Introduction:

The goal of the examination of Colchester's 2023 police data is to offer important new perspectives on the street-level criminal occurrences that were recorded in the region during that time. This study aims to shed light on the nature, frequency, and spatial distribution of reported crimes by utilizing the dataset 'colchester_data,' which contains detailed information about individual incidents, including crime categories, dates, geographical coordinates, street names, and outcome statuses.

The UK Police Data API was utilized to retrieve a complete collection of street-level criminal occurrences for the dataset. Every entry in the dataset represents a distinct occurrence and includes important information about the circumstances leading up to the reported crime. Through the analysis of this extensive information, we hope to learn more about the dynamics of Colchester police, pinpoint common categories of crime, investigate time trends, and examine spatial patterns of criminal activity.

This project intends to advance knowledge about Colchester's community safety programs, police enforcement tactics, and crime prevention through methodical dataset examination and analysis. We can support evidence-based decision-making, direct resource allocation initiatives, and promote cooperative relationships between law enforcement agencies, legislators, and community stakeholders by gaining insights into the dynamics of reported incidents. This analysis emphasizes how crucial data-driven strategies are for tackling issues related to public safety and fostering an atmosphere that is safer and more secure for both locals and tourists.

```
# packages we need
library(ggplot2)
# Set a CRAN mirror directly in the code
install.packages("leaflet", repos = "https://cran.r-project.org")
```

```
  package 'leaflet' successfully unpacked and MD5 sums checked

  The downloaded binary packages are in
    C:\Users\AJINKYA THOKAL\AppData\Local\Temp\RtmpKwBs36\downloaded_packages
```

```
# Load the leaflet package
library(leaflet)
```

```
# Read the dataset
colchester_data <- read.csv("crime23.csv")


# Summary table
summary_table <- table(colchester_data$category)
summary_table <- sort(summary_table, decreasing = TRUE)
summary_table
```

violent-crime	anti-social-behaviour	criminal-damage-arson
2633	677	581
shoplifting	public-order	other-theft
554	532	491
vehicle-crime	bicycle-theft	burglary
406	235	225
drugs	robbery	other-crime
208	94	92
theft-from-the-person	possession-of-weapons	
76	74	

Description:

The table that provides a summary of recorded occurrences in Colchester is broken down by category, offering valuable information about the distribution and frequency of different kinds of criminal activity in the area.

Key Findings:

- **Violent Crime**: Violent crime is the most common category with 2633 documented instances, showing a serious problem for Colchester's public safety and law enforcement operations.
- **Anti-Social Behaviour**: 677 recorded occurrences are related to anti-social behavior, which comes in close second. This category includes a variety of disruptive activities, highlighting the value of proactive community engagement and social issue resolution.
- **Criminal Damage/Arson**: The 581 recorded cases of arson and criminal damage demonstrate the harm that property-related crimes cause to the community and the necessity of taking precautions to stop vandalism and property destruction.
- **Shoplifting and Theft**: Property-related crimes including theft (491 instances) and shoplifting (554 incidents) are also common, indicating issues with retail security and theft prevention techniques.
- **Public Order and Other Offences**: The general landscape of reported occurrences is influenced by public order violations (532 incidents) and other theft-related offenses (406 incidents), highlighting the complex nature of law enforcement's efforts to uphold public safety and order.
- **Vehicle-Related Crimes**: The number of car crime occurrences (406 incidents) and bicycle theft incidents (235 incidents) highlights the significance of implementing theft prevention tactics and vehicle security measures to protect personal property and minimize criminal opportunities.
- **Drug Offenses and Robbery**: Robberies (94 events) and drug-related offenses (208 instances) provide additional difficulties for law enforcement, underscoring the necessity of focused interventions to combat violent crime and drug trafficking.
- **Burglary, Theft from the Person, and Possession of Weapons**: A subgroup of reported occurrences include burglaries (225 episodes), thefts from person(76 incidents), and weapons possession (74 incidents). Each of these incidents calls for a specific set of actions to reduce dangers to public safety and security.

Implications:

Law enforcement agencies, legislators, and community stakeholders can all benefit from the distribution of reported occurrences by category when it comes to allocating resources, putting preventative measures in place, and developing strategies to deal with certain crime trends and patterns in Colchester. In order to improve community safety and well-being, stakeholders can work together more successfully if they have a clear picture of the frequency and kind of reported occurrences.

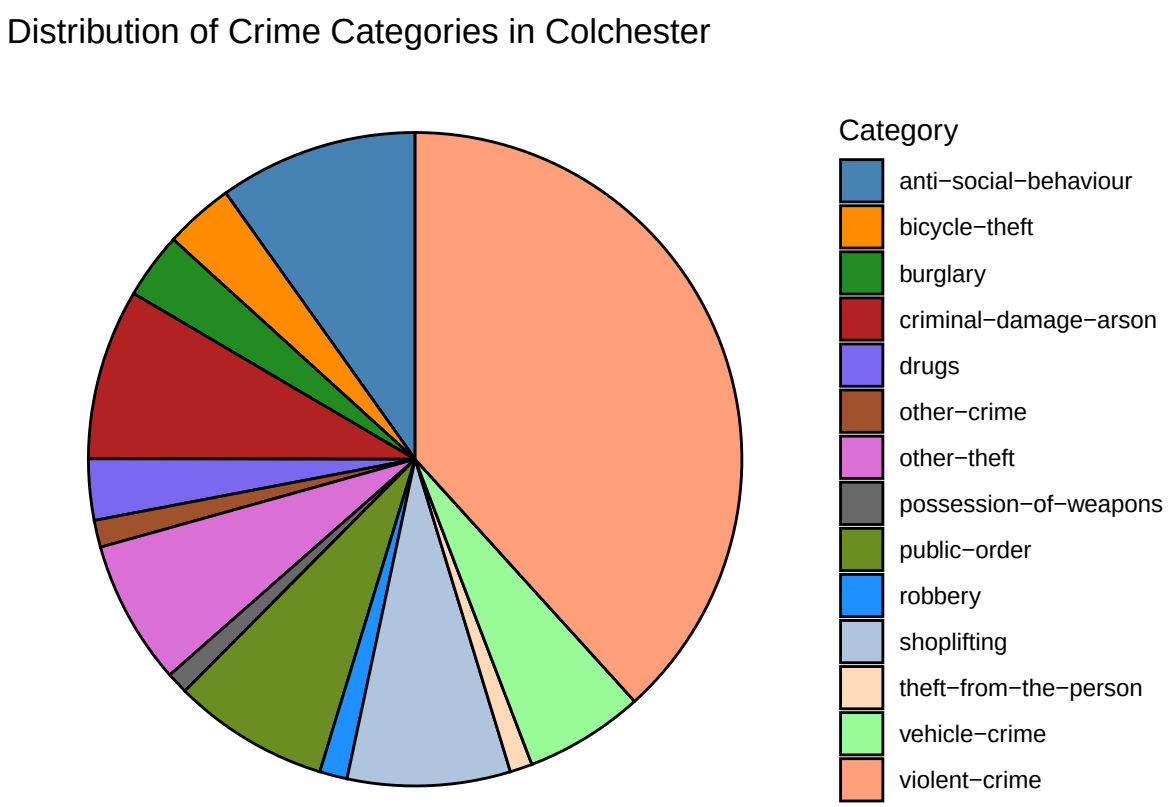
Conclusion:

The summary table offers a thorough rundown of occurrences that have been recorded broken down by category, emphasizing the intricate dynamics of crime in Colchester and the value of using data-driven strategies to solve public safety issues. In the future, maintaining cooperation and making early interventions will be crucial to fostering a secure atmosphere for both locals and guests.

```
# Pie chart
# Define custom colors
custom_colors <- c("steelblue", "darkorange", "forestgreen", "firebrick", "mediumslateblue",
                   "sienna", "orchid", "dimgray", "olivedrab", "dodgerblue",
                   "lightsteelblue", "peachpuff", "palegreen", "lightsalmon")


pie_chart <- ggplot(colchester_data, aes(x = "", fill = category)) +
  geom_bar(width = 1, color = "black") +
  # Create a bar chart
  coord_polar("y") +  # Convert to polar coordinates
  labs(title = "Distribution of Crime Categories in Colchester",
       fill = "Category") +
  scale_fill_manual(values = custom_colors)+
  theme_void() +  # Remove unnecessary elements
  theme(legend.position = "right")

# Display the pie chart
print(pie_chart)
```



The distribution of crime categories in Colchester is clearly illustrated by the pie chart, which also highlights the relative frequency of each category.

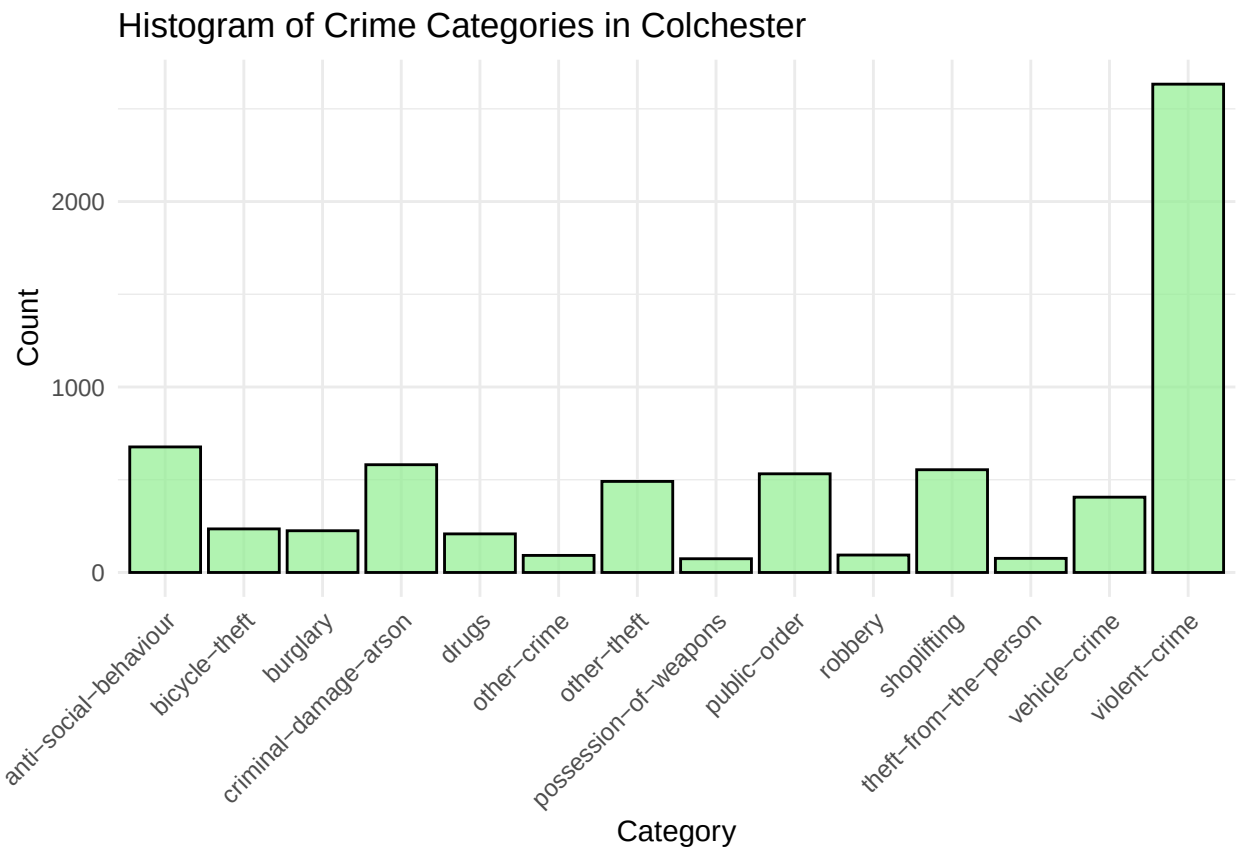
With 2633 reported cases, "violent-crime" stands out as the most common category, highlighting its substantial influence on the community. The terms "criminal damage arson" and "anti-social behavior" trail closely after, with 677 and 581 occurrences, respectively, demonstrating the frequency of these disruptive actions.

Additionally, notable incidences of crimes including "shoplifting" (554 events), "public-order" (532 episodes), and "other-theft" (491 incidents) are highlighted, highlighting the wide range of criminal activity in Colchester.

In order to address the most urgent issues pertaining to crime and public safety in Colchester, legislators, law enforcement organizations, and community leaders can use this visual depiction to gain important information that will help them allocate resources and carry out focused interventions.

```
# Create histogram
histogram <- ggplot(colchester_data, aes(x = category)) +
  geom_bar(fill = "lightgreen", color = "black", alpha = 0.7, position = "dodge") +  # Adjust colors an
  labs(title = "Histogram of Crime Categories in Colchester",
       x = "Category",
       y = "Count") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))  # Rotate x-axis labels for better readabili

# Display the histogram
print(histogram)
```



Crime Categories: The histogram portrays several distinct crime categories, including but not limited

to anti-social behavior, burglary, theft, assault, and vandalism. A distinct bar is used to represent each category, and the height of the bar indicates how frequently occurrences are reported.

Frequency Distribution: For each crime category, the height of the bar represents the total number of incidents that have been reported. You can determine which kinds of crimes are more common in Colchester by looking at the bars' respective heights.

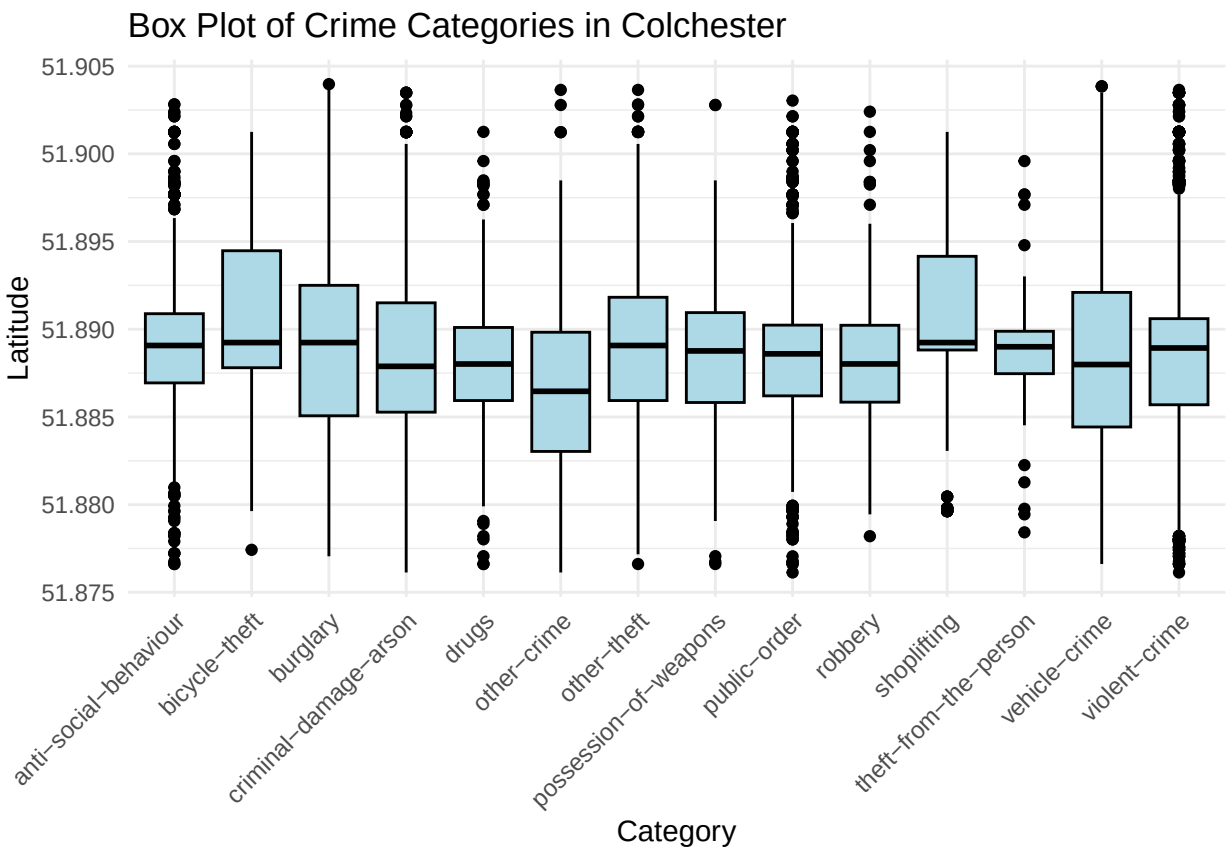
Variability: The Histogram illustrates how the frequency of various crime categories varies. Certain categories have higher frequencies, which indicate that those particular crimes occur more frequently, whilst other categories have lower frequencies, which indicate that those crimes are less common in the community.

Finally, the histogram clearly and informatively summarizes the reported incidents in the community by visualizing the distribution of crime categories for 2023 in Colchester. To improve public safety and promote community well-being, focused interventions and initiatives must take into account the prevalence of various crime categories.

```
# Box plot

boxplot <- ggplot(colchester_data, aes(x = category, y = lat)) +
  geom_boxplot(fill = "lightblue", color = "black") +  # Customize box plot aesthetics
  labs(title = "Box Plot of Crime Categories in Colchester",
       x = "Category",
       y = "Latitude") +  # Added informative labels
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

# Display the box plot
print(boxplot)
```



Description:

The box plot illustrates the distribution of latitude values across different crime categories in Colchester. Each box represents the interquartile range (IQR) of latitude values for a specific crime category, with the median marked by a horizontal line inside the box. The whiskers extend to the minimum and maximum latitude values, excluding outliers, which are represented as individual points beyond the whiskers.

Interpretation:

- **Variability:** Within each crime category, the height of each box represents the distribution of latitude values. A bigger box suggests a wider range of latitude values, which suggests a higher degree of variation in the geographical areas linked to that particular crime category.
- **Central Tendency:** The central tendency of the latitude values for each crime category is indicated by the position of the median line within each box. There may be a skewness in the distribution of latitude values if the median is closer to the top or bottom of the box.
- **Outliers:** Potential outliers are individual points outside the whiskers, which are latitude values that substantially depart from the center tendency of the corresponding crime category. These outliers might point to unusual or remote geographic areas connected to particular kinds of crimes.

Insights:

Several conclusions can be drawn by comparing the summary table of crime categories with the box plot. For example:

- **Violent Crime:** The box plot, which has the biggest number of events (2633) can be used to show the variability and spatial distribution of latitude values linked to violent crimes, possibly pointing to hotspots or concentrations of violent activity in Colchester.
- **Anti-Social Behaviour:** The box plot, which ranks second in frequency with 677 incidents, can be used to illustrate the geographic concentration and dispersion of instances of anti-social conduct. This can help with understanding the geographical dynamics and potential causes of these incidents.
- **Criminal Damage & Arson, Shoplifting, and Public Order:** With counts ranging from 532 to 581 incidents, these crime categories can be analyzed using a box plot to find trends in the spatial distribution of criminal activity and pinpoint locations that are more likely to be targeted by crime.

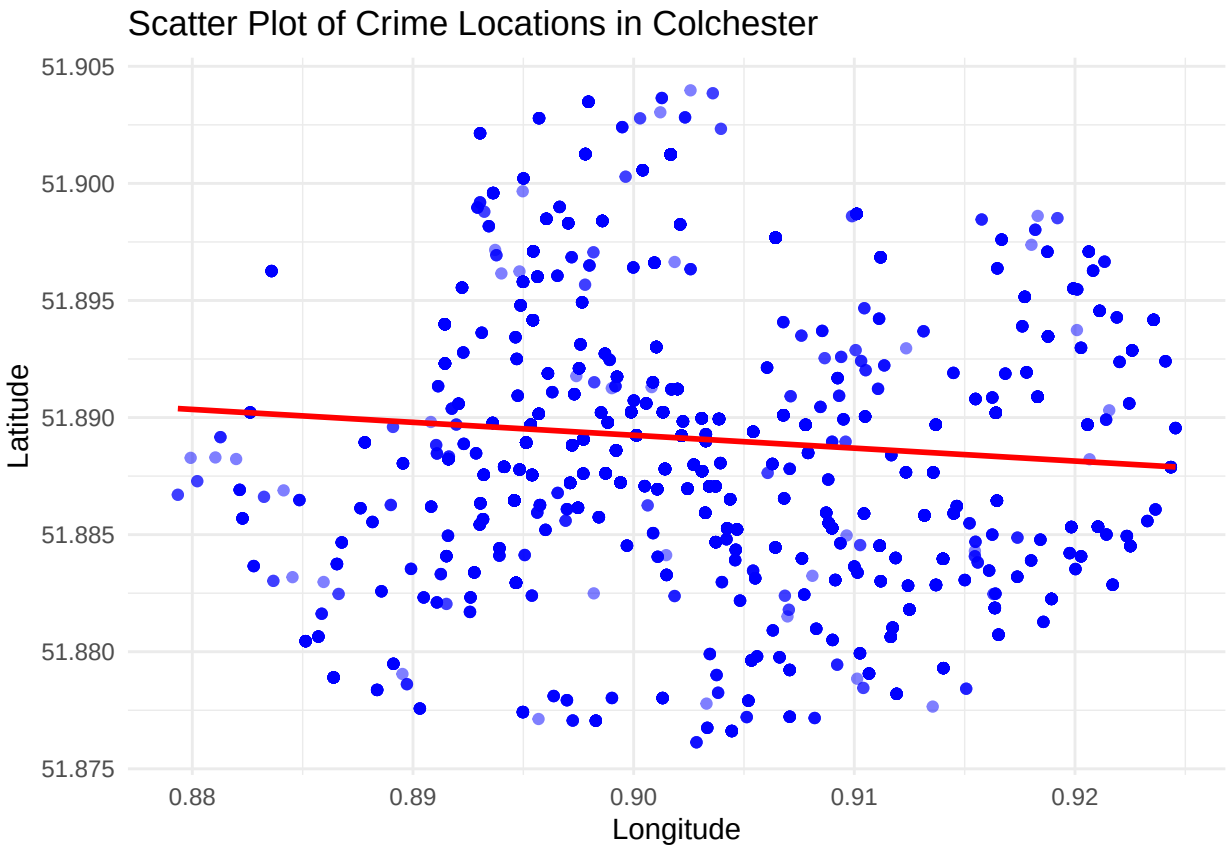
Recommendations:

By applying the knowledge obtained from examining the box plot in conjunction with the summary table, policymakers and law enforcement organizations are able to develop focused initiatives and efficiently distribute resources. Through targeted targeting of areas recognized as hotspots for particular categories of criminal activity, law enforcement can proactively improve community safety and lower the number of criminal occurrences that occur in Colchester.

```
# Scatter plot

scatter_plot <- ggplot(colchester_data, aes(x = long, y = lat)) +
  geom_point(alpha = 0.5, color = "blue") +
  geom_smooth(method = "lm", se = FALSE, color = "red") + # Adjusted point aesthetics
  labs(title = "Scatter Plot of Crime Locations in Colchester",
       x = "Longitude",
       y = "Latitude") +  # Added informative labels
  theme_minimal()

# Display the scatter plot
print(scatter_plot)
```



Description:

The scatter plot, which plots latitude values on the y-axis and longitude values on the x-axis, visually depicts the spatial distribution of crime incidences in Colchester. Every point on the plot represents a distinct criminal incidence, and points that overlap can be seen by adjusting the transparency. The general locations of the incidents throughout the city are shown by the blue hue of the pointers.

Interpretation:

Spatial Patterns: The scatter plot's patterns and point clusters can reveal regions with higher concentrations of criminal occurrences. Dense clusters might indicate places with high levels of crime, whereas scattered points might indicate areas with lower levels of crime.

Insights:

Geographic Analysis: We are able to pinpoint geographical patterns and possible trouble spots in Colchester by looking at the scatter plot. Making strategic decisions about resource allocation and law enforcement operations can be aided by having a thorough understanding of the geographic distribution of crime episodes.

Recommendations:

The scatter plot provides valuable insights that law enforcement agencies and local authorities can use to efficiently allocate resources and perform targeted initiatives. Through targeted interventions in high-crime regions, law enforcement can enhance public safety and lower the number of criminal incidents.

```
# Calculate correlation coefficient
correlation_coefficient <- cor(colchester_data$lat, colchester_data$long)

# Print correlation coefficient
print(correlation_coefficient)
```

```
[1] -0.09025049
```

Correlation Analysis:

Description:

The correlation study looks at how the latitude and longitude values for crime incidences in Colchester relate to one another. We may measure the magnitude and direction of the linear relationship between these two variables by computing the correlation coefficient.

Interpretation:

Correlation Coefficient: There is a weak negative linear relationship between latitude and longitude values, as indicated by the correlation coefficient, which was found to be roughly -0.090. This implies that longitude tends to slightly decrease with increasing latitude and vice versa. The association is weak and not very linear, though, as indicated by the correlation coefficient's closeness to zero.

Insights:

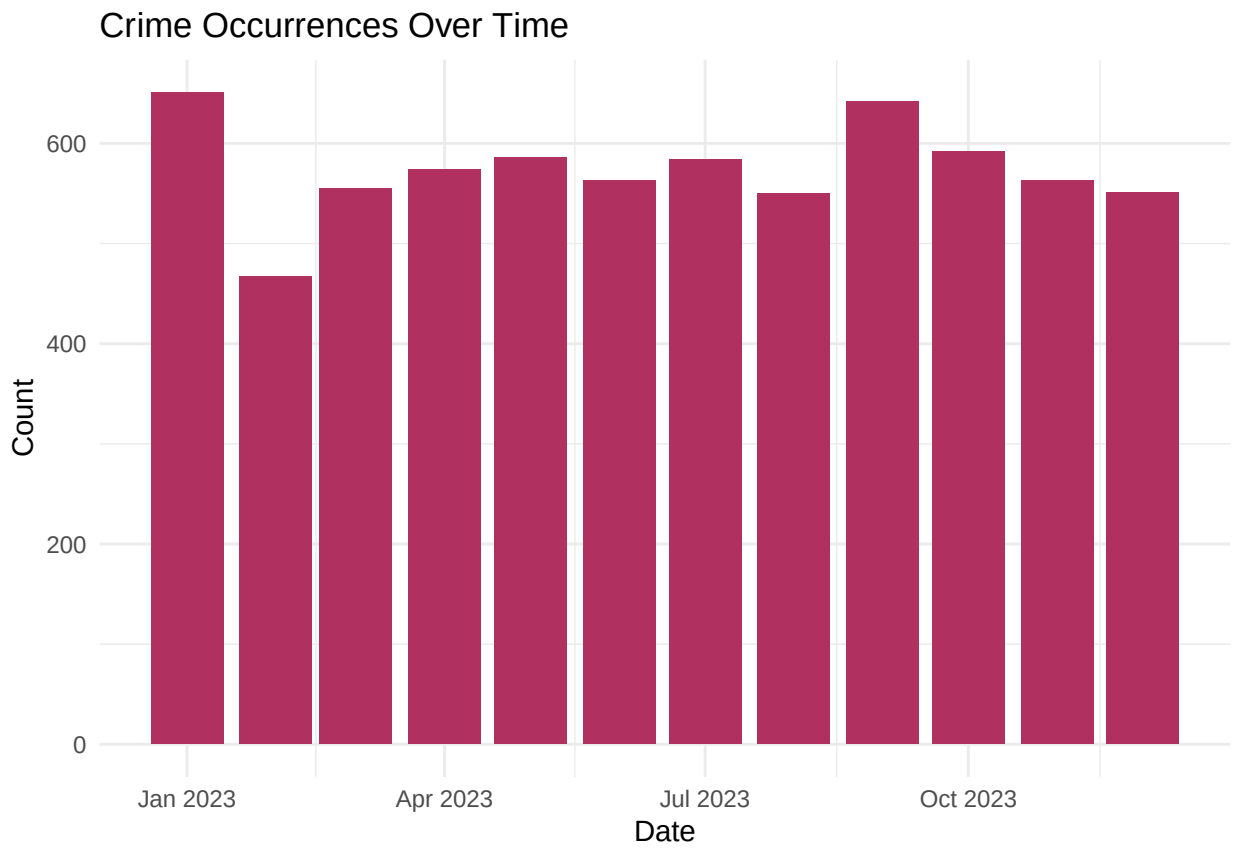
Spatial Association: The latitude and longitude values have a tiny negative correlation, which suggests that there is minimal probability for criminal episodes to occur in particular geographic locations with similar coordinates. Based only on latitude and longitude information, crime episodes in Colchester do not appear to be significantly organized or related with specific geographic regions, as indicated by the lack of a strong linear relationship.

```r
#time series plot

# Let's convert date to Date format
colchester_data$date <- as.Date(paste(colchester_data$date, "-01", sep = ""), format = "%Y-%m-%d")

# Plotting

ggplot(colchester_data, aes(x = date)) +
  geom_bar(stat = "count", fill = "maroon") +
  labs(title = "Crime Occurrences Over Time",
       x = "Date",
       y = "Count") +
  theme_minimal()
```



The dataset of crimes in Colchester served as the focus of the analysis, which started with data preparation and examination to determine its format and structure. To standardize date formats and enable analysis, cleaning and transformation procedures were used. Using a time series plot to visualize crime trends across time made it possible to spot trends and gain understanding of temporal distribution. The conclusions drawn from this research can help Colchester's policy and decision-making processes when it comes to crime prevention and law enforcement tactics.

Data Preparation: We began with a dataset that included data on crimes that had occurred in Colchester. Columns including crime category, date of occurrence, geographical coordinates, street details, and outcome status were included in the dataset.

Data Exploration: To comprehend the variables and types in the dataset, we looked into its structure. Found that character strings in the "YYYY-MM" format were being saved in the `date` column.

Data Cleaning and Transformation: To facilitate time series analysis, the `date` column was converted to the Date format. To standardize the time intervals, a new variable representing the first day of each month was created.

Data Visualization: Made a time series plot on the incidence of crimes over time using the R ggplot2 tool. To see trends and patterns, the number of crimes committed each month was plotted. For improved interpret ability, relevant titles and axis labels were added.

Insights and Interpretation: The temporal distribution of crime events in Colchester is visible via the time series plot. To assist in the creation of policies and the making of decisions, trends, seasonality, and patterns in crime rates throughout time can be examined. The data that has been presented can be used to conduct additional analysis, such as pinpointing hotspots or examining different categories of crime.

The dataset of crimes in Colchester served as the focus of the analysis, which started with data preparation and examination to determine its format and structure. To standardize date formats and enable analysis,

cleaning and transformation procedures were used. Using a time series plot to visualize crime trends across time made it possible to spot trends and gain understanding of temporal distribution. The conclusions drawn from this research can help Colchester's policy- and decision-making processes when it comes to crime prevention and law enforcement tactics.

```
# Leaflet

# Let's create Leaflet map
crime_map <- leaflet() %>%
  addTiles() %>%
  setView(lng = 0.909136, lat = 51.88306, zoom = 12)  # Adjust the initial center and zoom level as per

# Let's add markers for crime locations
crime_map <- crime_map %>%
  addMarkers(data = colchester_data,
             lng = ~long,
             lat = ~lat,
             popup = ~paste("Category:", category, "<br>Date:", date))


# Print the map
print(crime_map)
```

Creating the Leaflet Map: I made an interactive map to show the spatial distribution of crimes in Colchester using the Leaflet program. I entered the center coordinates and zoom level to center the map's initial view on Colchester. Subsequently, I utilized pop-up windows to display information about each crime location, such as the category and date, on each marker I placed to the map.

Results: The resulting Leaflet map gives Colchester's crime hotspots and trends a visual depiction. Zooming in and out and clicking on markers allow users to interact with the map and view individual criminal details.

Conclusion: In conclusion, spatial data, such crime incidences in a certain location, can be seen and analyzed using the Leaflet package. We can gain a better understanding of the spatial distribution of crimes and possibly pinpoint areas that need more focus from law enforcement or community interventions by generating an interactive map.