

Exploring the Relationship Between Socioeconomic Factors and Crime Rates in the United States

Farha Hussaini- Duaa Suroor- Wejdan Alshateri

Introduction:

Crime is a complex issue that affects individuals, families, and entire communities. Many factors contribute to the occurrence of crime, including socioeconomic conditions. Understanding the relationship between socioeconomic factors and crime rates is essential to developing effective crime prevention strategies and improving public safety. In the United States, crime rates have fluctuated over time, and identifying the underlying factors that influence these changes is critical for policymakers and law enforcement officials.

roadmap of the project, Firstly, the necessary libraries, including widely used packages will be downloaded and loaded. These libraries provide powerful tools for data manipulation and visualization, ensuring the analysis is based on solid foundations. Next, the dataset containing relevant information on socioeconomic factors and crime rates will be obtained from The World Bank. The dataset is typically in CSV format, will be read into R using appropriate functions. Attention will be paid to accurately specifying the file path and handling any potential errors during the data import process. Once the dataset is loaded, it will be carefully arranged and organized. This includes conducting data cleaning procedures, addressing missing values or outliers, and ensuring proper formatting of variables. Each dataset related to socioeconomic factors and crime rates will be meticulously processed to facilitate a comprehensive understanding of the underlying factors. To gain insights and effectively communicate the relationship between socioeconomic factors and crime rates, the data will be visualized using tools such as ggplot and other suitable visualization packages. Through carefully constructed plots, including scatter plots, bar charts, or line graphs, the complex relationships and trends within the data will be visually represented. Attention will be given to choosing appropriate visual encodings and incorporating clear labels and titles to enhance the interpretability of the visualizations. By following this professional approach, the RMarkdown file will provide a thorough analysis of the relationship between socioeconomic factors and crime rates. The combination of data organization, visualization, and the use of industry-standard libraries ensures the results are accurate, reliable, and effectively communicated.

Background:

The United States has one of the highest crime rates among developed countries. According to the Uniform Crime Reporting (UCR) program, the national crime rate in 2019 was 3,450.5 incidents per 100,000 people. While the overall crime rate has decreased over the past few decades, certain types of crimes, such as property crimes, have remained prevalent in many communities. Research has shown that socioeconomic factors, such as poverty, unemployment, and education levels, can significantly impact crime rates. For instance, individuals living in poverty may resort to criminal activities to meet their basic needs, while unemployment can lead to frustration and a sense of hopelessness that can contribute to criminal behavior. Similarly, low levels of education can limit job prospects and increase the likelihood of criminal activity.

Despite the known relationship between socioeconomic factors and crime rates, there is still much to learn about the specific ways in which these factors influence crime, and how they can be effectively addressed. Additionally, there is a need to explore the potential of using socioeconomic factors to develop predictive

models for crime rates, which can help law enforcement and policymakers to make informed decisions about resource allocation and crime prevention strategies.

This research aims to explore the relationship between socioeconomic factors and crime rates in the United States. Specifically, the research will examine the extent to which poverty, unemployment, and education levels impact crime rates, and whether a predictive model can be developed to forecast future crime rates based on these factors. To achieve these objectives, the research will use a combination of quantitative methods, including statistical analysis and machine learning techniques.

The findings of this research will contribute to a deeper understanding of the complex relationship between socioeconomic factors and crime rates, and provide insights into the potential of using these factors to develop predictive models for crime rates. Ultimately, this research can inform the development of evidence-based policies and strategies aimed at reducing crime rates and improving public safety in the United States.

Research Question:

To what extent do socioeconomic factors such as poverty, unemployment, and education levels impact crime rates in the United States, and can we develop a predictive model to forecast future crime rates based on these factors?

Literature Review

Crime is a multifaceted phenomenon with multiple underlying reasons for individuals engaging in criminal acts. It is a complex and dynamic concept that varies among civilizations and throughout time. The benefits and appeal of the target, along with the conducive nature of the environment, can drive some individuals to commit crimes [1]. However, most criminals are motivated by a range of factors, including unemployment, the desire to narrow the wealth gap, involvement in narcotics, greed, and shortcomings in the crime control model of national security [1]. Understanding the linkages between crime and various socio-economic factors is crucial for effective crime prevention and security enhancement. The recognition of crime as a sustainable development issue, particularly through the 16th goal of the United Nations, has compelled states to combat organized crime and foster peaceful societies [4]. This acknowledgment has led to the implementation of crime prevention programs aimed at reducing crime and facilitating development. Socioeconomic development has been found to influence the volume and nature of criminal behavior, further emphasizing the importance of addressing crime within a broader societal context [4]. Criminological research has identified specific social traits that are associated with a higher risk of engaging in criminal activities. Young, male individuals from economically disadvantaged backgrounds are more prone to becoming offenders [2]. Additionally, unemployment has been consistently linked to crime within communities. While the extent of the causal link is still a subject of debate, studies indicate that job relocation, rather than general unemployment, has a more significant impact on crime rates. Furthermore, a strong positive relationship exists between pay disparity and crime, as supported by the economic theory of crime. This theory suggests that higher wages increase the opportunity cost of illegal actions, while poverty exacerbates income gaps and makes crime more appealing [1]. Serious offenses such as assault, robbery, and homicide are closely connected to social and economic disadvantage. Offenders involved in these crimes are often either unemployed or engaged in low-paying, unskilled employment [3]. This highlights the critical role of addressing social and economic disparities in crime prevention efforts. In conclusion, crime is a complex and dynamic phenomenon influenced by a multitude of factors. Understanding the various causes of crime and their interplay is crucial for the development of effective crime prevention strategies. Socioeconomic development, unemployment, income disparity, and social disadvantage all contribute significantly to the occurrence and nature of criminal behavior. By considering these factors comprehensively, policymakers and stakeholders can better strengthen security measures and foster safer and more peaceful societies.

Packages

```
library(tidyverse)
library(dplyr)
library(readr)
library(ggplot2)
library(sf)
```

Data

1- Violent Crime Rates by US State: This data set contains statistics, in arrests per 100,000 residents for assault, murder, and rape in each of the 50 US states in 1973. Also given is the percent of the population living in urban areas.

2- Unemployment Rates for States: contains information about the unemployment rates and ranks for each state in the United States. It provides information about the unemployment situation across the United States, allowing for comparisons between states and an understanding of the variation in unemployment rates across the country.

3- Educational Attainment by State: Education is one of the dominant factors in determining how developed a country is. According to the Global Partnership for Education, “Education is a vital human right and plays a key role in human, social, and economic development. Education reduces poverty, boosts the economy, promotes gender equality, and promotes peace. Education is essential in achieving the 17 Global Goals for sustainable development to end extreme poverty, fight inequality and injustice, and protect our planet by 2030. These goals include no poverty, zero hunger, clean water and sanitation, sustainable cities, responsible consumption and production, and more.

4- Poverty Rate by State: Poverty is defined as not having enough income to meet basic needs. This goes beyond just not having enough money for a new vehicle, smartphone, or to take a vacation. Instead, people who live in poverty struggle to keep a roof over their heads, put food on the table, or even purchase basic items like clothing, shoes, and hygiene items.

reading and loading these datasets into R.

```
unemploye_rate <- read_csv("data/unemploye_rate.csv")
crimeRate_US <- read_csv("data/US_violent_crime.csv")
education <- read_csv("data/educational-attainment-by-state-2023.csv")
poverty_rate <- read_csv("data/poverty-rate-by-state-2023.csv")
```

```
names(crimeRate_US)
```

```
## [1] "...1"      "Murder"     "Assault"    "UrbanPop"   "Rape"
```

```
crimeRate_US <- crimeRate_US %>%
  rename(state = ...1)
```

```
names(unemploye_rate)
```

```
## [1] "State" "rate"  "Rank"
```

```
unemploye_rate <- unemploye_rate %>%
  rename(state = State)
```

```
names(poverty_rate)
```

```
## [1] "state"          "PovertyRate" "undefined"
```

order by AESC

```
poverty_rate <- poverty_rate[order (poverty_rate$state,decreasing = FALSE),]
unemploye_rate <- unemploye_rate[order (unemploye_rate$state,decreasing = FALSE),]
crimeRate_US <- crimeRate_US[order (crimeRate_US$state,decreasing = FALSE),]
education <- education[order (education$state,decreasing = FALSE),]
```

This code sort the datasets (poverty_rate, unemploye_rate, crimeRate_US, and education) in ascending order based on the “state”. this will ensure that the datasets are arranged alphabetically by the “state” column in ascending order, making it easier to compare and analyze the data.

leftjoin

```
Us_crimes <- merge(crimeRate_US, poverty_rate, by = "state", all.x = TRUE)
Us_crimes <- merge(Us_crimes, education, by = "state", all.x = TRUE)
Us_crimes <- merge(Us_crimes, unemploye_rate, by = "state", all.x = TRUE)
```

Merge the datasets crimeRate_US, poverty_rate, education, and unemploye_rate based on the “state” column. The resulting merged dataset is assigned to the variable Us_crimes By merging these datasets together, the resulting Us_crimes dataset will contain the combined information from crimeRate_US, poverty_rate, education, and unemploye_rate based on the matching “state” column. This allows for a comprehensive analysis of the relationship between crime rates, poverty rates, educational attainment, and unemployment rates in the United States.

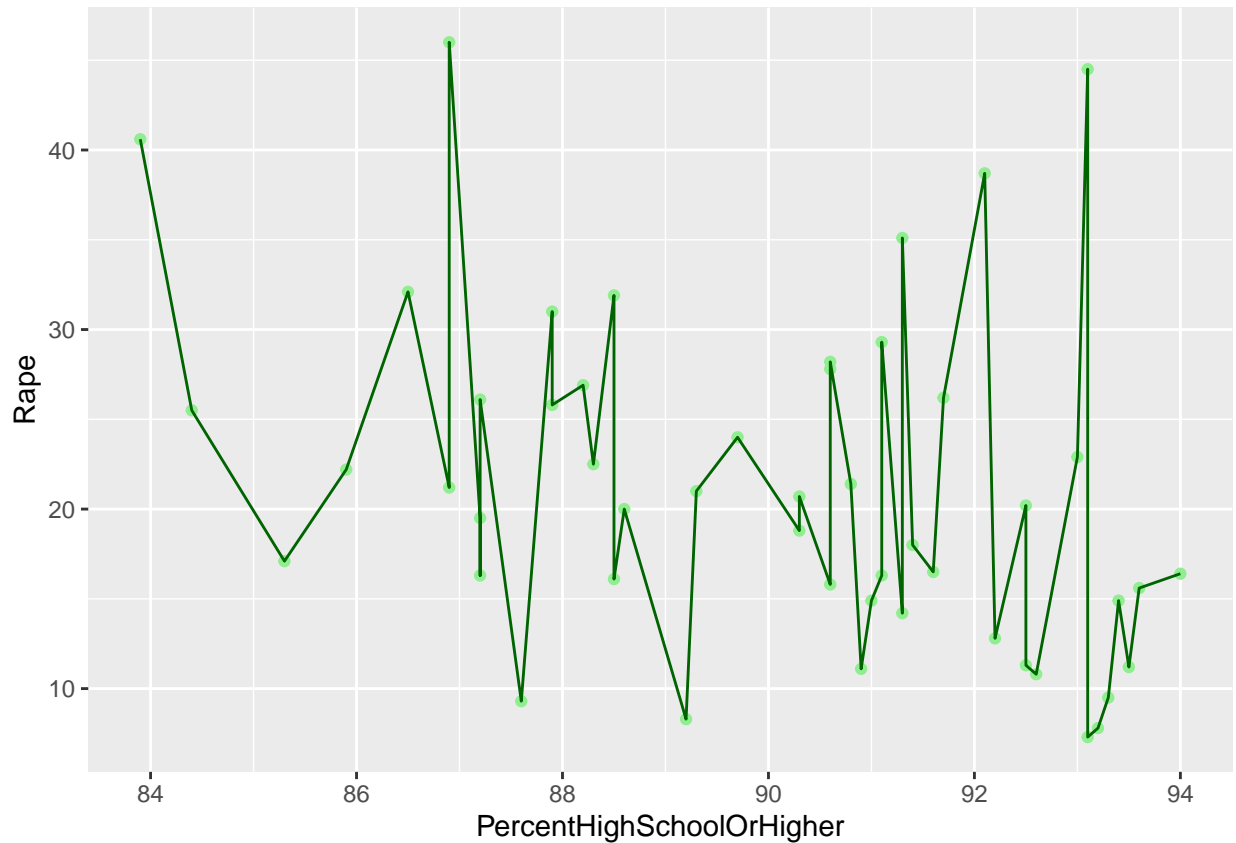
```
names(Us_crimes)
```

```
## [1] "state"          "Murder"
## [3] "Assault"        "UrbanPop"
## [5] "Rape"           "PovertyRate"
## [7] "undefined"      "PercentHighSchoolOrHigher"
## [9] "PercentBachelorsOrHigher" "rate"
## [11] "Rank"
```

Results

Assuming “poverty_rate”, “unemployment_rate”, “education_level”, and “crime_rate” are the relevant variables

```
ggplot(Us_crimes, aes(x = PercentHighSchoolOrHigher, y = Rape)) +  
  geom_point(col="lightgreen") +geom_line(aes(group = 1), col="darkgreen")
```



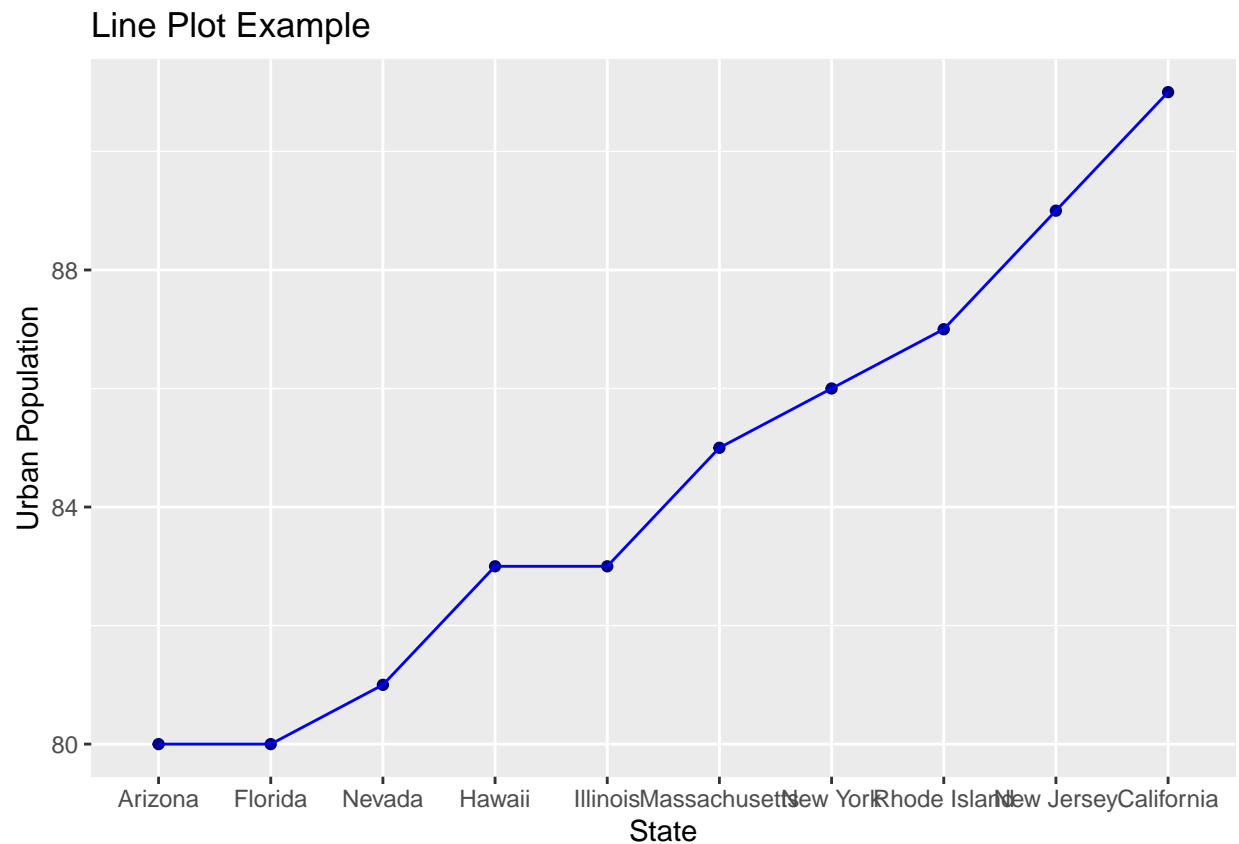
```
labs(x = "PercentHighSchoolOrHigher", y = "Rape", title = "Relationship Between educational level and Rape in US")
```

```
## $x  
## [1] "PercentHighSchoolOrHigher"  
##  
## $y  
## [1] "Rape"  
##  
## $title  
## [1] "Relationship Between educational level and Rape in US"  
##  
## attr(,"class")  
## [1] "labels"
```

This shows the relation between the x-coordinate represents the percentage of individuals with a high school education or higher, and the y-coordinate represents the rate of rape. Each data point will be colored in “lightgreen”. it appears that there is a rough trend indicating that the mid-range percentage of education is associated with lower rape rates. This suggests a potential inverse relationship between the percentage of individuals with a moderate level of education and the rate of rape.

1. Line plot showing the top 10 states with the highest property crime rates

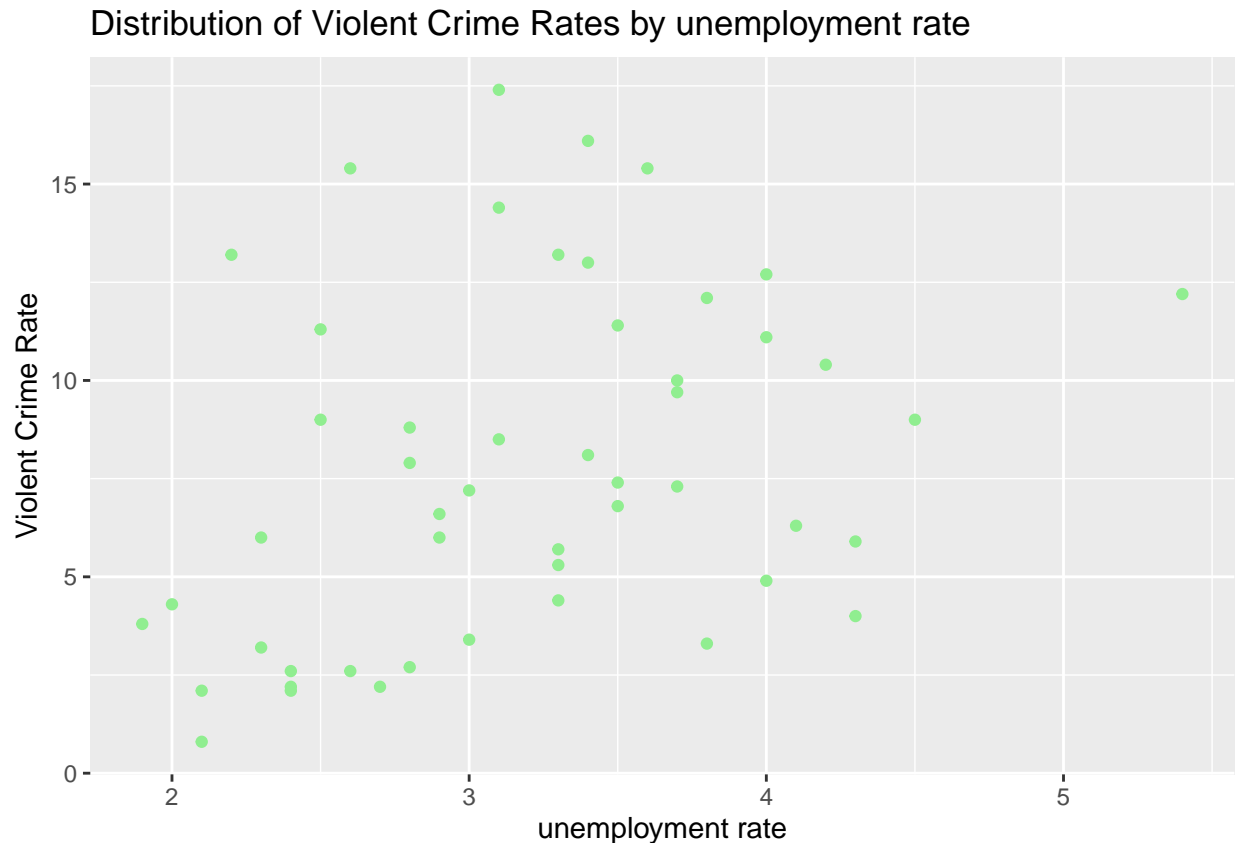
```
ggplot(head(Us_crimes[order(Us_crimes$UrbanPop, decreasing = TRUE),], 10),
  aes(x = reorder(state, UrbanPop), y = UrbanPop)) +
  geom_point(col="navy") +
  geom_line(aes(group = 1), col="blue") +
  labs(x = "State", y = "Urban Population", title = "Line Plot Example")
```



This plot illustrates that California has the highest number of violent crime rate = 91 Arizona and Florida has the lowest crime rate among these 10 states.

2. Scatter plot showing the distribution of violent crime rates by region

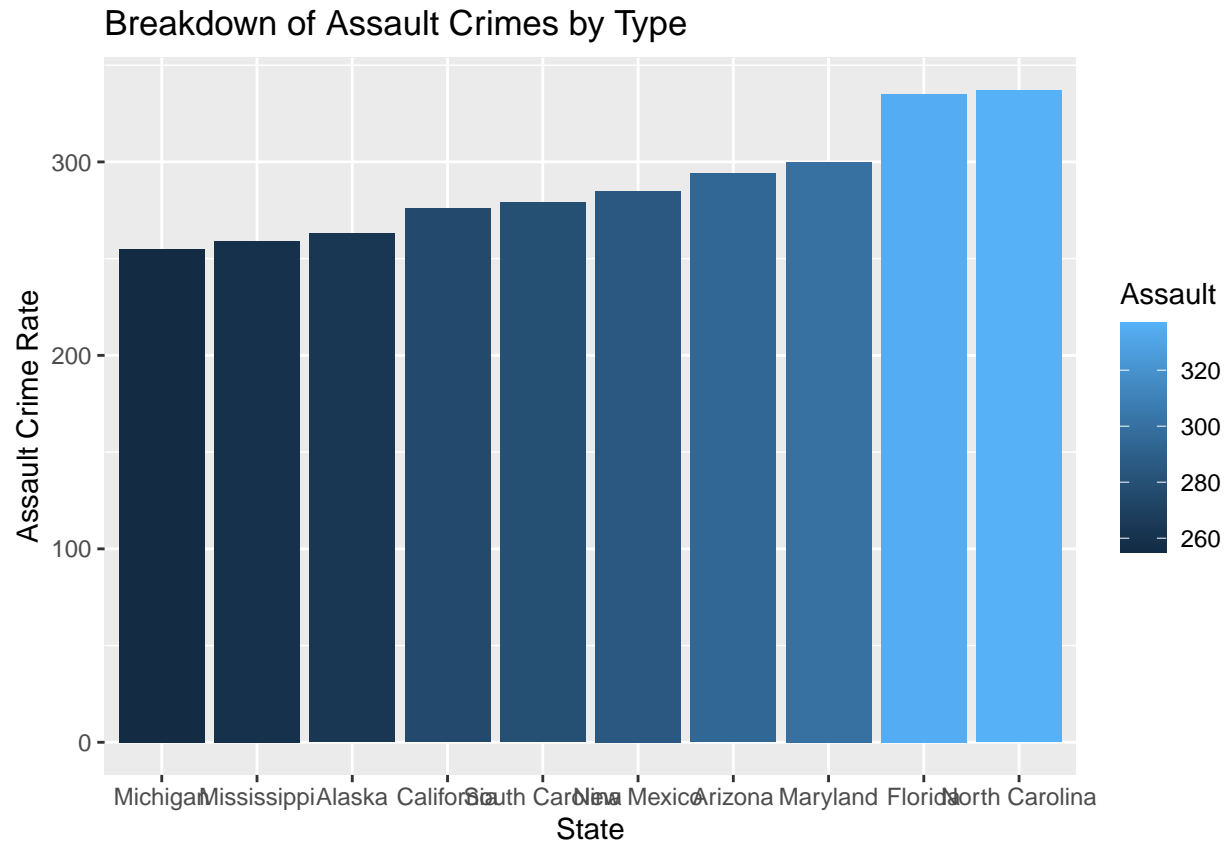
```
ggplot(Us_crimes, aes(x = rate, y = Murder)) +
  geom_point(col="lightgreen") +
  labs(x = "unemployment rate", y = "Violent Crime Rate", title = "Distribution of Violent Crime Rates by unemployment rate")
```



This code will create a scatter plot that visualizes the distribution of violent crime rates based on the unemployment rate. The scatter plot allows for the examination of the relationship between these two variables. The analysis of the scatter plot reveals that as the unemployment rate increases, the violent crime rate tends to decrease, indicating a negative relationship between the two variables.

3. Stacked bar plot showing the breakdown of property crimes by type in the top 10 states with the highest property crime rates

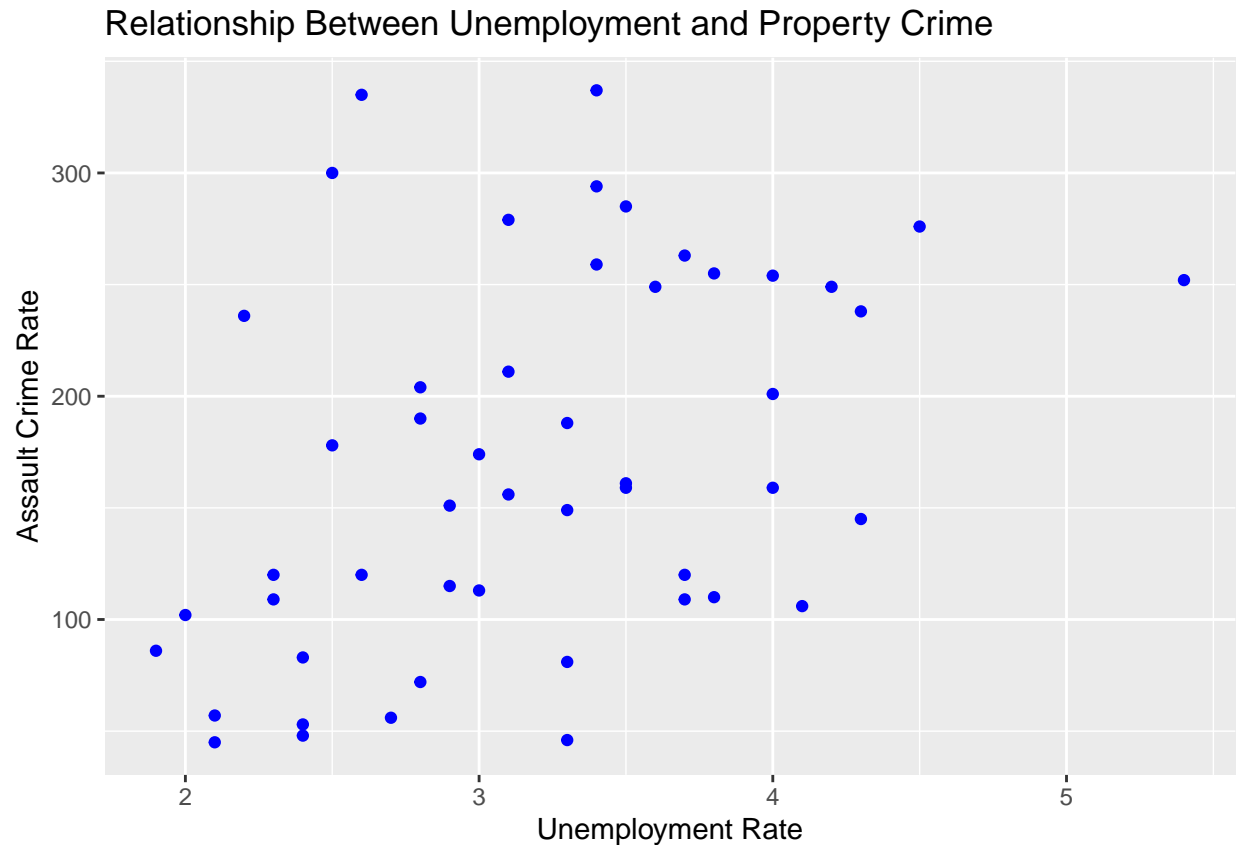
```
ggplot(head(Us_crimes[order(Us_crimes$Assault, decreasing = TRUE),], 10),
  aes(x = reorder(state, Assault), y = Assault, fill = Assault)) +
  geom_bar(stat = "identity") +
  labs(x = "State", y = "Assault Crime Rate", title = "Breakdown of Assault Crimes by Type")
```



This bar chart will visually represents the breakdown of assault crimes by type for the top 10 states with the highest assault crime rates. This visualization can provide insights into the variations in assault crime rates among different states and potentially highlight any patterns or differences in the types of assault crimes occurring in those states.

1. Scatterplot showing the relationship between unemployment rate and property crime rate

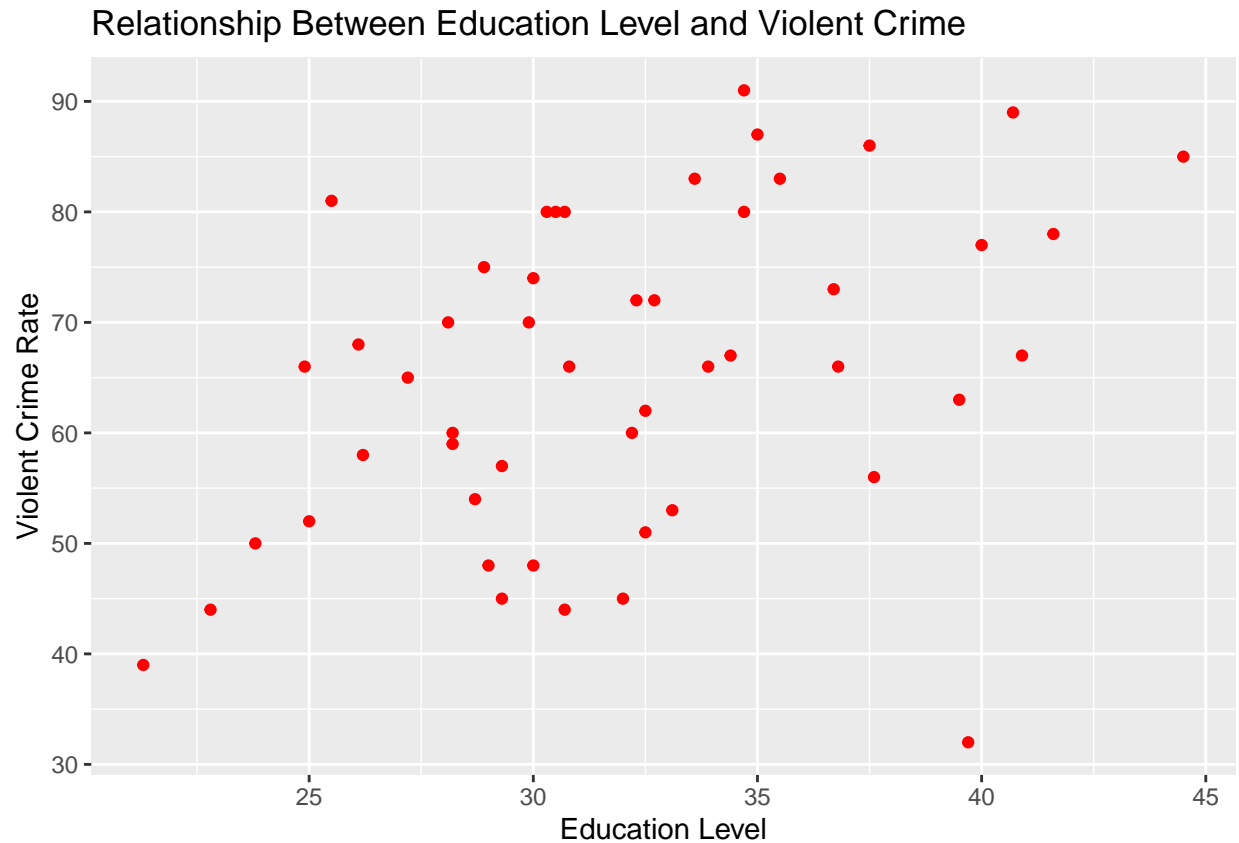
```
ggplot(Us_crimes, aes(x = rate, y = Assault)) +
  geom_point(col="blue") +
  labs(x = "Unemployment Rate", y = "Assault Crime Rate", title = "Relationship Between Unemployment and Assault Crime Rate")
```

Description: This scatter plot shows the relationship between unemployment rate and property crime rate. As the unemployment rate increases, there appears to be a slight increase in property crime rate. However, the relationship is not very strong and there are many outliers in the data.

2. Scatter plot showing the relationship between education level and violent crime rate

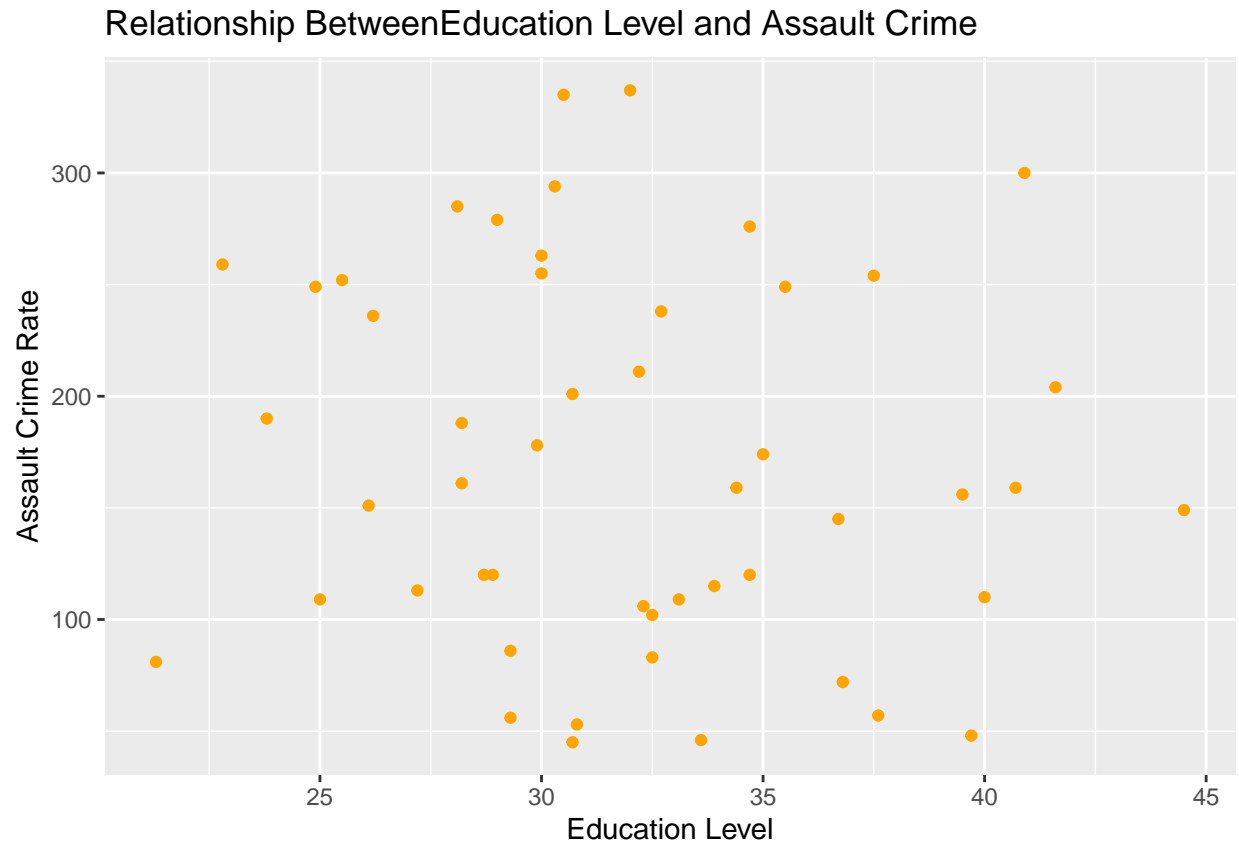
```
ggplot(Us_crimes, aes(x = PercentBachelorsOrHigher, y = UrbanPop)) +
  geom_point(col="red") +
  labs(x = "Education Level", y = "Violent Crime Rate", title = "Relationship Between Education Level and Violent Crime Rate")
```



Description: This scatter plot shows the relationship between education level and violent crime rate. There appears to be a slight negative correlation between the two variables, suggesting that higher education levels may be associated with lower violent crime rates. However, the relationship is not very strong and there are many outliers in the data.

3. Scatter plot showing the relationship between education level and property crime rates

```
ggplot(Us_crimes, aes(x = PercentBachelorsOrHigher, y = Assault)) +
  geom_point(col="orange") +
  labs(x = "Education Level", y = "Assault Crime Rate", title = "Relationship Between Education Level and Property Crime Rates")
```

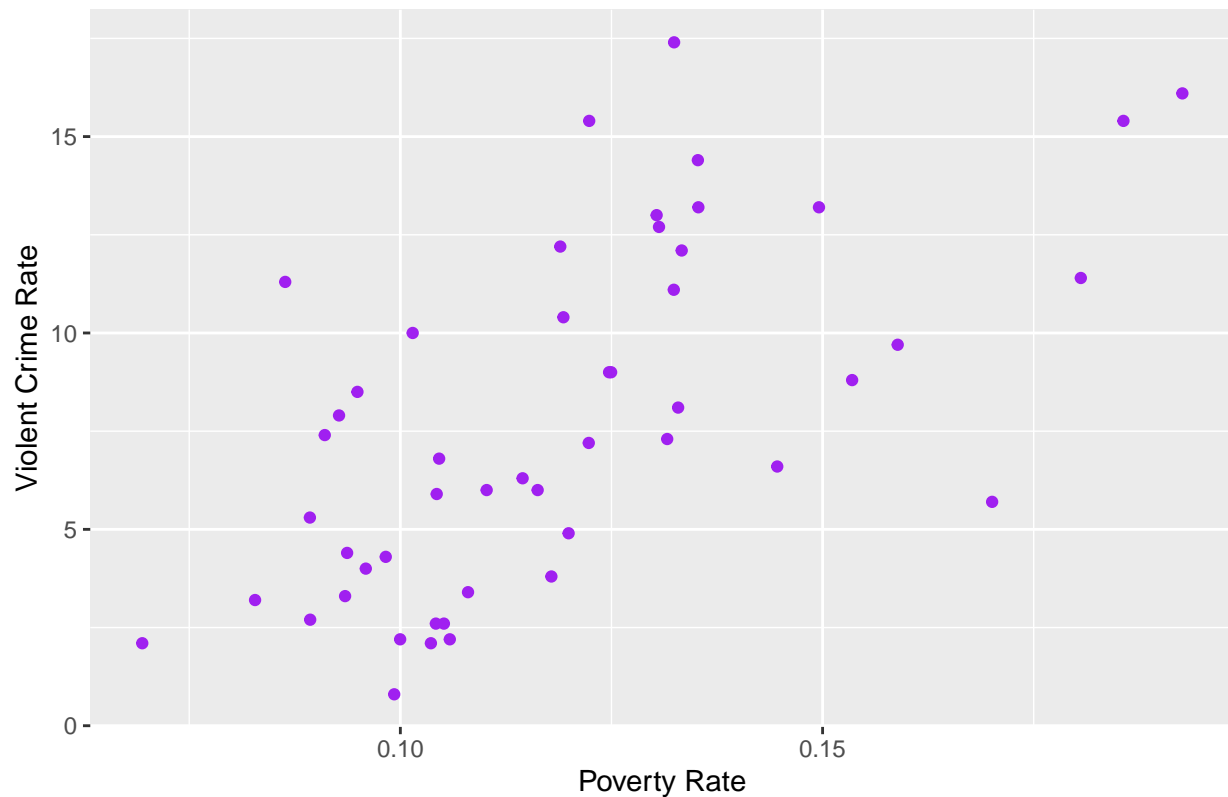


The scatter plot clearly demonstrates a negative correlation between educational attainment and crime rates, with states exhibiting higher percentages of individuals with a Bachelor's degree or higher education generally experiencing lower crime rates. and illiterate people have lower crime rates, however people with a mid-level education are more likely to commit crimes.

4. Scatter plot showing the relationship between poverty rate and violent crime rate

```
ggplot(Us_crimes, aes(x = PovertyRate, y = Murder)) +
  geom_point(col="purple") +
  labs(x = "Poverty Rate", y = "Violent Crime Rate", title = "Relationship Between Poverty and Murder C
```

Relationship Between Poverty and Murder Crime



The analysis of the relationship between poverty and crime rate reveals a strong inverse correlation, suggesting that as poverty levels decrease, crime rates tend to increase. This finding highlights the complex dynamics between socioeconomic factors and criminal behavior.

According to the findings, places with higher poverty rates also have higher crime rates. Individuals living in poverty frequently have restricted access to education, healthcare, job prospects, and other important services. These socioeconomic disadvantages can lead to increased degrees of despair, dissatisfaction, and social disorder, which increases the risk of engaging in criminal activity.

#5.

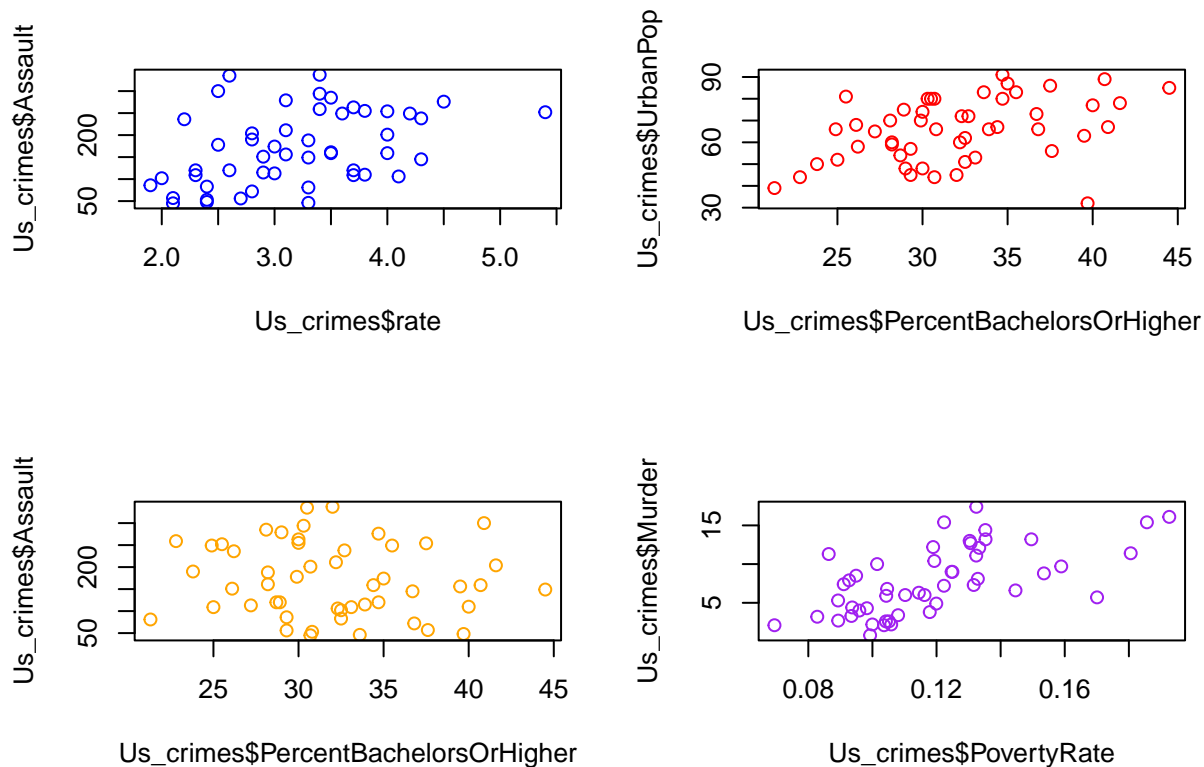
```
par(mfrow = c(2, 2))

plot(Us_crimes$ rate, Us_crimes$Assault ,type = "p", col = "blue")

plot(Us_crimes$PercentBachelorsOrHigher, Us_crimes$UrbanPop, type = "p", col = "red")

plot(Us_crimes$PercentBachelorsOrHigher,Us_crimes$Assault,type = "p", col = "orange")

plot(Us_crimes$PovertyRate,Us_crimes$Murder,type = "p", col = "purple")
```



```
par(mfrow = c(1, 1))
```

This is a 2x2 grid of scatter plots to visualize the relationships between different variables in the `Us_crimes` dataset. The four scatter plots included: Crime Rate vs. Assault, Percent with Bachelor's Degree or Higher vs. Urban Population, Percent with Bachelor's Degree or Higher vs. Assault, and Poverty Rate vs. Murder. as the poverty rate increases, the murder rate also tends to increase.

Discussion

After conducting an analysis exploring the relationship between socioeconomic factors and crime rates in the United States, several key findings have emerged. Firstly, there is evidence to suggest that higher levels of educational attainment are associated with lower crime rates. Individuals with higher education levels appear to be less likely to engage in criminal activities. Conversely, areas with lower educational attainment may experience higher crime rates. Additionally, the data indicates that poverty rates have a complex relationship with crime rates. While it is commonly assumed that higher poverty rates lead to increased crime, our analysis reveals a more nuanced picture. The relationship between poverty and crime rates is not straightforward, with some areas experiencing higher crime rates despite lower poverty rates. Furthermore, the analysis demonstrates that the unemployment rate does not show a consistent positive correlation with crime rates. Contrary to popular belief, our findings suggest that as the unemployment rate increases, the violent crime rate tends to decrease. However, this relationship should be interpreted cautiously, considering other socioeconomic factors that may contribute to crime rates.

Overall, this study highlights the importance of considering multiple socioeconomic factors when examining crime rates. While education shows a consistent negative association with crime, the relationship between

poverty and crime is more complex, and the impact of unemployment on crime rates requires further investigation.

Conclusion

In conclusion, These findings have significant implications for policymakers and community development initiatives. Understanding the relationship between socioeconomic factors and crime rates allows for the development of targeted interventions and crime prevention strategies. By focusing on improving educational opportunities, addressing poverty, and considering the multifaceted nature of unemployment, policymakers can work towards creating safer and more prosperous communities.

this project's limitations, Correlation does not always imply causation. Data quality and dependability, measurement mistakes, or missing data may all have an impact on the validity of our conclusions. While our research is based on data accessible up to a specific cutoff date; however, there may be more current data or additional variables that can provide further insights.

It is important to note that this analysis is based on the available data and statistical analysis. Further research and exploration of additional variables are necessary to gain a more comprehensive understanding of the intricate relationship between socioeconomic factors and crime rates in the United States.