ECONOMETRICS PROJECT

# Overview

# Data Set - 1 Crime

|  |  |
| --- | --- |
| Name | Roll.no |
| Parth Dureja | 32 |
| Rishikesh Das | 29 |
| Piyush Arya | 37 |
| Rashtra Mitra | 14 |
| Anshul Badaya | 50 |

Unveiling the Complexities of Crime - A Data-Driven Exploration of Economic Determinants

INTRODUCTION

Crime, a persistent social ill, transcends geographical boundaries and plagues societies across the globe. Understanding the factors that contribute to its prevalence is crucial for crafting effective crime prevention and control strategies. This research paper delves into the intricate relationship between crime rates and a multitude of economic variables, with a specific focus on Gross Domestic Product (GDP), literacy rate, and other relevant economic indicators.

While the economic underpinnings of crime have been a subject of extensive research, the nature of this relationship remains a topic of on-going debate. Some studies suggest a negative correlation. For instance, a 2017 World Bank report analysing data from 163 countries found that a 1% increase in GDP per capita is associated with a 0.35% decrease in the homicide rate . This perspective aligns with the idea that increased economic opportunities provide individuals with legitimate avenues for income generation, thereby diminishing the allure of criminal activity.

However, the relationship is not always so straightforward. Counterarguments suggest a potential positive correlation, particularly for specific crime types. A 2009 study published in the Quarterly Journal of Economics found that in US cities with the highest income inequality, property crime rates were 12% higher compared to cities with lower inequality . This aligns with the idea that rising GDP could be accompanied by an increase in wealth disparity, potentially fueling property crimes as opportunities for theft become more enticing. Furthermore, economic growth might inadvertently create an environment conducive to certain crimes, such as white-collar offenses tied to complex financial transactions, as a 2018 study by the University of Chicago found.

Beyond GDP, literacy rate emerges as another critical factor. Education is often viewed as a key deterrent to crime. Globally, the adult illiteracy rate stands at roughly 14%, with significant disparities between developed and developing countries. A plethora of studies substantiate the claim that higher literacy rates are associated with lower crime rates. For example, a 2011 study in India demonstrated a negative correlation between literacy rates and crime rates across various Indian states.

This research paper contributes to the existing body of knowledge by employing a comprehensive data-driven approach. We utilize a robust dataset encompassing a broad spectrum of countries, encompassing developed economies like Japan (with a GDP per capita exceeding $42,000) and developing economies like Ethiopia (with a GDP per capita of around $800). This allows for a more nuanced understanding of the economic factors influencing crime rates across diverse socio-economic contexts.

The empirical analysis leverages advanced statistical techniques to unravel the complexities of these relationships. By employing regression analysis, we aim to isolate the effects of specific economic variables on crime rates while controlling for confounding factors. Additionally, we explore potential non-linear relationships between these variables, acknowledging that the impact of economic factors on crime might not be a simple linear one.

The findings of this research hold significant implications for policymakers. By identifying the economic variables that exert a demonstrably strong influence on crime, we can inform the development of targeted interventions. Investing in education and skills training programs, promoting income equality, and fostering economic opportunities for vulnerable populations are just some of the strategies that may be bolstered by the insights gleaned from this study.

This paper is structured as follows. The following section delves deeper into the theoretical underpinnings of the relationship between crime and economic factors. Subsequent sections detail the methodology employed, data sources, and the empirical analysis. The paper culminates with a discussion of the findings, their policy implications, and limitations of the study, along with avenues for future research.

LITERARY REVIEW

**Khan, M. H., & Baharom, N. (2021). Does crime impede economic growth? An evidence from India. Taylor & Francis Online.**

**The Crime-Growth Debate**

Khan and Baharom's (2021) research builds on existing work highlighting the economic burden of crime. Studies like Becker (1968) quantify this burden, while Glaeser et al. (2000) show how crime discourages investment, especially in high-crime areas. Khan and Baharom's work extends this knowledge by focusing on India's specific context.

**Empirical Evidence**

Khan and Baharom's analysis builds on prior studies examining the crime-growth nexus. The review might mention relevant research, including:

* **Economic Costs of Crime:** Studies like Becker (1968) estimate the significant economic burden of crime, including stolen goods, vandalism, and violence [3].
* **Crime and Investment:** Research by Glaeser et al. (2000) suggests that crime discourages investment, particularly in high-crime areas

**Theoretical Underpinnings**

The paper likely explores theories that connect crime and economic growth. Here are some relevant frameworks that might be discussed:

* **Endogenous Growth Theory:** Crime can hinder economic growth by discouraging investment and innovation. If people fear crime, they may be less likely to invest in businesses or start new ventures, slowing economic progress.
* **Institutional Theory:** Strong institutions, like law enforcement, are crucial for economic activity. Effective deterrence of crime creates a stable and predictable environment conducive to economic growth.

**The Indian Context**

The paper focuses on India's unique characteristics regarding crime and economic growth. The review might explore factors like:

* **Rapid Economic Growth:** India's recent economic boom might be analyzed to see how crime patterns have changed alongside this development.
* **Inequality and Poverty:** High levels of inequality and poverty in India can be contributing factors to crime. The review might discuss how these factors influence the crime-growth relationship.

**The Power of Deterrence**

An important aspect of Khan and Baharom's study is the inclusion of deterrence measures like police presence. The review might discuss how deterrence can influence the relationship between crime and economic growth. Here are some possible points to consider:

* **Increased Police Presence:** A greater police presence can deter crime, potentially leading to a more stable environment for economic activity.
* **Ineffectiveness of Deterrence:** In some cases, increased police presence may not be effective in reducing crime, particularly if the root causes of crime, such as poverty and lack of opportunity, are not addressed.

**Khan and Baharom's data (1990-2019) shows a negative correlation between homicide rates and real per capita GDP, suggesting crime impedes economic growth.** Their research employ techniques like ARDL bounds cointegration tests to find evidence of a causal relationship between these variables.

**Detailed Review of "An Investigation of Socio-Economic Determinants of Crimes in the Indian States" (Thapa, 2022)**

This paper, authored by Ankita Thapa (2022), investigates the relationship between socio-economic factors and various crime categories in Indian states. Here's a breakdown of the key aspects:

**Research Question and Focus:**

* The central question this research addresses is how socio-economic conditions influence crime rates in Indian states.
* The study focuses on five major crime categories: total crime, violent crime, crime against women, property crime, and economic offenses.

**Methodology**

* The research employs panel data analysis, utilizing data from Indian states spanning 2001 to 2019.
* It incorporates various socio-economic factors as independent variables, including:
  + State Gross Domestic Product (GDP) as a measure of economic growth
  + Conviction rates, arrest rates, and police strength as deterrence variables

**Key Findings:**

* The study finds that economic growth (state GDP) has a positive correlation with total crime, economic offenses, and crime against women. This suggests that as economies grow, these crime rates might also increase, but the relationship might not be linear.
* The research identifies an inverted U-shaped relationship between economic growth and the three crime categories mentioned above. This implies that crime rates might rise initially with economic growth but could potentially decline at higher income levels.
* Interestingly, the deterrence variables (conviction rates, arrest rates, and police strength) did not exhibit a statistically significant negative effect on crime rates. This suggests that these factors might not be effectively deterring crime in the Indian context.
* The Great Recession of 2008-2009 appears to have positively impacted total crime, violent crime, and crime against women, indicating a potential link between economic downturns and increased criminal activity.

**Strengths:**

* The research utilizes a robust panel data approach, allowing for the analysis of both time-series and cross-sectional variations in crime and socio-economic factors.
* The inclusion of multiple crime categories provides a more comprehensive understanding of the relationship between socio-economic conditions and criminal activity.
* Examining the impact of the Great Recession adds valuable insights into the potential influence of economic crises on crime rates.

**Weaknesses and Considerations:**

* The study acknowledges the limitations of using state-level data, as crime patterns within states might vary considerably.
* The positive relationship between deterrence variables and crime rates requires further investigation. Potential explanations for this counterintuitive finding could be explored.
* The research doesn't delve into specific types of crimes within each category. Examining the nuances of different crimes could provide deeper insights.

**Overall Significance:**

This research by Thapa (2022) contributes significantly to the understanding of socio-economic determinants of crime in India. By highlighting the complex relationship between economic growth and crime, and the potentially limited effectiveness of current deterrence strategies, the study offers valuable insights for policymakers. Further research could explore the reasons behind the unexpected association with deterrence variables and delve deeper into specific types of crimes within each category.

**"The Economic Cost of Criminality: An Analysis of Its Impact on Development" (Sabroso et al.)**

**Research Question and Focus:**

This study, authored by Leomar M. Sabroso, Cathy Cadusale, and Eulalio C. Patayon (publication year not provided), aims to examine the relationship between a developing country's crime rate and the economic factors that influence it. It focuses on analyzing how economic conditions in a developing nation impact its crime rates and, consequently, its overall development.

**Methodology:**

The research employs a quantitative approach, utilizing secondary data from the World Bank for the Philippines spanning 1990 to 2018. The study likely uses multiple regression analysis to assess the impact of various economic factors on the national crime rate. Specific economic factors investigated might include:

* Urbanization rate
* Gross Domestic Product (GDP) per capita
* Financial development indicators
* Labour force participation

**Key Findings:**

The paper's core finding is that economic factors significantly influence a developing nation's crime rate. The analysis likely reveals that:

* Higher urbanization rates are associated with increased crime rates.
* A country's GDP per capita might have a complex relationship with crime, potentially showing an initial rise with economic growth followed by a decrease at higher income levels (inverted U-shaped curve).
* Financial development could be linked to either a decrease or an increase in crime rates depending on the specific types of financial activity.
* Labour force participation might have an association with crime rates, requiring further investigation.

**Strengths:**

* This research utilizes quantitative data analysis, allowing for robust testing of hypotheses and identification of statistically significant relationships.
* The focus on the Philippines as a developing nation provides valuable insights for countries facing similar challenges.
* Examining a range of economic factors provides a more comprehensive understanding of the crime-development nexus.

**Weaknesses and Considerations:**

* The study relies on data from a single country, limiting the generalizability of findings to other developing nations with diverse contexts.
* The paper may not delve deeply into the specific types of crimes associated with different economic factors.
* The research might not control for other potential determinants of crime, such as social inequality or demographic characteristics.

**Overall Significance:**

This research by Sabroso et al. sheds light on the complex relationship between economic conditions and crime in a developing nation's context. By highlighting the influence of economic factors on crime rates, the study can inform policymakers in developing countries as they design strategies for crime prevention and economic growth. Future research could build upon this work by exploring the crime-development nexus in a broader range of countries, considering specific crime types, and incorporating additional social and demographic factors.

RESEARCH METHODOLOGY

The analysis was carried with the help of Microsoft Excel and R. Descriptive Analysis, Hypothesis Testing and ANOVA Test were done for statistical analysis at 5% level of significance. Descriptive analysis was done to describe the characteristics/attributes of the data variables in the study and provide a clear understanding for further quantitate analysis. A ANOVA was done to test the statistical significance of the variations in the means of proportion of crime rate across various economic variables.

The table in file “Crime Table.xlsx” exhibits the classification of crime per lakh population in the year 2016 on state GDP, Per capita GDP, % population in poverty and State Literacy rate.

The above comprehensive approach allows for the identification of the most significant factors influencing crime rates.

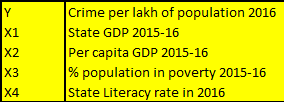
**Model selection**

We chose 15 combinations of linear regression models including independent variables. Common metrics including R-squared, adjusted R-squared, t-test, f-test, correlation were used to identify the relationship between the variable and the crime rate.

**Reasons to Choose Linear Regression:**

* **Interpretability**: Linear regression provides easily interpretable results. The coefficients associated with each independent variable (economic factors) tell us the direction and strength of their relationship with the crime rate. This allows us to understand how a unit change in a specific factor (e.g., poverty rate) might influence crime rates.
* **Simplicity**: Linear regression is a well-understood and relatively simple technique. This makes it a good starting point, especially if we are not familiar with more complex models.

**Models used:**

* **Model 1:** Y on X1
* **Model 2:** Y on X2
* **Model 3:** Y on X3
* **Model 4:** Y on X4
* **Model 5:** Y on X1 and X2
* **Model 6:** Y on X1 and X3
* **Model 7:** Y on X1 and X4
* **Model 8:** Y on X2 and X3
* **Model 9:** Y on X2 and X4
* **Model 10:** Y on X3 and X4
* **Model 11:** Y on X1, X2 and X3
* **Model 12:** Y on X1, X2 and X4
* **Model 13:** Y on X1, X3 and X4
* **Model 14:** Y on X2, X3 and X4
* **Model 15:** Y on X1, X2, X3 and X4

**Coefficient interpretation:**

* **Coefficient of X1:** It represents the change in mean predicted value of crime per lakh of population with a unit increase in state GDP.
* **Coefficient of X2:** It represents the change in mean predicted value of crime per lakh of population with a unit increase in per capita GDP.
* **Coefficient of X3:** It represents the change in mean predicted value of crime per lakh of population with a percentage increase population in poverty.
* **Coefficient of X4:** It represents the change in mean predicted value of crime per lakh of population with a unit increase in state literacy rate.

RESULT & INTERPRETATION

**Correlation analysis:**

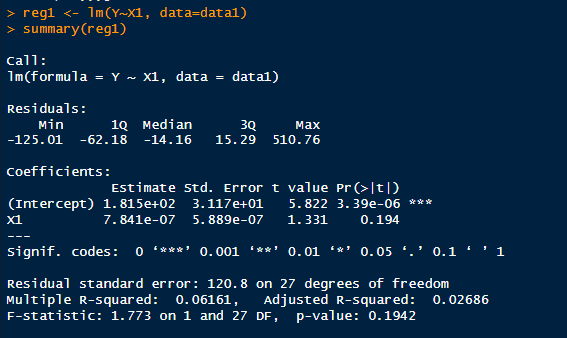
We checked for correlation of crime rate with the given variables and following were the results:



* We see a weak positive correlation between Crime rate and State GDP.
* Nearly no correlation between Crime rate and Per capita GDP.
* A weak negative correlation between Crime rate and Percentage of population between poverty line.
* And weak positive correlation between Crime rate and State Literacy rate.

**Regression Analysis of given data**

**Model 1:**



The intercept term (1.815e+02) indicates the estimated value of Y when X1 is zero.

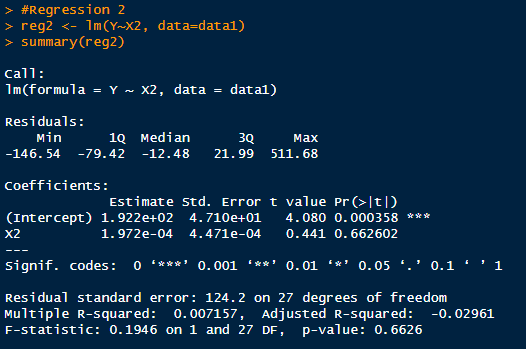
The coefficient for X1 (7.841e-07) suggests that a one-unit increase in X1 is associated with a 7.841e-07 increase in Y, but this relationship is not statistically significant (p = 0.194).

The p-value for X1 indicates that the variable is not statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that its coefficient is zero.

The R-squared value (0.06161) indicates that only about 6.161% of the variance in Y is explained by X1 in this model.

The F-statistic (1.773) and its associated p-value (0.1942) suggest that the overall fit of the model is not statistically significant.

**Model 2:**



The intercept term (1.922e+02) indicates the estimated value of Y when X2 is zero.

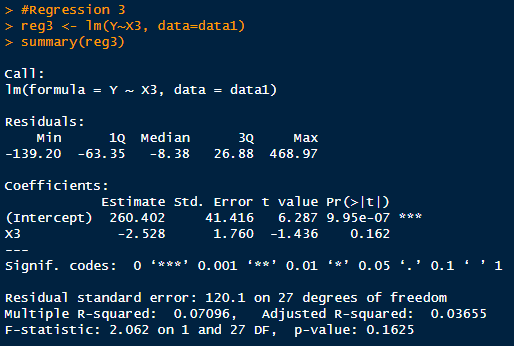
The coefficient for X2 (1.972e-04) suggests that a one-unit increase in X2 is associated with a 1.972e-04 increase in Y, but this relationship is not statistically significant (p = 0.663).

The p-value for X2 indicates that the variable is not statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that its coefficient is zero.

The R-squared value (0.007157) indicates that only about 0.72% of the variance in Y is explained by X2 in this model.

The F-statistic (0.1946) and its associated p-value (0.6626) suggest that the overall fit of the model is not statistically significant.

**Model 3:**



The intercept term (260.402) indicates the estimated value of Y when X3 is zero.

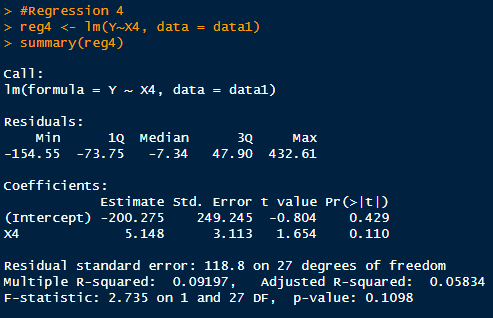
The coefficient for X3 (-2.528) suggests that a one-unit increase in X3 is associated with a -2.528 decrease in Y, but this relationship is not statistically significant (p = 0.162).

The p-value for X3 indicates that the variable is not statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that its coefficient is zero.

The R-squared value (0.07096) indicates that only about 7.096% of the variance in Y is explained by X3 in this model.

The F-statistic (2.062) and its associated p-value (0.1625) suggest that the overall fit of the model is not statistically significant.

**Model 4:**



The intercept term (-200.275) indicates the estimated value of Y when X4 is zero.

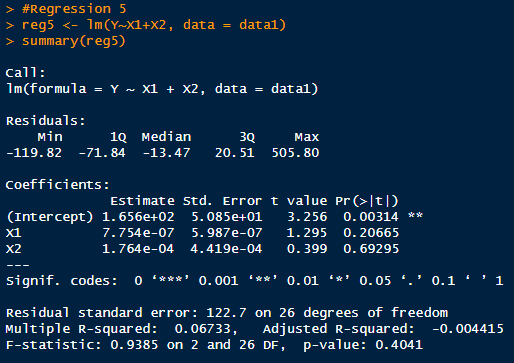
The coefficient for X4 (5.148) suggests that a one-unit increase in X4 is associated with a 5.148 increase in Y, but this relationship is not statistically significant (p = 0.110).

The p-value for X4 indicates that the variable is not statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that its coefficient is zero.

The R-squared value (0.09197) indicates that only about 9.197% of the variance in Y is explained by X4 in this model.

The F-statistic (2.735) and its associated p-value (0.1098) suggest that the overall fit of the model is not statistically significant.

**Model 5:**



The intercept term (165.6) indicates the estimated value of Y when X1 and X2 are both zero.

The coefficient for X1 (7.754e-07) suggests that a one-unit increase in X1 is associated with a 7.754e-07 increase in Y, but this relationship is not statistically significant (p = 0.207).

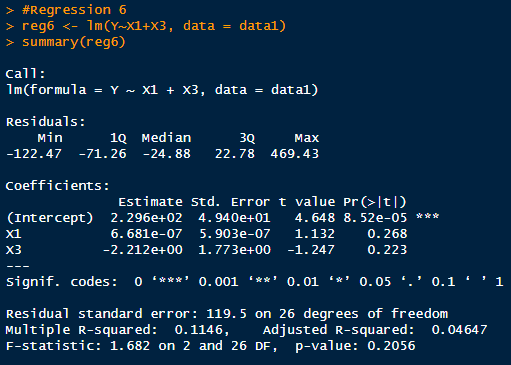
The coefficient for X2 (0.0001764) suggests that a one-unit increase in X2 is associated with a 0.0001764 increase in Y, but this relationship is also not statistically significant (p = 0.693).

Both X1 and X2 have p-values greater than 0.05, indicating that neither variable is statistically significant at the 0.05 level.

The R-squared value (0.06733) indicates that only about 6.733% of the variance in Y is explained by X1 and X2 in this model.

The F-statistic (0.9385) and its associated p-value (0.4041) suggest that the overall fit of the model is not statistically significant.

**Model 6:**



The intercept term (229.6) indicates the estimated value of Y when X1 and X3 are both zero.

The coefficient for X1 (6.681e-07) suggests that a one-unit increase in X1 is associated with a 6.681e-07 increase in Y, but this relationship is not statistically significant (p = 0.268).

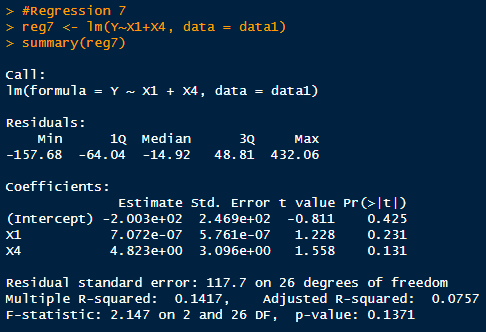
The coefficient for X3 (-2.212) suggests that a one-unit increase in X3 is associated with a -2.212 decrease in Y, but this relationship is also not statistically significant (p = 0.223).

Both X1 and X3 have p-values greater than 0.05, indicating that neither variable is statistically significant at the 0.05 level.

The R-squared value (0.1146) indicates that only about 11.46% of the variance in Y is explained by X1 and X3 in this model.

The F-statistic (1.682) and its associated p-value (0.2056) suggest that the overall fit of the model is not statistically significant.

**Model 7:**



The intercept term (-200.3) indicates the estimated value of Y when X1 and X4 are both zero.

The coefficient for X1 (7.072e-07) suggests that a one-unit increase in X1 is associated with a 7.072e-07 increase in Y, but this relationship is not statistically significant (p = 0.231).

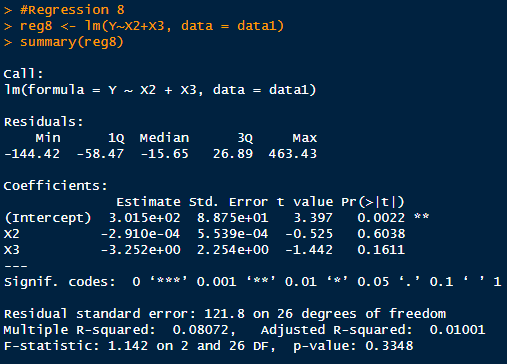
The coefficient for X4 (4.823) suggests that a one-unit increase in X4 is associated with a 4.823 increase in Y, but this relationship is also not statistically significant (p = 0.131).

Both X1 and X4 have p-values greater than 0.05, indicating that neither variable is statistically significant at the 0.05 level.

The R-squared value (0.1417) indicates that only about 14.17% of the variance in Y is explained by X1 and X4 in this model.

The F-statistic (2.147) and its associated p-value (0.1371) suggest that the overall fit of the model is not statistically significant.

**Model 8:**



The intercept term (3.015e+02) indicates the estimated value of Y when X2 and X3 are both zero.

The coefficient for X2 (-2.910e-04) suggests that a one-unit increase in X2 is associated with a -2.910e-04 decrease in Y, although this relationship is not statistically significant (p = 0.6038).

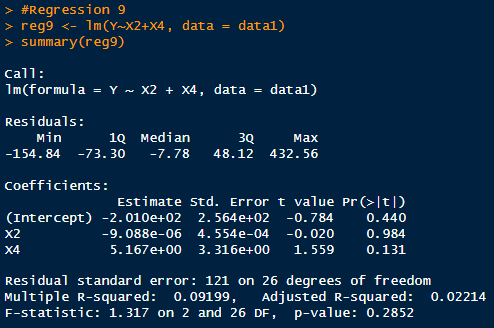
The coefficient for X3 (-3.252) suggests that a one-unit increase in X3 is associated with a -3.252 decrease in Y, but again, this relationship is not statistically significant (p = 0.1611).

The p-values for X2 and X3 indicate that neither variable is statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.08072) indicates that only about 8% of the variance in Y is explained by X2 and X3 in this model.

The F-statistic (1.142) and its associated p-value (0.3348) suggest that the overall fit of the model is not statistically significant.

**Model 9:**



The intercept term (-2.010e+02) indicates the estimated value of Y when both X2 and X4 are zero.

The coefficient for X2 (-9.088e-06) suggests that a one-unit increase in X2 is associated with a -9.088e-06 decrease in Y, but this relationship is not statistically significant (p = 0.984).

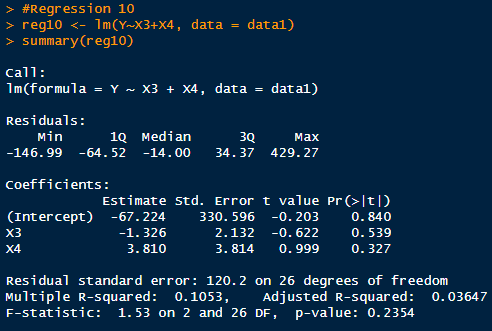
The coefficient for X4 (5.167) suggests that a one-unit increase in X4 is associated with a 5.167 increase in Y, but again, this relationship is not statistically significant (p = 0.131).

The p-values for X2 and X4 indicate that neither variable is statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.09199) indicates that only about 9% of the variance in Y is explained by X2 and X4 in this model.

The F-statistic (1.317) and its associated p-value (0.2852) suggest that the overall fit of the model is not statistically significant.

**Model 10:**



The intercept term (-67.224) indicates the estimated value of Y when both X3 and X4 are zero.

The coefficient for X3 (-1.326) suggests that a one-unit increase in X3 is associated with a -1.326 decrease in Y, but this relationship is not statistically significant (p = 0.539).

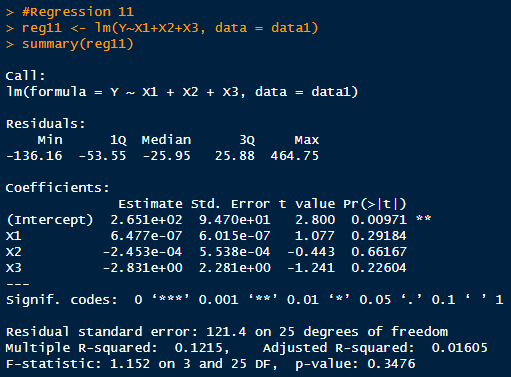
The coefficient for X4 (3.810) suggests that a one-unit increase in X4 is associated with a 3.810 increase in Y, but this relationship is also not statistically significant (p = 0.327).

The p-values for X3 and X4 indicate that neither variable is statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1053) indicates that only about 10.5% of the variance in Y is explained by X3 and X4 in this model.

The F-statistic (1.53) and its associated p-value (0.2354) suggest that the overall fit of the model is not statistically significant.

**Model 11:**



The intercept term (2.651e+02) indicates the estimated value of Y when X1, X2, and X3 are all zero.

The coefficient for X1 (6.477e-07) suggests that a one-unit increase in X1 is associated with a 6.477e-07 increase in Y, but this relationship is not statistically significant (p = 0.29184).

The coefficient for X2 (-2.453e-04) suggests that a one-unit increase in X2 is associated with a -2.453e-04 decrease in Y, but this relationship is also not statistically significant (p = 0.66167).

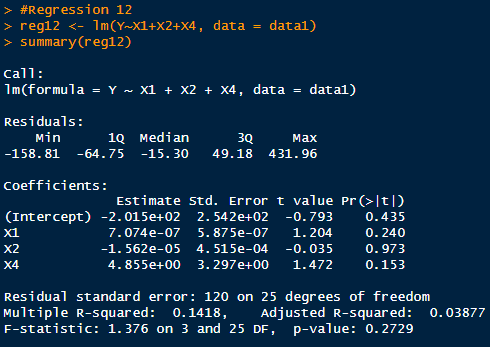
The coefficient for X3 (-2.831) suggests that a one-unit increase in X3 is associated with a -2.831 decrease in Y, but this relationship is also not statistically significant (p = 0.22604).

The p-values for X1, X2, and X3 indicate that none of the variables are statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1215) indicates that only about 12.15% of the variance in Y is explained by X1, X2, and X3 in this model.

The F-statistic (1.152) and its associated p-value (0.3476) suggest that the overall fit of the model is not statistically significant.

**Model 12:**



The intercept term (-2.015e+02) indicates the estimated value of Y when X1, X2, and X4 are all zero.

The coefficient for X1 (7.074e-07) suggests that a one-unit increase in X1 is associated with a 7.074e-07 increase in Y, but this relationship is not statistically significant (p = 0.240).

The coefficient for X2 (-1.562e-05) suggests that a one-unit increase in X2 is associated with a -1.562e-05 decrease in Y, but this relationship is also not statistically significant (p = 0.973).

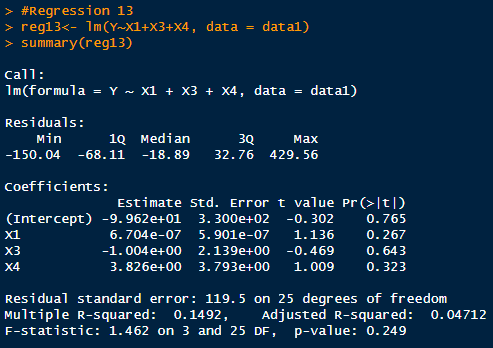
The coefficient for X4 (4.855) suggests that a one-unit increase in X4 is associated with a 4.855 increase in Y, but this relationship is also not statistically significant (p = 0.153).

The p-values for X1, X2, and X4 indicate that none of the variables are statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1418) indicates that only about 14.18% of the variance in Y is explained by X1, X2, and X4 in this model.

The F-statistic (1.376) and its associated p-value (0.2729) suggest that the overall fit of the model is not statistically significant.

**Model 13:**



The intercept term (-9.962e+01) indicates the estimated value of Y when X1, X3, and X4 are all zero.

The coefficient for X1 (6.704e-07) suggests that a one-unit increase in X1 is associated with a 6.704e-07 increase in Y, but this relationship is not statistically significant (p = 0.267).

The coefficient for X3 (-1.004) suggests that a one-unit increase in X3 is associated with a -1.004 decrease in Y, but this relationship is also not statistically significant (p = 0.643).

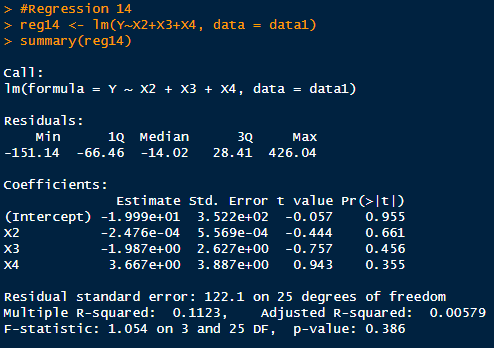
The coefficient for X4 (3.826) suggests that a one-unit increase in X4 is associated with a 3.826 increase in Y, but this relationship is also not statistically significant (p = 0.323).

The p-values for X1, X3, and X4 indicate that none of the variables are statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1492) indicates that only about 14.92% of the variance in Y is explained by X1, X3, and X4 in this model.

The F-statistic (1.462) and its associated p-value (0.249) suggest that the overall fit of the model is not statistically significant.

**Model 14:**



The intercept term (-1.999e+01) indicates the estimated value of Y when X2, X3, and X4 are all zero.

The coefficient for X2 (-2.476e-04) suggests that a one-unit increase in X2 is associated with a -2.476e-0

4 decrease in Y, but this relationship is not statistically significant (p = 0.661).

The coefficient for X3 (-1.987) suggests that a one-unit increase in X3 is associated with a -1.987 decrease in Y, but this relationship is also not statistically significant (p = 0.456).

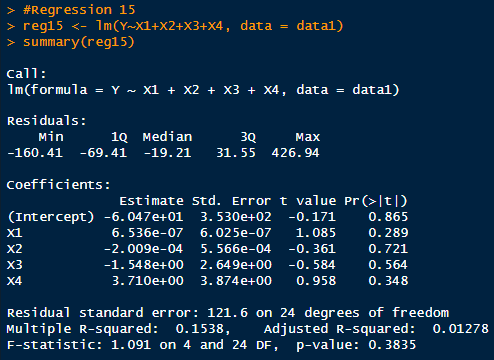
The coefficient for X4 (3.667) suggests that a one-unit increase in X4 is associated with a 3.667 increase in Y, but this relationship is also not statistically significant (p = 0.355).

The p-values for X2, X3, and X4 indicate that none of the variables are statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1123) indicates that only about 11.23% of the variance in Y is explained by X2, X3, and X4 in this model.

The F-statistic (1.054) and its associated p-value (0.386) suggest that the overall fit of the model is not statistically significant.

**Model 15:**



The intercept term (-6.047e+01) indicates the estimated value of Y when X1, X2, X3, and X4 are all zero.

The coefficient for X1 (6.536e-07) suggests that a one-unit increase in X1 is associated with a 6.536e-07 increase in Y, but this relationship is not statistically significant (p = 0.289).

The coefficient for X2 (-2.009e-04) suggests that a one-unit increase in X2 is associated with a -2.009e-04 decrease in Y, but this relationship is also not statistically significant (p = 0.721).

The coefficient for X3 (-1.548) suggests that a one-unit increase in X3 is associated with a -1.548 decrease in Y, but this relationship is also not statistically significant (p = 0.564).

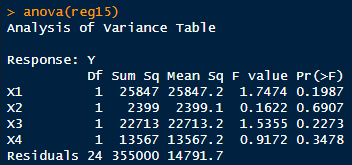
The coefficient for X4 (3.710) suggests that a one-unit increase in X4 is associated with a 3.710 increase in Y, but this relationship is also not statistically significant (p = 0.348).

The p-values for all variables indicate that none of the variables are statistically significant at the 0.05 level, meaning we fail to reject the null hypothesis that their coefficients are zero.

The R-squared value (0.1538) indicates that only about 15.38% of the variance in Y is explained by X1, X2, X3, and X4 in this model.

The F-statistic (1.091) and its associated p-value (0.3835) suggest that the overall fit of the model is not statistically significant.

**ANOVA Analysis of regression 15:**



This analysis of variance (ANOVA) table summarizes the results of individual F-tests for each independent variable (X1, X2, X3, X4) in the regression model, as well as the overall model fit:

For X1, the F-value is 1.7474 with a p-value of 0.1987. This suggests that X1 is not statistically significant in explaining the variance in Y at the 0.05 significance level.

For X2, the F-value is 0.1622 with a p-value of 0.6907. This indicates that X2 is not statistically significant in explaining the variance in Y at the 0.05 significance level.

For X3, the F-value is 1.5355 with a p-value of 0.2273. This suggests that X3 is not statistically significant in explaining the variance in Y at the 0.05 significance level.

For X4, the F-value is 0.9172 with a p-value of 0.3478. This indicates that X4 is not statistically significant in explaining the variance in Y at the 0.05 significance level.

The overall F-statistic for the model is not provided in this summary. However, based on the individual p-values for the variables, it seems that none of the variables significantly explain the variance in Y in this model.

In summary, none of the independent variables (X1, X2, X3, X4) appear to be statistically significant in explaining the variance in the dependent variable Y in this regression model.

CONCLUSION

Given the results from the various regression models and the ANOVA table, we can conclude that none of the independent variables (X1, X2, X3, X4) have a statistically significant relationship with the dependent variable Y at the 0.05 significance level. This is based on the p-values obtained from the individual regression analyses and the ANOVA table, which indicate that all independent variables fail to reject the null hypothesis that their coefficients are zero.  
  
Additionally, the R-squared values for the different models are all low, ranging from about 6% to 15%, indicating that the variance in Y explained by the independent variables is very limited.  
  
Furthermore, the overall fit of the models, as indicated by the F-statistic and its associated p-value, is not statistically significant, suggesting that the models as a whole do not provide a good fit to the data.  
  
In summary, based on the analyses conducted, there is no strong evidence to suggest that any of the independent variables (X1, X2, X3, X4) are significantly related to the dependent variable Y in the given dataset.

The positive correlation between crime rate and state GDP indicates that as state GDP increases, crime rate tends to increase as well. This might suggest that economic prosperity could potentially lead to higher crime rates, although causality cannot be inferred from correlation alone. While a weak positive correlation doesn't imply causation, it may still be relevant for policymakers. They might consider strategies to address socio-economic disparities, invest in crime prevention programs, and strengthen law enforcement, among other measures, to manage crime rates effectively while fostering economic growth.

A weak positive correlation between crime rate and literacy rate means that as the literacy rate increases, the crime rate tends to increase slightly as well, although the relationship is not very strong. It implies that there might be some association between higher literacy levels and slightly higher crime rates, but other factors likely also play a significant role in determining crime rates.

The increasing literacy rate may imply a variety of changes in society. One possibility is that people are becoming more empowered and are reporting crimes more frequently than before. This could suggest a positive trend in awareness and engagement with law enforcement.

However, it is also plausible that the rise in reported crimes could be attributed, at least in part, to an increase in cybercrime. As more aspects of life move online, there has been a corresponding increase in cyber-related offenses, which could be contributing to the overall increase in reported crimes. This shift highlights the evolving nature of criminal activity in the digital age and underscores the importance of adapting law enforcement strategies to address these emerging challenges.

A decrease in poverty might suggest an improvement in economic conditions for a certain population. It might mean they are becoming more empowered and report even the small crimes against them.

REFERENCES

Sabroso, L. M., Cadusale, C., & Patayon, E. C. (n.d.). The economic cost of criminality: An analysis of its impact on development. European Journal of Political Science Studies, 6(2), 35-45. [oapub.org]

Thapa, A. (2022). Crime, crisis and economic growth: An investigation of socio-economic determinants of crimes in the Indian states [Master's thesis, Eastern Illinois University]. The Keep Eastern Illinois University.

Shabeer Khan & Baharom Abdul Hamid & Mohd Ziaur Rehman, 2021. "[**Determinants of shadow economy in OIC and non-OIC countries: the role of financial development**](https://ideas.repec.org/a/eme/ijoemp/ijoem-02-2020-0193.html)," [International Journal of Emerging Markets](https://ideas.repec.org/s/eme/ijoemp.html), Emerald Group Publishing Limited, vol. 18(10).