.0909   -0.1495   -0.1563
   -0.0909    1.0000    0.7651    0.8572
   -0.1495    0.7651    1.0000    0.4661
   -0.1563    0.8572    0.4661    1.0000
```

Screenshot No-34: RMP Women & Child Figure



Screenshot No-35: SMP Women & Child Coefficient

```
Estimated Coefficients:
                    Estimate      SE       tStat      pValue
                    ________    ______    _______    _______

    (Intercept)       13378     34894     0.38338     0.7172
    Year            -6.7651     17.59     -0.3846    0.71635
    Ur               46.198    74.941     0.61646    0.56457
    Gdp               31.03    48.629      0.6381    0.55148


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 41.9
R-squared: 0.216,  Adjusted R-Squared -0.254
F-statistic vs. constant model: 0.46, p-value = 0.723

Correlation :

    1.0000    0.3710    0.3620    0.3889
    0.3710    1.0000    0.7651    0.8572
    0.3620    0.7651    1.0000    0.4661
    0.3889    0.8572    0.4661    1.0000
```

Screenshot No-36: SMP Women & Child Figure

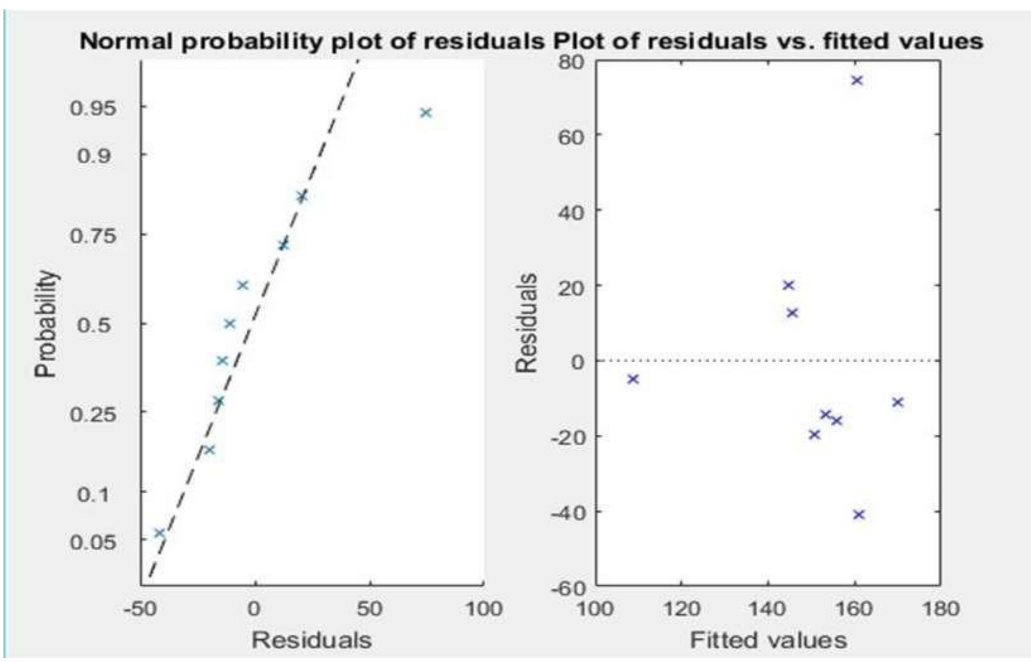


Table 4. 3 Women & Child Repression Prediction of Metropolitan Area

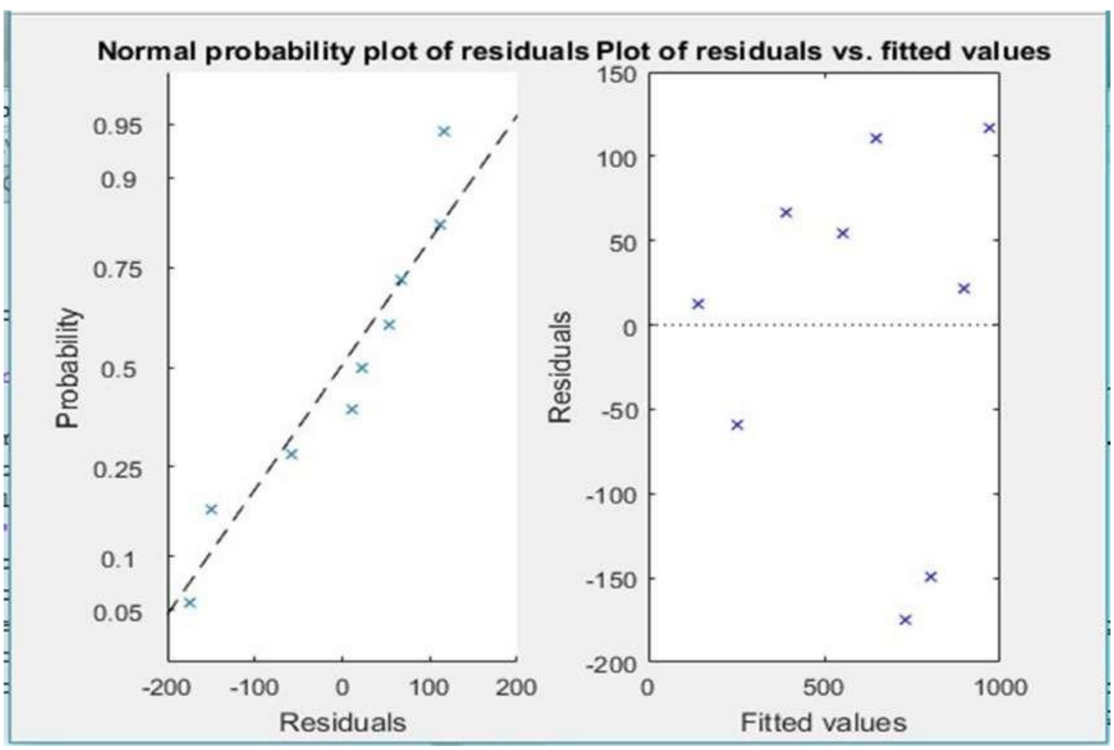
Actual value				Actual value		
year	DMP	Predicted Value		year	CMP	Predicted Value
2010	1370	1366		2010	455	476
2011	1538	1555		2011	408	361
2012	1637	1615		2012	281	310
2013	1631	1598		2013	244	288
2014	1611	1585		2014	387	331
2015	1550	1643		2015	357	336
2016	1651	1696		2016	344	356
2017	1779	1714		2017	350	378
2018	1782	1775		2018	405	394
2019		1810		2019		348
2020		1884		2020		346
Actual value				Actual value		
year	KMP	Predicted Value		year	RMP	Predicted Value
2010	153	160		2010	157	160
2011	171	170		2011	146	138
2012	199	176		2012	122	133
2013	153	179		2013	152	140
2014	206	178		2014	169	146
2015	146	179		2015	131	143
2016	158	179		2016	108	141
2017	248	179		2017	120	143
2018	145	179		2018	177	139
2019		185		2019		139
2020		186		2020		138
Actual value				Actual value		
year	BMP	Predicted Value		year	SMP	Predicted Value
2010	112	113		2010	104	109
2011	147	143		2011	158	146
2012	135	139		2012	140	156
2013	110	112		2013	131	151
2014	94	97		2014	165	145
2015	110	100		2015	139	153
2016	105	103		2016	235	161
2017	89	95		2017	120	161
2018	101	101		2018	159	170
2019		88		2019		175
2020		83		2020		181

4.1.4 Narcotics Result

Screenshot No-37: BMP Narcotics Coefficient

```
Estimated Coefficients:
                    Estimate        SE          tStat       pValue
                   __________    __________    ________    ________

    (Intercept)    -1.9781e+05   1.1197e+05    -1.7666     0.13755
    Year               98.365        56.446     1.7426     0.14186
    Ur                  123.3        240.48    0.51273     0.62998
    Gdp               -31.482        156.05   -0.20174     0.84807


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 134
R-squared: 0.88,  Adjusted R-Squared 0.808
F-statistic vs. constant model: 12.2, p-value = 0.00973

Correlation :

    1.0000    0.9300    0.7879    0.7490
    0.9300    1.0000    0.7651    0.8572
    0.7879    0.7651    1.0000    0.4661
    0.7490    0.8572    0.4661    1.0000
```

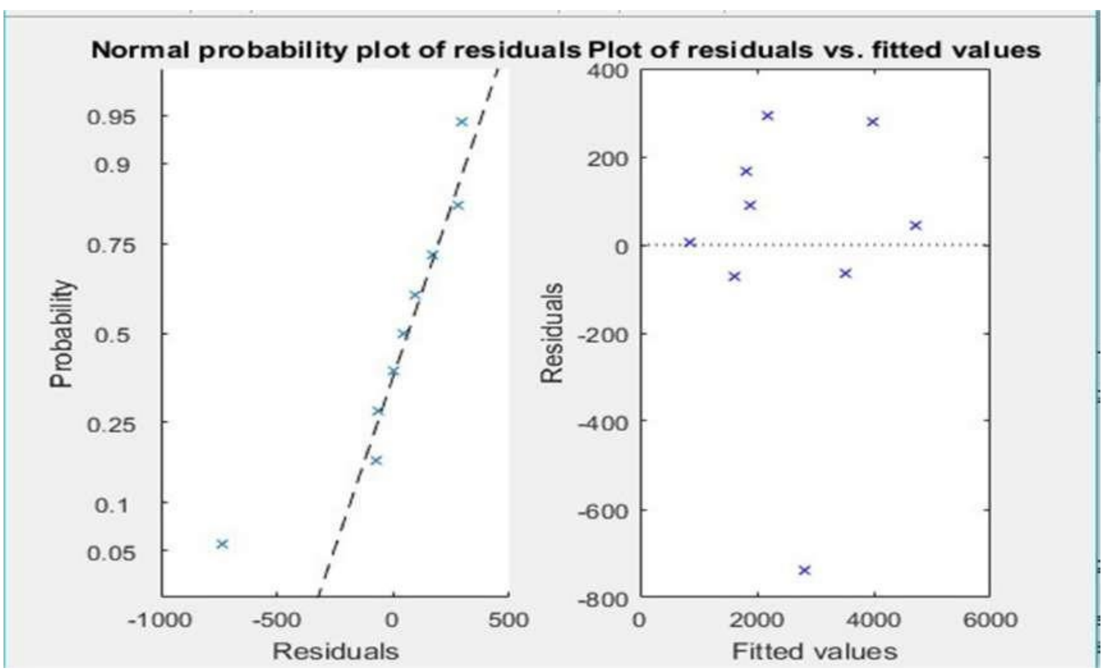

Screenshot No-38: BMP Narcotics Figure



Screenshot No-39: CMP Narcotics Coefficient

```
Estimated Coefficients:
                    Estimate        SE          tStat       pValue
                   __________   __________   _________   ________

    (Intercept)    -7.052e+05   3.2409e+05    -2.1759    0.081531
    Year               350.24       163.37     2.1438    0.084911
    Ur                -340.75       696.04   -0.48956     0.64517
    Gdp                581.91       451.66     1.2884     0.25401


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 389
R-squared: 0.943,  Adjusted R-Squared 0.909
F-statistic vs. constant model: 27.6, p-value = 0.00155

Correlation :

    1.0000    0.9479    0.6224    0.9177
    0.9479    1.0000    0.7651    0.8572
    0.6224    0.7651    1.0000    0.4661
    0.9177    0.8572    0.4661    1.0000
```

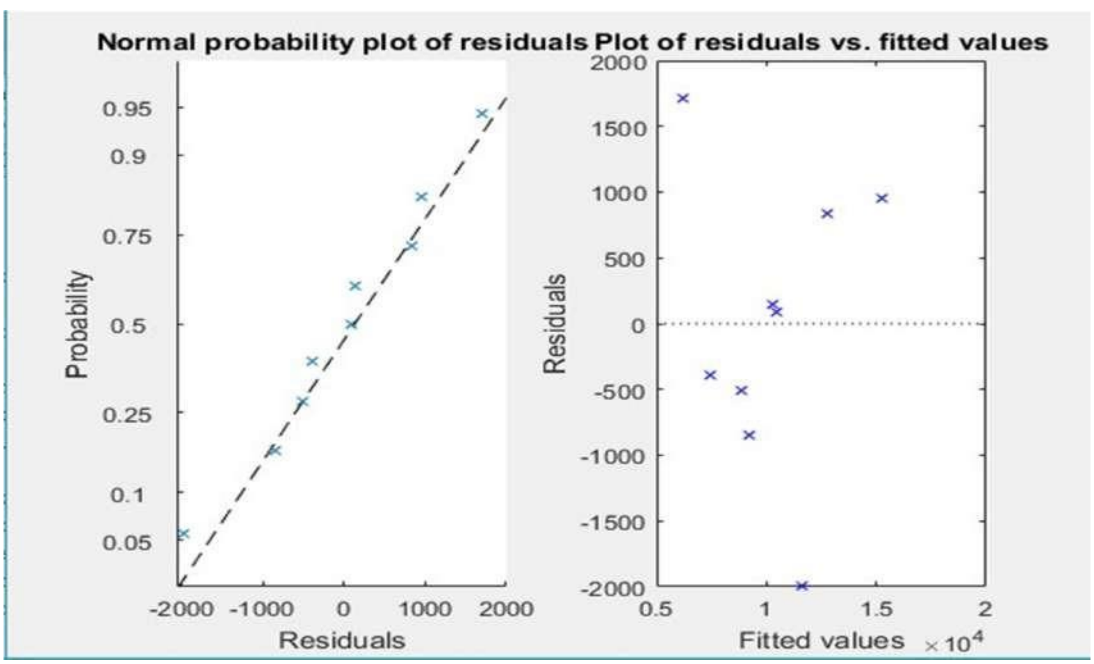
Screenshot No-40: CMP Narcotics Figure



Screenshot No-41: DMP Narcotics Coefficient

```
Estimated Coefficients:
                    Estimate        SE          tStat       pValue
                   __________    __________    ________    ________

    (Intercept)    -2.0733e+06    1.1581e+06    -1.7902     0.13343
    Year                1045.3        583.81     1.7905     0.13338
    Ur                   -7680        2487.3    -3.0877    0.027229
    Gdp                 1541.5          1614    0.95505     0.38341


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 1.39e+03
R-squared: 0.863,  Adjusted R-Squared 0.78
F-statistic vs. constant model: 10.5, p-value = 0.0136

Correlation :

    1.0000    0.5454   -0.0559    0.7509
    0.5454    1.0000    0.7651    0.8572
   -0.0559    0.7651    1.0000    0.4661
    0.7509    0.8572    0.4661    1.0000
```

Screenshot No-42: DMP Narcotics Figure



Screenshot No-43: KMP Narcotics Coefficient

Estimated Coefficients:

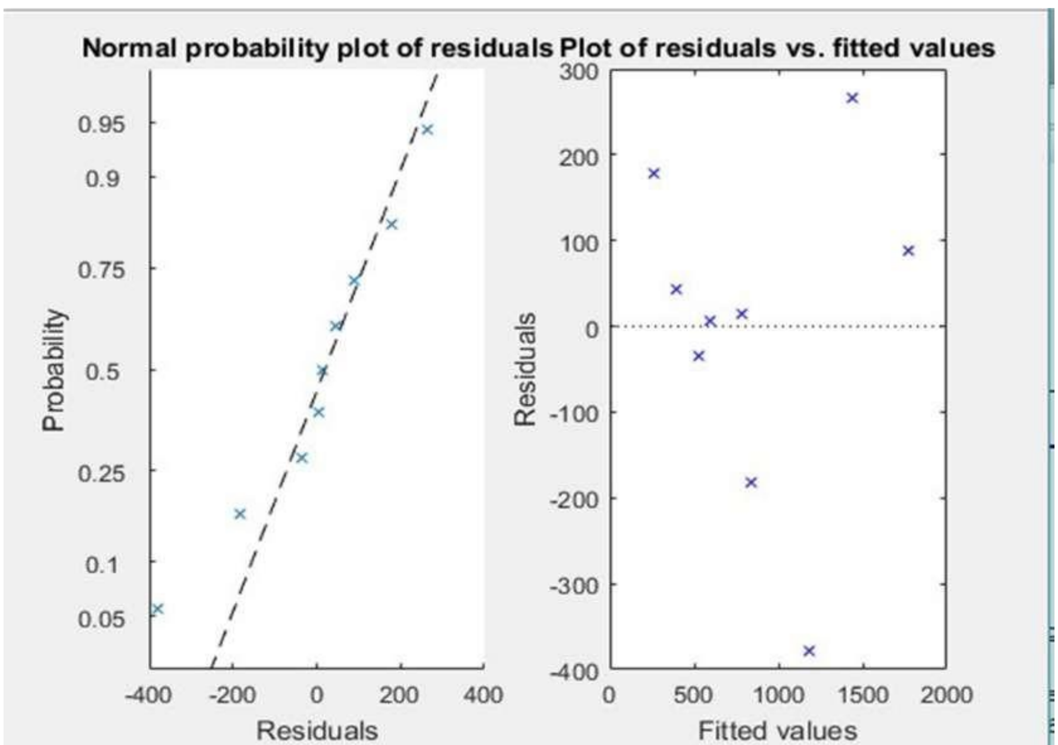
	Estimate	SE	tStat	pValue
(Intercept)	-6.2747e+05	2.006e+05	-3.1279	0.026019
Year	315.15	101.12	3.1166	0.026355
Ur	-1344	430.82	-3.1196	0.026263
Gdp	-120.94	279.56	-0.43262	0.68332

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 241
R-squared: 0.875, Adjusted R-Squared 0.8
F-statistic vs. constant model: 11.6, p-value = 0.0108

Correlation :

1.0000	0.7502	0.2171	0.7775
0.7502	1.0000	0.7651	0.8572
0.2171	0.7651	1.0000	0.4661
0.7775	0.8572	0.4661	1.0000

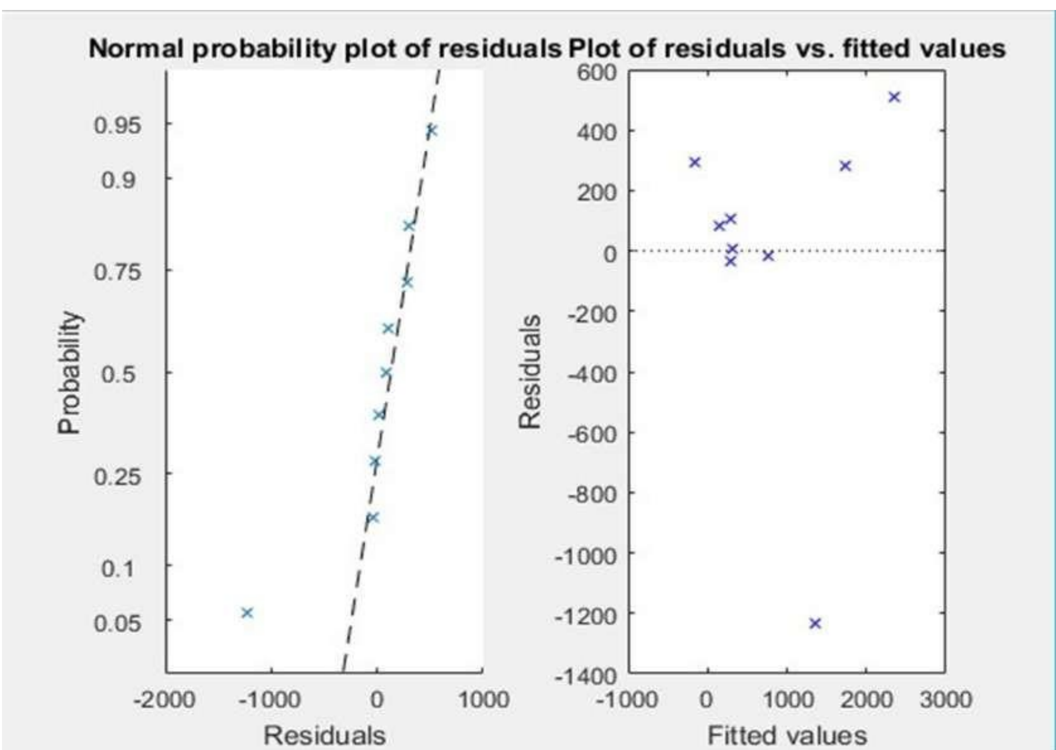
Screenshot No-44: KMP Narcotics Figure



Screenshot No-45: RMP Narcotics Coefficient

```
Estimated Coefficients:
                   Estimate        SE          tStat       pValue
                  __________    __________    ________    _______

    (Intercept)   -8.099e+05    5.2147e+05     -1.5531    0.18111
    Year              405.52        262.87      1.5427    0.18355
    Ur               -1673.5        1119.9     -1.4943    0.19534
    Gdp               139.41        726.73     0.19183    0.85542


Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 626
R-squared: 0.742,  Adjusted R-Squared 0.587
F-statistic vs. constant model: 4.79, p-value = 0.0622

Correlation :

    1.0000    0.7366    0.2775    0.7808
    0.7366    1.0000    0.7651    0.8572
    0.2775    0.7651    1.0000    0.4661
    0.7808    0.8572    0.4661    1.0000
```

Screenshot No-46: RMP Narcotics Figure



Screenshot No-47: SMP Narcotics Coefficient

Estimated Coefficients:

	Estimate	SE	tStat	pValue
(Intercept)	-72858	85852	-0.84865	0.43481
Year	36.451	43.278	0.84227	0.43806
Ur	-101.89	184.38	-0.55259	0.60434
Gdp	16.982	119.64	0.14194	0.89267

Number of observations: 9, Error degrees of freedom: 5
Root Mean Squared Error: 103
R-squared: 0.527, Adjusted R-Squared 0.244
F-statistic vs. constant model: 1.86, p-value = 0.254

Correlation :

1.0000	0.6848	0.3709	0.6754
0.6848	1.0000	0.7651	0.8572
0.3709	0.7651	1.0000	0.4661
0.6754	0.8572	0.4661	1.0000

Screenshot No-48: SMP Narcotics Figure

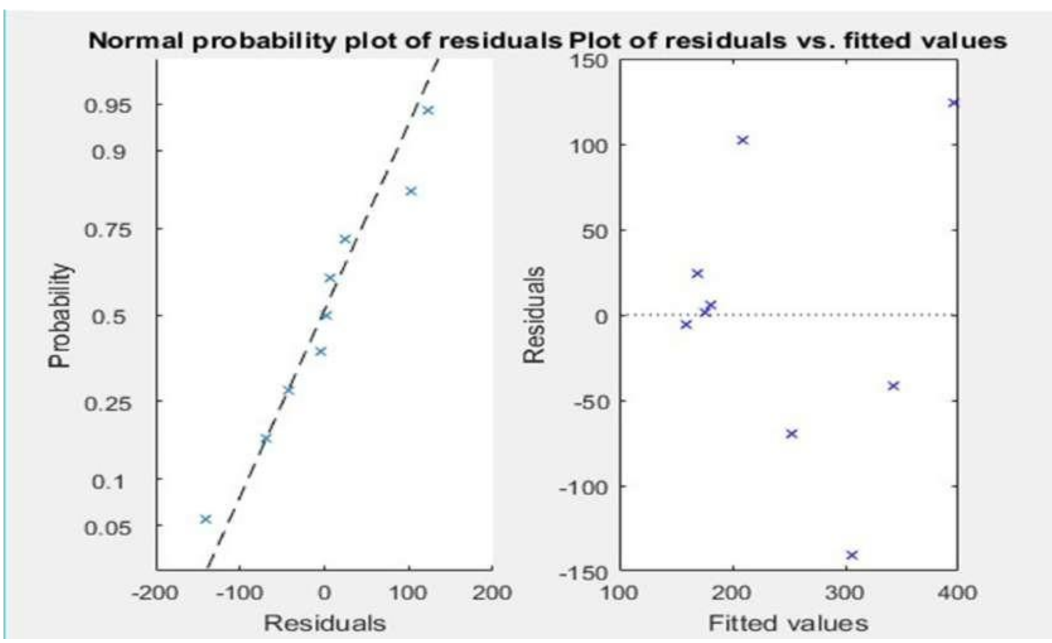


Table 4. 4 Narcotics Prediction of Metropolitan Area

Actual Value				Actual Value		
YEAR	BMP	Predicted Value		YEAR	CMP	Predicted Value
2010	155	143		2010	866	861
2011	196	255		2011	1541	1614
2012	459	392		2012	1976	1886
2013	607	553		2013	1977	1811
2014	759	648		2014	2488	2195
2015	557	732		2015	2090	2829
2016	655	805		2016	3457	3522
2017	921	899		2017	4260	3981
2018	1088	971		2018	4747	4702
2019		1120		2019		4830
2020		1225		2020		5276
Actual Value				Actual Value		
YEAR	DMP	Predicted Value		YEAR	KMP	Predicted Value
2010	10535	10442		2010	792	778
2011	10405	10256		2011	495	530
2012	8345	8860		2012	437	395
2013	7927	6216		2013	442	263
2014	7063	7453		2014	599	593
2015	8365	9215		2015	659	842
2016	9627	11615		2016	801	1179
2017	13638	12800		2017	1706	1440
2018	16215	15261		2018	1858	1769
2019		13187		2019		1602
2020		13778		2020		1749
Actual Value				Actual Value		
YEAR	RMP	Predicted Value		YEAR	SMP	Predicted Value
2010	332	323		2010	154	160
2011	253	280		2011	178	176
2012	229	147		2012	186	180
2013	144	-150		2013	194	170
2014	395	287		2014	311	208
2015	733	752		2015	183	253
2016	114	1344		2016	164	305
2017	2021	1741		2017	301	343
2018	2847	2338		2018	520	396
2019		2096		2019		392
2020		2358		2020		421

Chapter 5: Future Work and Conclusion

5.1 Future work

In future we will work with more data for more accuracy of Prediction, and we will work with the effect of independent variables.

5.2 Conclusion

This paper uses Multiple Linear Regression (MLR) model for predicting data in next two years. This study gives a deliberate methodology that helps police, detective agencies who are looking for an assumption about future crime.We got the value of prediction which is very closely related to the actual value. Prediction for 2019 and 2020 was dependent on the Unemployment rate (Ur) 4.68 and 4.78 also for the Gross domestic product (GDP) 7.736 and 7.961 using linear regression with two variables. The study of this paper will help the people to aware Metropolitan crime area.

Reference

[1] "https://www.police.gov.bd/storage/upload/announcement/lfE2cLvtf6LfV0e5h33wq2vpRJgeuXbHtReH124s.pdf".

[2] H. B. F. D. a. A. Suruliandi, "Survey on crime analysis and prediction using data mining techniques," *ICTACT JOURNAL ON SOFT COMPUTING,* vol. 07, no. 03, 2017.

[3] M. A. Awal, J. Rabbi and I. Rana, "Using Data Mining Technique to Analyze Crime of Bangladesh," *IJCSN - International Journal of Computer Science and Network,* vol. 6, no. 4, 2017.

[4] N. M. L. McClendon, "Using Machine Learning Algorithms to Analyze Crime Data," *Machine Learning and Applications: An International Journal (MLAIJ),* vol. 2, no. 1, 2015.

[5] S. V. Nath, "Crime Pattern Detection Using Data Mining," *International Conference on Web Intelligence and Intelligent Agent Technology,* pp. 41-44, 2006.

[6] L. T. S. U. S. Chainey, "The utility of hotspot mapping for predicting spatial patterns of crime," *Security Journal,* vol. 21, p. 4–28.

[7] T. K. C. W. A. K. J. L. J. K. J. S. D. Bruin, "Data mining approaches to criminal career analysis," in *the Sixth International Conference on Data Mining (ICDM'06)*, 2006.

[8] S. S. P. Thongtae, "An Analysis of Data Mining Applications in Crime Domain," in *IEEE International Conference on Computer and Information Technology Workshops*, 2006.

[9] M. W. W. M. M. W. D. C. H. Yu, "Crime Forecasting Using Data Mining Techniques," in *IEEE 11th International Conference on Data Mining Workshops (ICDMW' 11)*, 2011.

[10] " https://www.investopedia.com/terms/m/mlr.asp".

[11] "https://www.investopedia.com/terms/c/correlation.asp".

[12] "https://www.police.gov.bd/en/crime_statistic/year/2010".

[13] "https://www.police.gov.bd/en/crime_statistic/year/2011".

[14] "https://www.police.gov.bd/en/crime_statistic/year/2012".

[15] "https://www.police.gov.bd/en/crime_statistic/year/2013".

[16] "https://www.police.gov.bd/en/crime_statistic/year/2014".

[17] "https://www.police.gov.bd/en/crime_statistic/year/2015".

[18] "https://www.police.gov.bd/en/crime_statistic/year/2016".

[19] "https://www.police.gov.bd/en/crime_statistic/year/2017".

[20] "https://www.police.gov.bd/en/crime_statistic/year/2018".

[21] " https://tradingeconomics.com/bangladesh/gdp-growth".

[22] "https://www.ceicdata.com/en/indicator/bangladesh/unemployment-rate".

Appendix

```
%-------- Murder data -----

%--------BMP_MURDER--------------

clear all;

clc;

bmpTable=xlsread('BMP_MURDER.xlsx');

tbl  =  table(bmpTable(:,2),bmpTable(:,3),bmpTable(:,4),bmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','Murder'});

mdl = fitlm(tbl);

disp(mdl);

fprintf('\n');

correlation=corr(bmpTable);

disp('Correlation :');

fprintf('\n');

disp(correlation);

figure(1)

subplot(1,2,1),plotResiduals(mdl,'probability')

subplot(1,2,2),plotResiduals(mdl,'fitted')

%-------- CMP_MURDER---------

clear all;

clc;

cmpTable=xlsread('CMP_MURDER.xlsx');

tbl  =  table(cmpTable(:,2),cmpTable(:,3),cmpTable(:,4),cmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','Murder'});

mdl = fitlm(tbl);

disp(mdl);
```

```
fprintf('\n');
correlation=corr(cmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------- DMP_MURDER--------
clear all;
clc;
dmpTable=xlsread('DMP_MURDER.xlsx');
tbl  =  table(dmpTable(:,2),dmpTable(:,3),dmpTable(:,4),dmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Murder'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(dmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------- KMP_MURDER---------
clear all;
clc;
```

```
kmpTable=xlsread('KMP_MURDER.xlsx');
tbl  =  table(kmpTable(:,2),kmpTable(:,3),kmpTable(:,4),kmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Murder'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(kmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%--------RMP_MURDER  ---------
clear all;
clc;
rmpTable=xlsread('RMP_MURDER.xlsx');
tbl  =  table(rmpTable(:,2),rmpTable(:,3),rmpTable(:,4),rmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Murder'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(rmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
```

```
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')

%-------SMP_MURDER--------
clear all;
clc;
smpTable=xlsread('SMP_MURDER.xlsx');
tbl = table(smpTable(:,2),smpTable(:,3),smpTable(:,4),smpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Murder'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(smpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------NARCOTICS DATA------------
%--------BMP_NARCOTICS-----------
clear all;
clc;
bmpTable=xlsread('BMP_NARCOTICS.xlsx');
tbl = table(bmpTable(:,2),bmpTable(:,3),bmpTable(:,4),bmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Narcotics'});
mdl = fitlm(tbl);
```

```
disp(mdl);
fprintf('\n');
correlation=corr(bmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%------CMP_NARCOTICS---------
clear all;
clc;
cmpTable=xlsread('CMP_NARCOTICS.xlsx');
tbl = table(cmpTable(:,2),cmpTable(:,3),cmpTable(:,4),cmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Narcotics'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(cmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-----DMP_NARCOTICS----------
clear all;
clc;
```

```
dmpTable=xlsread('DMP_NARCOTICS.xlsx');
tbl = table(dmpTable(:,2),dmpTable(:,3),dmpTable(:,4),dmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Narcotics'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(dmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------KMP_NARCOTICS----------
clear all;
clc;
kmpTable=xlsread('KMP_NARCOTICS.xlsx');
tbl = table(kmpTable(:,2),kmpTable(:,3),kmpTable(:,4),kmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Narcotics'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(kmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
```

```
subplot(1,2,2),plotResiduals(mdl,'fitted')

%------RMP_NARCOTICS---------

clear all;

clc;

rmpTable=xlsread('RMP_NARCOTICS.xlsx');

tbl = table(rmpTable(:,2),rmpTable(:,3),rmpTable(:,4),rmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','Narcotics'});

mdl = fitlm(tbl);

disp(mdl);

fprintf('\n');

correlation=corr(rmpTable);

disp('Correlation :');

fprintf('\n');

disp(correlation);

figure(1)

subplot(1,2,1),plotResiduals(mdl,'probability')

subplot(1,2,2),plotResiduals(mdl,'fitted')

%--------SMP_NARCOTICS--------

clear all;

clc;

smpTable=xlsread('SMP_NARCOTICS.xlsx');

tbl = table(smpTable(:,2),smpTable(:,3),smpTable(:,4),smpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','Narcotics'});

mdl = fitlm(tbl);

disp(mdl);

fprintf('\n');

correlation=corr(smpTable);

disp('Correlation :');
```

```
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%.................ROBBERY............
%----------BMP_ROBBERY--------
clear all;
clc;
bmpTable=xlsread('BMP_ROBBERY.xlsx');
tbl = table(bmpTable(:,2),bmpTable(:,3),bmpTable(:,4),bmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Robbery'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(bmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%----------CMP_ROBBER---------
clear all;
clc;
cmpTable=xlsread('CMP_ROBBERY.xlsx');
tbl = table(cmpTable(:,2),cmpTable(:,3),cmpTable(:,4),cmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Robbery'});
```

```
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(cmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------DMP_ROBBERY-----------
clear all;
clc;
dmpTable=xlsread('DMP_ROBBERY.xlsx');
tbl = table(dmpTable(:,2),dmpTable(:,3),dmpTable(:,4),dmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Robbery'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(dmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-----------KMP_ROBBERY------------
```

```
clear all;
clc;
kmpTable=xlsread('KMP_ROBBERY.xlsx');
tbl = table(kmpTable(:,2),kmpTable(:,3),kmpTable(:,4),kmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Robbery'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(kmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-------------RMP_ROBBERY---------
clear all;
clc;
rmpTable=xlsread('RMP_ROBBERY.xlsx');
tbl = table(rmpTable(:,2),rmpTable(:,3),rmpTable(:,4),rmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','Robbery'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(rmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
```

```
figure(1)

subplot(1,2,1),plotResiduals(mdl,'probability')

subplot(1,2,2),plotResiduals(mdl,'fitted')

%------------SMP_ROBBERY------------

clear all;

clc;

smpTable=xlsread('SMP_ROBBERY.xlsx');

tbl = table(smpTable(:,2),smpTable(:,3),smpTable(:,4),smpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','Robbery'});

mdl = fitlm(tbl);

disp(mdl);

fprintf('\n');

correlation=corr(smpTable);

disp('Correlation :');

fprintf('\n');

disp(correlation);

figure(1)

subplot(1,2,1),plotResiduals(mdl,'probability')

subplot(1,2,2),plotResiduals(mdl,'fitted')

%-----------WOMEN AND CHILD REPRESSION----------

%-----------BMP_WOMEN-----------------------

clear all;

clc;

bmpTable=xlsread('BMP_WOMEN.xlsx');

tbl = table(bmpTable(:,2),bmpTable(:,3),bmpTable(:,4),bmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','WomenandChidRepression'});

mdl = fitlm(tbl);

disp(mdl);
```

```
fprintf('\n');
correlation=corr(bmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-----------CMP_WOMEN-------------------
clear all;
clc;
cmpTable=xlsread('CMP_WOMEN.xlsx');
tbl = table(cmpTable(:,2),cmpTable(:,3),cmpTable(:,4),cmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','WomenandChidRepression'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(cmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%----------FOR PREDECTION Demo Code for Women&ChildRepression CMP ----------
% WomenandChidRepression = cmpTable(:,1);
% year = cmpTable(:,2);
% Ur=cmpTable(:,3);
% Gdp=cmpTable(:,4);
```

```
% n=length(year);

% disp('Prediction for Women&ChildRepression Cmp');

% for i=1:n

%     y=648.77-.31417*year-1.1459*Ur+.81512*Gdp;

% end

% disp(y);

%............DMP_WOMEN.................

clear all;

clc;

dmpTable=xlsread('DMP_WOMEN.xlsx');

tbl  =  table(dmpTable(:,2),dmpTable(:,3),dmpTable(:,4),dmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','WomenandChidRepression'});

mdl = fitlm(tbl);

disp(mdl);

fprintf('\n');

correlation=corr(dmpTable);

disp('Correlation :');

fprintf('\n');

disp(correlation);

figure(1)

subplot(1,2,1),plotResiduals(mdl,'probability')

subplot(1,2,2),plotResiduals(mdl,'fitted')

%..............KMP_WOMEN................

clear all;

clc;

kmpTable=xlsread('KMP_WOMEN.xlsx');

tbl  =  table(kmpTable(:,2),kmpTable(:,3),kmpTable(:,4),kmpTable(:,1),'VariableNames',...

{'Year','Ur','Gdp','WomenandChidRepression'});
```

```
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(kmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%-----------------RMP_WOMEN----------------
clear all;
clc;
rmpTable=xlsread('RMP_WOMEN.xlsx');
tbl = table(rmpTable(:,2),rmpTable(:,3),rmpTable(:,4),rmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','WomenandChidRepression'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(rmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
%---------------SMP_WOMEN---------------------
clear all;
```

```
clc;
rmpTable=xlsread('RMP_WOMEN.xlsx');
tbl = table(rmpTable(:,2),rmpTable(:,3),rmpTable(:,4),rmpTable(:,1),'VariableNames',...
{'Year','Ur','Gdp','WomenandChidRepression'});
mdl = fitlm(tbl);
disp(mdl);
fprintf('\n');
correlation=corr(rmpTable);
disp('Correlation :');
fprintf('\n');
disp(correlation);
figure(1)
subplot(1,2,1),plotResiduals(mdl,'probability')
subplot(1,2,2),plotResiduals(mdl,'fitted')
```