Colchester Crime(2023) Analysis

Musaddik Maulavi

2024-04-18

INTRODUCTION

The Colchester crime dataset for 2023 offers important information about the frequency of street-level criminal acts in the community. This dataset, which was taken from the UK Police interface, includes comprehensive data on a range of crimes that have been recorded all year long. The dataset has variables for every item, including the kind of crime, date of occurrence, street name, location type, latitude and longitude, and result status, among others.

The goal of this dataset's study is to find patterns, trends, and correlations in the crime data so that we may better understand Colchester's crime scene throughout the given time frame. Through the application of statistical analysis and data visualization tools, we are able to determine crime hotspots, investigate correlations between various factors, and find periodic trends in crime occurrences.

This report will conduct a thorough examination of the Colchester policing dataset in this paper using a range of statistical and data visualization approaches. I want to give useful insights into the dynamics of crime in the Colchester region in 2023 by looking at the distribution of crime categories, utilizing maps to explore geographical patterns, assessing temporal trends through time series plots, and looking into relationships between factors. In order to improve the way our findings are presented and provide an interesting and educational look into the dataset, we will also make use of sophisticated visuals and interactive components.

EXPLORING DATASET

The dataset comprises 6878 observations with 12 variables, providing detailed information on street-level crime incidents reported in Colchester during the year 2023.

This dataset offers a comprehensive view of the crime landscape in Colchester, allowing for analysis and exploration of various aspects of street-level crime incidents occurring throughout the year 2023. Through the utilization of data visualization techniques and statistical analysis, report aim to derive meaningful insights and patterns from the data to better understand the dynamics of crime within the area.

```
# Importing the libraries
crimeData <- read.csv('crime23.csv')

# Printing first few rows of our dataset
head(crimeData)
```

```
##                category persistent_id    date      lat     long street_id
## 1 anti-social-behaviour               2023-01 51.88306 0.909136   2153366
## 2 anti-social-behaviour               2023-01 51.90124 0.901681   2153173
## 3 anti-social-behaviour               2023-01 51.88907 0.897722   2153077
## 4 anti-social-behaviour               2023-01 51.89122 0.901988   2153186
## 5 anti-social-behaviour               2023-01 51.89416 0.895433   2153012
## 6 anti-social-behaviour               2023-01 51.88050 0.909014   2153379
##                      street_name context        id location_type
## 1      On or near Military Road      NA 107596596         Force
## 2                    On or near      NA 107596646         Force
## 3 On or near Culver Street West      NA 107595950         Force
## 4       On or near Ryegate Road      NA 107595953         Force
## 5       On or near Market Close      NA 107595979         Force
## 6         On or near Lisle Road      NA 107595985         Force
##   location_subtype outcome_status
## 1                            <NA>
## 2                            <NA>
## 3                            <NA>
## 4                            <NA>
## 5                            <NA>
## 6                            <NA>
```

```
str(crimeData)
```

```
## 'data.frame':    6878 obs. of  12 variables:
##  $ category        : chr  "anti-social-behaviour" "anti-social-behaviour" "anti-social-behaviour"
"anti-social-behaviour" ...
##  $ persistent_id   : chr  "" "" "" "" ...
##  $ date            : chr  "2023-01" "2023-01" "2023-01" "2023-01" ...
##  $ lat             : num  51.9 51.9 51.9 51.9 51.9 ...
##  $ long            : num  0.909 0.902 0.898 0.902 0.895 ...
##  $ street_id       : int  2153366 2153173 2153077 2153186 2153012 2153379 2153105 2153541 2152937 2
153107 ...
##  $ street_name     : chr  "On or near Military Road" "On or near " "On or near Culver Street West"
"On or near Ryegate Road" ...
##  $ context         : logi  NA NA NA NA NA NA ...
##  $ id              : int  107596596 107596646 107595950 107595953 107595979 107595985 107596603 107
596291 107596305 107596453 ...
##  $ location_type   : chr  "Force" "Force" "Force" "Force" ...
##  $ location_subtype: chr  "" "" "" "" ...
##  $ outcome_status  : chr  NA NA NA NA ...
```

CLEANING AND MANIPULATING DATASET

Before Visualization, there is a desperate need of cleaning the dataset. For this I have printed the NA values in each column. As you can see in output below all the values in context is NA. So I have removed that column. Also ‘date’ column is in string format, I have converted in date format using ‘anytime’ library and ‘yearmon’ function so the resultant format can be used easily for plotting graphs.

```
# Printing the null values in each column.
na_count <- colSums(is.na(crimeData))
print(na_count)
```

```
##         category    persistent_id             date              lat
##                0                0                0                0
##             long        street_id      street_name          context
##                0                0                0             6878
##               id    location_type location_subtype   outcome_status
##                0                0                0              677
```

```
# Coverting the 'date' column into DATE format which was previously in 'string' using 'anydate' and 'y
earmon' library
Sys.setlocale("LC_TIME", "C")
```

```
## [1] "C"
```

```
crimeData$date <- anydate(crimeData$date)
crimeData$date <- as.yearmon(crimeData$date)
```

Also there are some entries in 'outcome_status' which are NA. So I have replaced it with "Investigation complete; no suspect Identified". Then I have again checked the NA values in every column as you can see below code and output.

```
# unique(crimeData$outcome_status)

crimeData$outcome_status[is.na(crimeData$outcome_status)] <- "Investigation complete; no suspect ident
ified"

# Checking the null values again
na_count <- colSums(is.na(crimeData))
print(na_count)
```

```
##         category    persistent_id             date              lat
##                0                0                0                0
##             long        street_id      street_name          context
##                0                0                0             6878
##               id    location_type location_subtype   outcome_status
##                0                0                0                0
```

In the below code I have cleaned the 'street_name', as there are many incomplete entries so I have replaced it with 'Location Unidentified'. I have removed the 'context' and 'persistent_id' as persistent id has the unique IDs which wont be helpful for out Data Visualization.

```
head(crimeData$street_name,10)
```

```
##  [1] "On or near Military Road"      "On or near "
##  [3] "On or near Culver Street West" "On or near Ryegate Road"
##  [5] "On or near Market Close"       "On or near Lisle Road"
##  [7] "On or near "                   "On or near Gantry Close"
##  [9] "On or near Stable Road"        "On or near Trinity Street"
```

```
# Replacing incomplete entries with 'Location unidentified'
crimeData$street_name <- gsub("^On or near $", "Location unidentified", crimeData$street_name)


# All the entries in 'Context' column is NA. Therefore we can remove this column.

crimeData <- subset(crimeData, select = -context)
crimeData <- subset(crimeData, select = -persistent_id)
```

DATA VISUALISATION

A. TABLE

```
# Table for 'category' column
print(table(crimeData$category))
```

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

This table gives a insight about the number of times the particular crime has occurred in the year 2023 in Colchester. `Violent crime` has occurred for highest number of time, whereas `possession-of-weapon` type of crime has occured for least number of times.

B. TWO-WAY TABLE

```
# Two-way table for category column and outcome_status
print(table(crimeData$category, crimeData$outcome_status))
```

```
##
##                         Action to be taken by another organisation
##   anti-social-behaviour                                          0
##   bicycle-theft                                                  0
##   burglary                                                       0
##   criminal-damage-arson                                          5
##   drugs                                                          2
##   other-crime                                                    1
##   other-theft                                                    1
##   possession-of-weapons                                          1
##   public-order                                                   6
##   robbery                                                        0
##   shoplifting                                                    5
##   theft-from-the-person                                          0
##   vehicle-crime                                                  0
##   violent-crime                                                 83
##
##                         Awaiting court outcome Court result unavailable
##   anti-social-behaviour                      0                         0
##   bicycle-theft                              0                         1
##   burglary                                   1                        15
##   criminal-damage-arson                     31                        22
##   drugs                                     18                        17
##   other-crime                                6                         3
##   other-theft                                2                         6
##   possession-of-weapons                      9                        10
##   public-order                              18                        17
##   robbery                                    4                         4
##   shoplifting                               51                        45
##   theft-from-the-person                      0                         0
##   vehicle-crime                              6                         2
##   violent-crime                            114                        64
##
##                         Formal action is not in the public interest
##   anti-social-behaviour                                            0
##   bicycle-theft                                                    0
##   burglary                                                         0
##   criminal-damage-arson                                            0
##   drugs                                                            1
##   other-crime                                                      0
##   other-theft                                                      1
##   possession-of-weapons                                            0
##   public-order                                                     2
##   robbery                                                          0
##   shoplifting                                                      1
##   theft-from-the-person                                            0
##   vehicle-crime                                                    0
##   violent-crime                                                    4
##
##                         Further action is not in the public interest
##   anti-social-behaviour                                             0
##   bicycle-theft                                                     0
##   burglary                                                          2
##   criminal-damage-arson                                             2
##   drugs                                                            10
##   other-crime                                                       6
##   other-theft                                                       1
##   possession-of-weapons                                             1
##   public-order                                                     12
##   robbery                                                           0
```

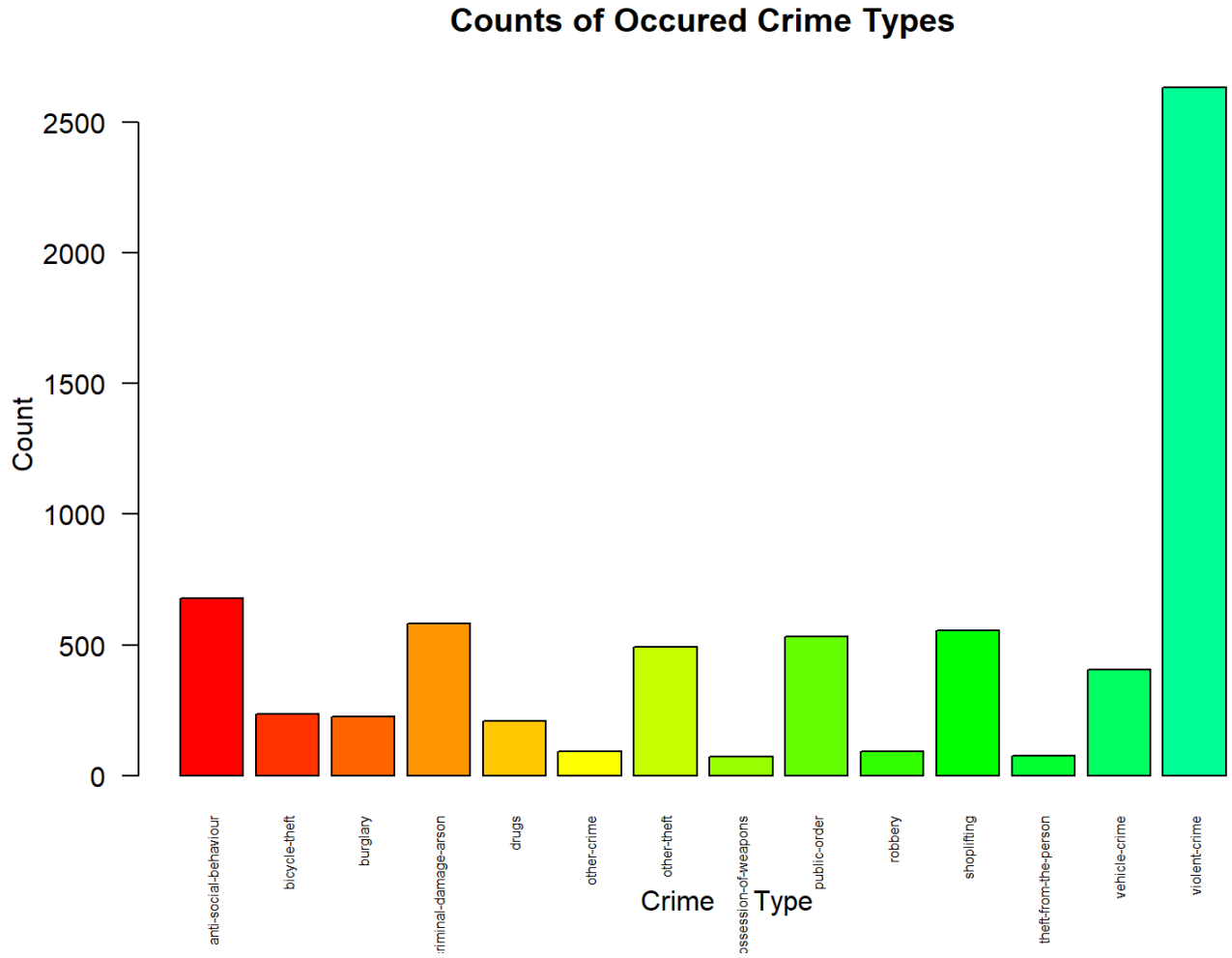
```
##   shoplifting                                                    11
##   theft-from-the-person                                           0
##   vehicle-crime                                                   0
##   violent-crime                                                  37
##
##                        Further investigation is not in the public interest
##   anti-social-behaviour                                                   0
##   bicycle-theft                                                           0
##   burglary                                                                0
##   criminal-damage-arson                                                   0
##   drugs                                                                   0
##   other-crime                                                             8
##   other-theft                                                             0
##   possession-of-weapons                                                   0
##   public-order                                                            0
##   robbery                                                                 0
##   shoplifting                                                             0
##   theft-from-the-person                                                   0
##   vehicle-crime                                                           0
##   violent-crime                                                           0
##
##                        Investigation complete; no suspect identified
##   anti-social-behaviour                                                 677
##   bicycle-theft                                                         216
##   burglary                                                              154
##   criminal-damage-arson                                                 363
##   drugs                                                                  15
##   other-crime                                                            13
##   other-theft                                                           350
##   possession-of-weapons                                                   9
##   public-order                                                          193
##   robbery                                                                38
##   shoplifting                                                           299
##   theft-from-the-person                                                  61
##   vehicle-crime                                                         350
##   violent-crime                                                         595
##
##                        Local resolution Offender given a caution
##   anti-social-behaviour                0                        0
##   bicycle-theft                        1                        0
##   burglary                             0                        0
##   criminal-damage-arson               14                        4
##   drugs                               98                        6
##   other-crime                          0                        2
##   other-theft                          1                        2
##   possession-of-weapons               11                        6
##   public-order                         7                        1
##   robbery                              1                        0
##   shoplifting                         34                        5
##   theft-from-the-person                0                        0
##   vehicle-crime                        1                        0
##   violent-crime                       71                       35
##
##                        Status update unavailable
##   anti-social-behaviour                          0
##   bicycle-theft                                  8
##   burglary                                      13
##   criminal-damage-arson                          6
##   drugs                                          9
##   other-crime                                   10
##   other-theft                                   12
```

```
##   possession-of-weapons                        5
##   public-order                                14
##   robbery                                      5
##   shoplifting                                  4
##   theft-from-the-person                        2
##   vehicle-crime                                5
##   violent-crime                               84
##
##                         Suspect charged as part of another case
##   anti-social-behaviour                                      0
##   bicycle-theft                                              0
##   burglary                                                   0
##   criminal-damage-arson                                      0
##   drugs                                                      0
##   other-crime                                                0
##   other-theft                                                0
##   possession-of-weapons                                      0
##   public-order                                               0
##   robbery                                                    0
##   shoplifting                                                1
##   theft-from-the-person                                      0
##   vehicle-crime                                              0
##   violent-crime                                              0
##
##                         Unable to prosecute suspect Under investigation
##   anti-social-behaviour                           0                   0
##   bicycle-theft                                   8                   1
##   burglary                                       20                  20
##   criminal-damage-arson                         117                  17
##   drugs                                          13                  19
##   other-crime                                    34                   9
##   other-theft                                    94                  21
##   possession-of-weapons                          12                  10
##   public-order                                  218                  44
##   robbery                                        35                   7
##   shoplifting                                    76                  22
##   theft-from-the-person                          10                   3
##   vehicle-crime                                  23                  19
##   violent-crime                                1299                 247
```

The two-way table provides a breakdown of crime categories (rows) and their corresponding outcome statuses (columns). Each cell in the table represents the count of crime incidents belonging to a specific category and having a particular outcome status.

C. i) BARPLOT

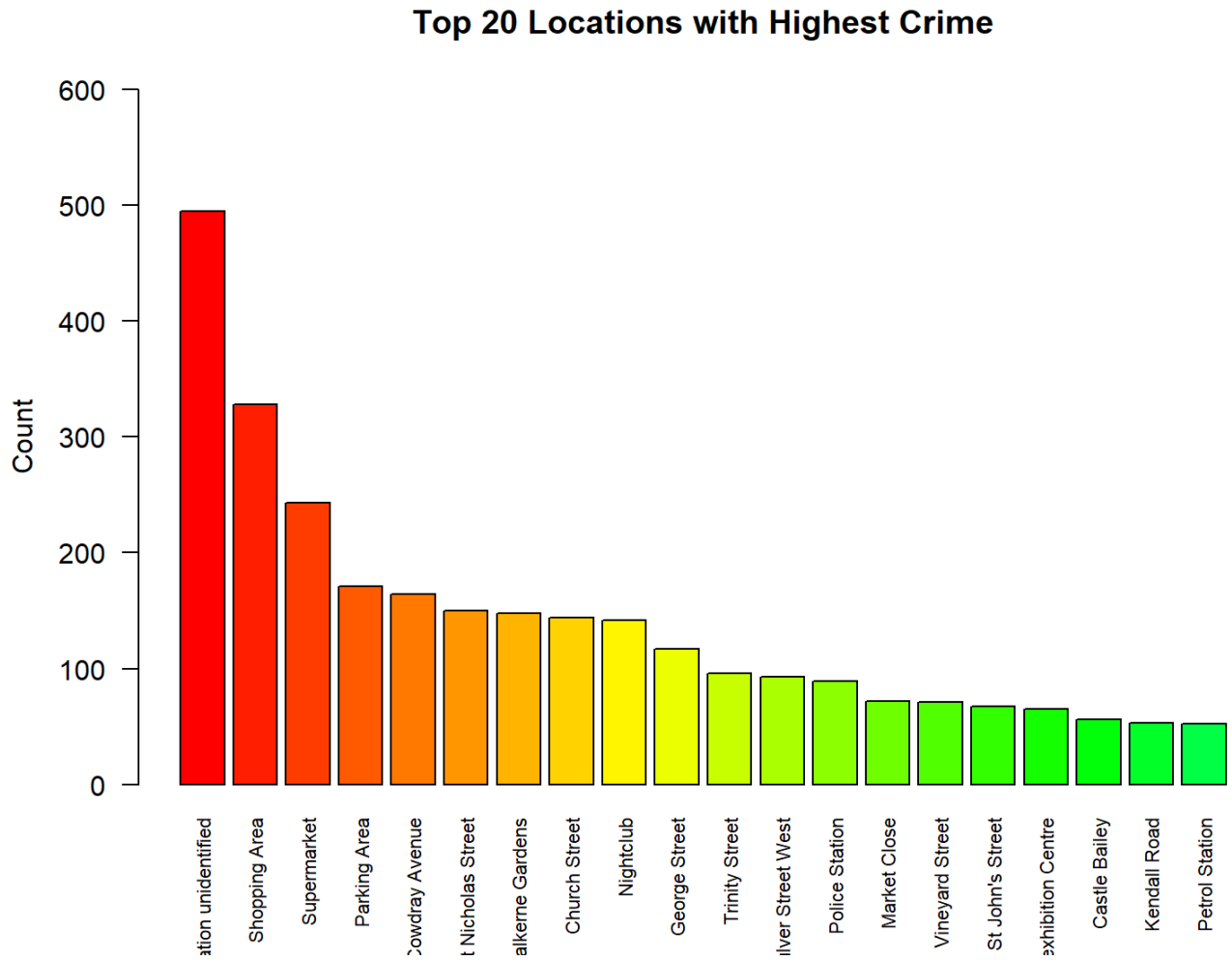
```
barplot(table(crimeData$category),names=c("anti-social-behaviour", "bicycle-theft", "burglary",
                                          "criminal-damage-arson", "drugs", "other-crime",
                                          "other-theft", "possession-of-weapons", "public-order",
                                          "robbery", "shoplifting", "theft-from-the-person",
                                          "vehicle-crime", "violent-crime"),
        main = "Counts of Occured Crime Types",
        xlab = "      Crime     Type",
        cex.names = 0.5,
        las=2,
        ylab = "Count",
        col = rainbow(30))
```



The above barplot shows the count of each crime type occured. As you can see violent crime has occured more number of time than rest of the crime types. I have used rainbow color so that each bar is of different color. This is making our barplot more aesthetic and attractive.

ii)

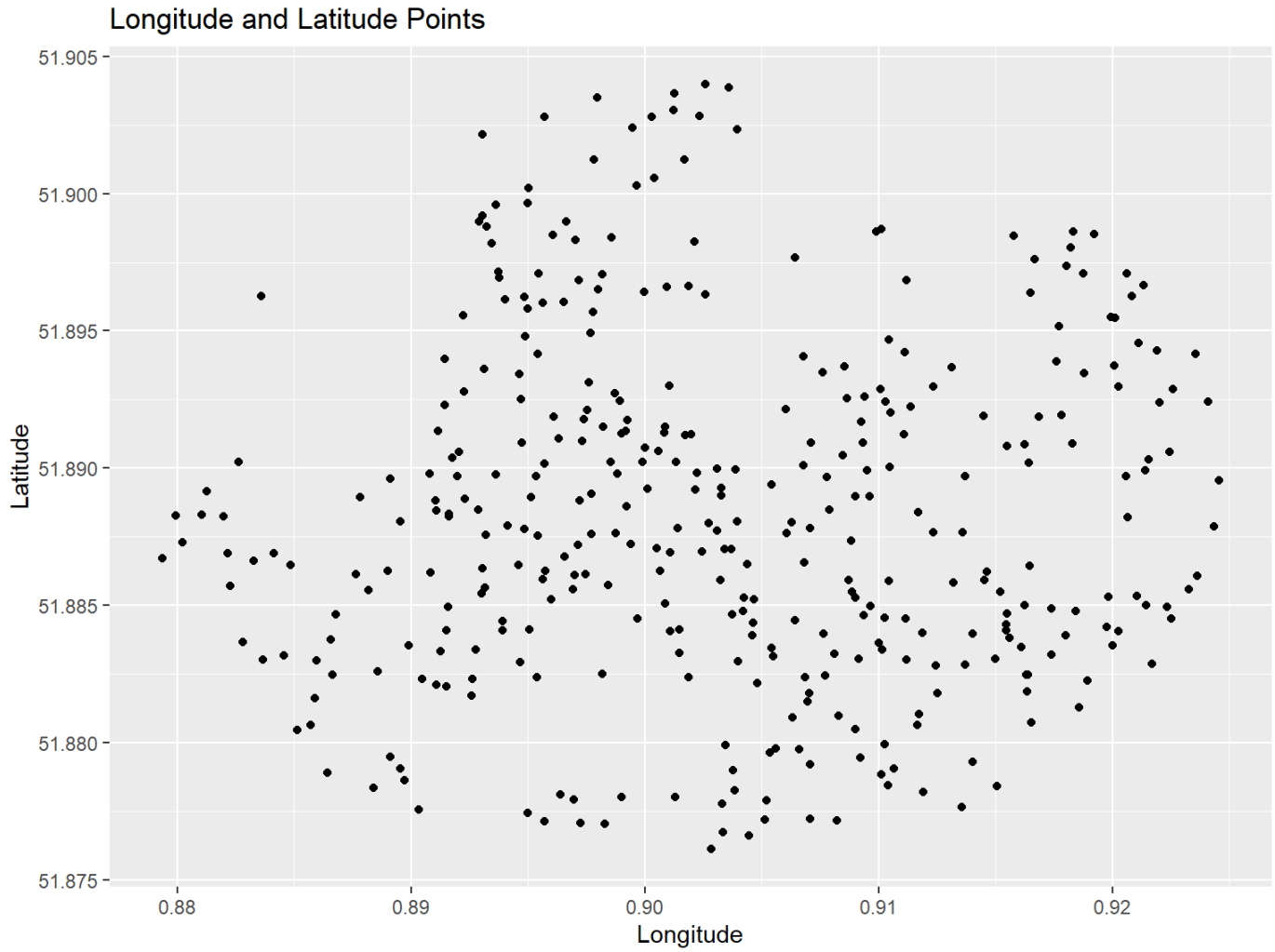
```
street_counts <- table(crimeData$street_name)

street_crime <- data.frame(location=names(street_counts),counts=as.numeric(street_counts))

# Sort the data by count in descending order
street_crime <- street_crime[order(-street_crime$counts), ]
street_crime$location <- gsub("^On or near ", "", street_crime$location)

# Select top N locations to visualize
N <- 20
top_locations <- head(street_crime, N)


barplot(top_locations$count,
        names.arg = top_locations$location,
        main = "Top 20 Locations with Highest Crime",
        ylab = "Count",
        ylim = c(0,600),
        las = 2,  # Rotate x-axis labels vertically for better readability
        col = rainbow(50),
        cex.names = 0.7)
```



This is another bar plot showing the top 20 locations where the crime has occurred in Colchester. There is a frequency on y-axis which represents the count of crime occurred. Shopping area is the identified location where the crime occured for highest number of times.

D. SCATTER PLOT

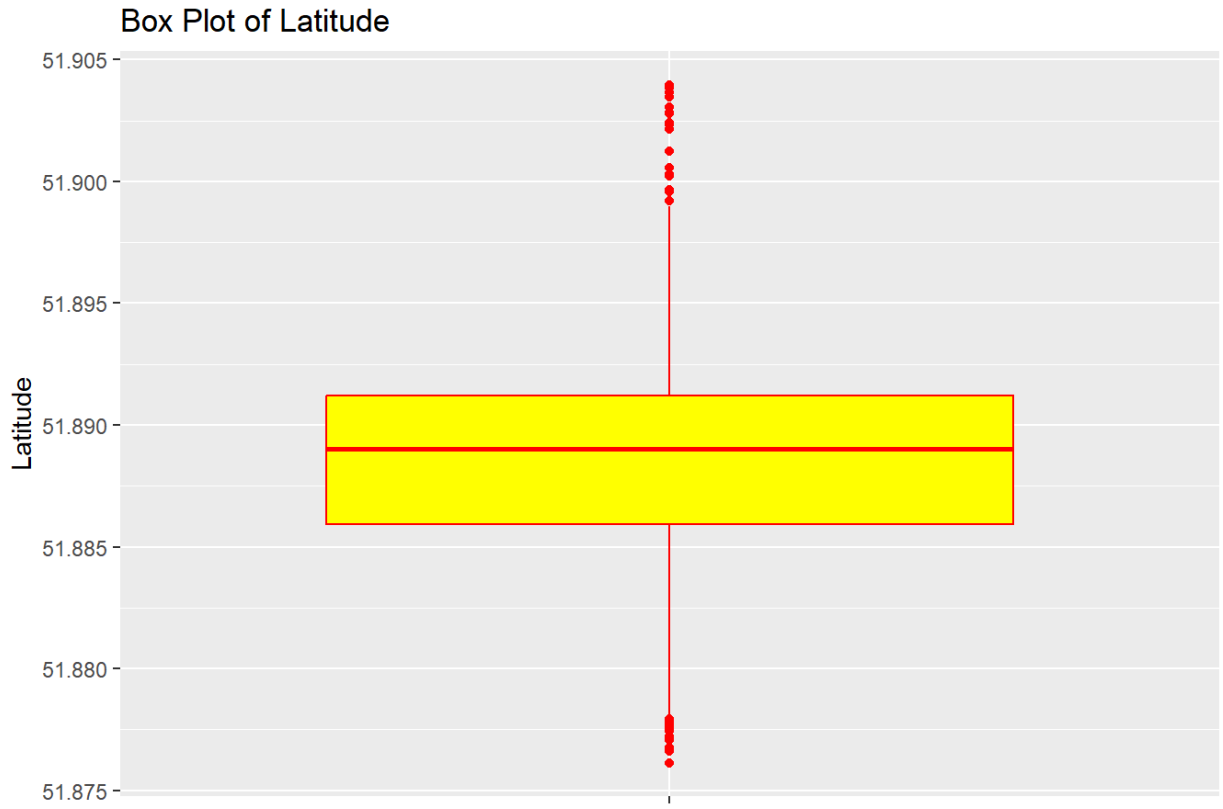
```
# Plotting longitude and latitude points on a map
ggplot(crimeData, aes(x = long, y = lat)) +
  geom_point() +
  labs(x = "Longitude", y = "Latitude") +
  ggtitle("Longitude and Latitude Points")
```



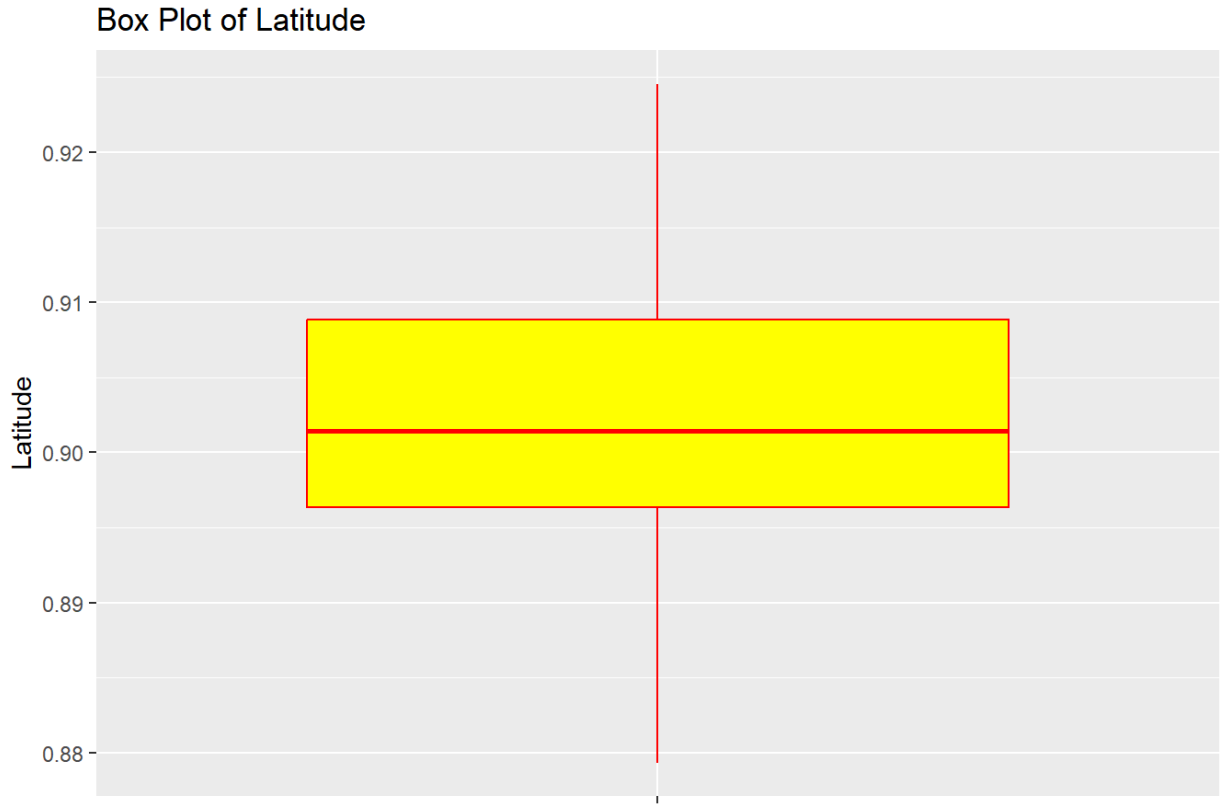
The geographic distribution of criminal occurrences recorded in Colchester in 2023 is shown visually by the scatter plot. With the x-axis representing the longitude and the y-axis representing the latitude coordinates, each point on the plot symbolizes a particular place where a crime was committed. Plot points arranged in clusters or patterns may represent spikes or places in the city with a greater rate of criminal activity. Law enforcement organizations and politicians might find regions that need more monitoring or focused action by analyzing such geographical patterns.

E. Box Plot

```
#Box plot for latitude
ggplot(crimeData, aes(x = "", y = lat)) +
  geom_boxplot(fill = "yellow", color = "red") +
  labs(title = "Box Plot of Latitude",
       x = "",
       y = "Latitude")
```



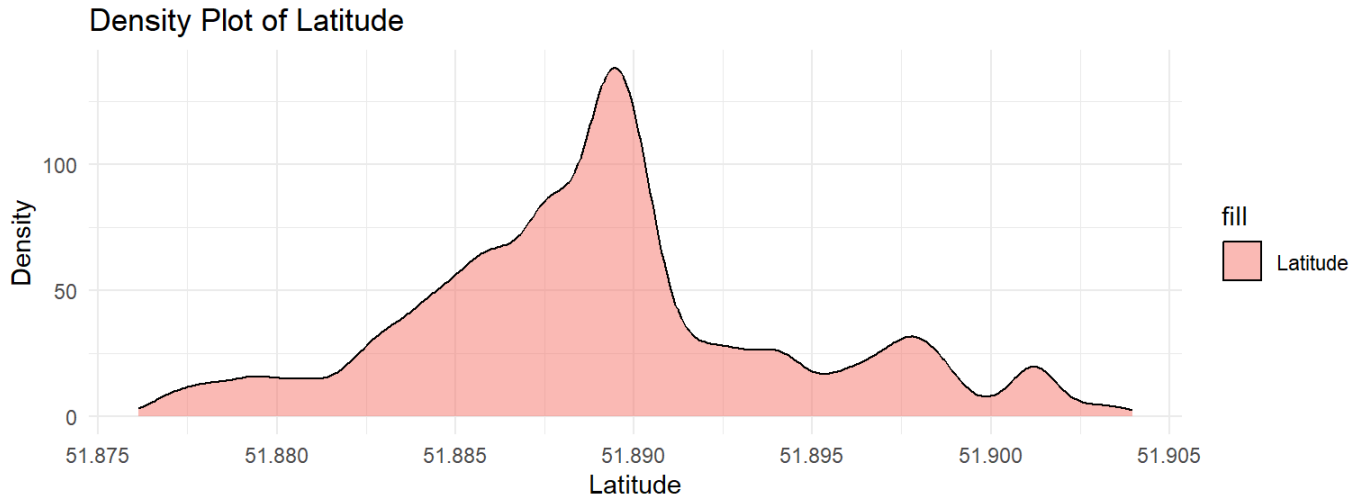

```
#Box plot for Longitude
ggplot(crimeData, aes(x = "", y =long)) +
  geom_boxplot(fill = "yellow", color = "red") +
  labs(title = "Box Plot of Latitude",
       x = "",
       y = "Latitude")
```

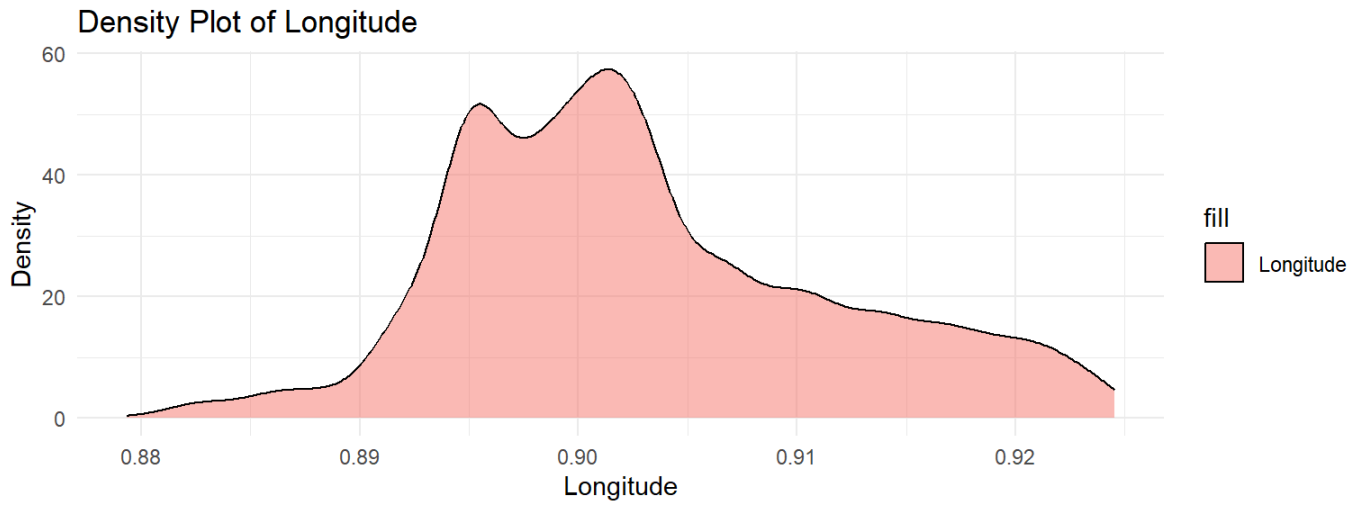


The above box plots illustrate the distribution of latitude and longitude coordinates associated with crime incidents reported in Colchester in 2023. Each box plot provides insights into the central tendency, spread, and variability of the geographical coordinates recorded for the reported crimes.

F. DENSITY PLOT

```
# Create density plots for Latitude and Longitude with different colors
density_plot_lat <- ggplot(crimeData, aes(x = lat, fill = "Latitude")) +
  geom_density(alpha = 0.5) +
  labs(title = "Density Plot of Latitude",
       x = "Latitude",
       y = "Density") +
  theme_minimal()

density_plot_long <- ggplot(crimeData, aes(x = long, fill = "Longitude")) +
  geom_density(alpha = 0.5) +
  labs(title = "Density Plot of Longitude",
       x = "Longitude",
       y = "Density") +
  theme_minimal()

# Combine the density plots using facets
combined_density_plot <- cowplot::plot_grid(density_plot_lat, density_plot_long, nrow = 2)

# Display the combined density plot
print(combined_density_plot)
```





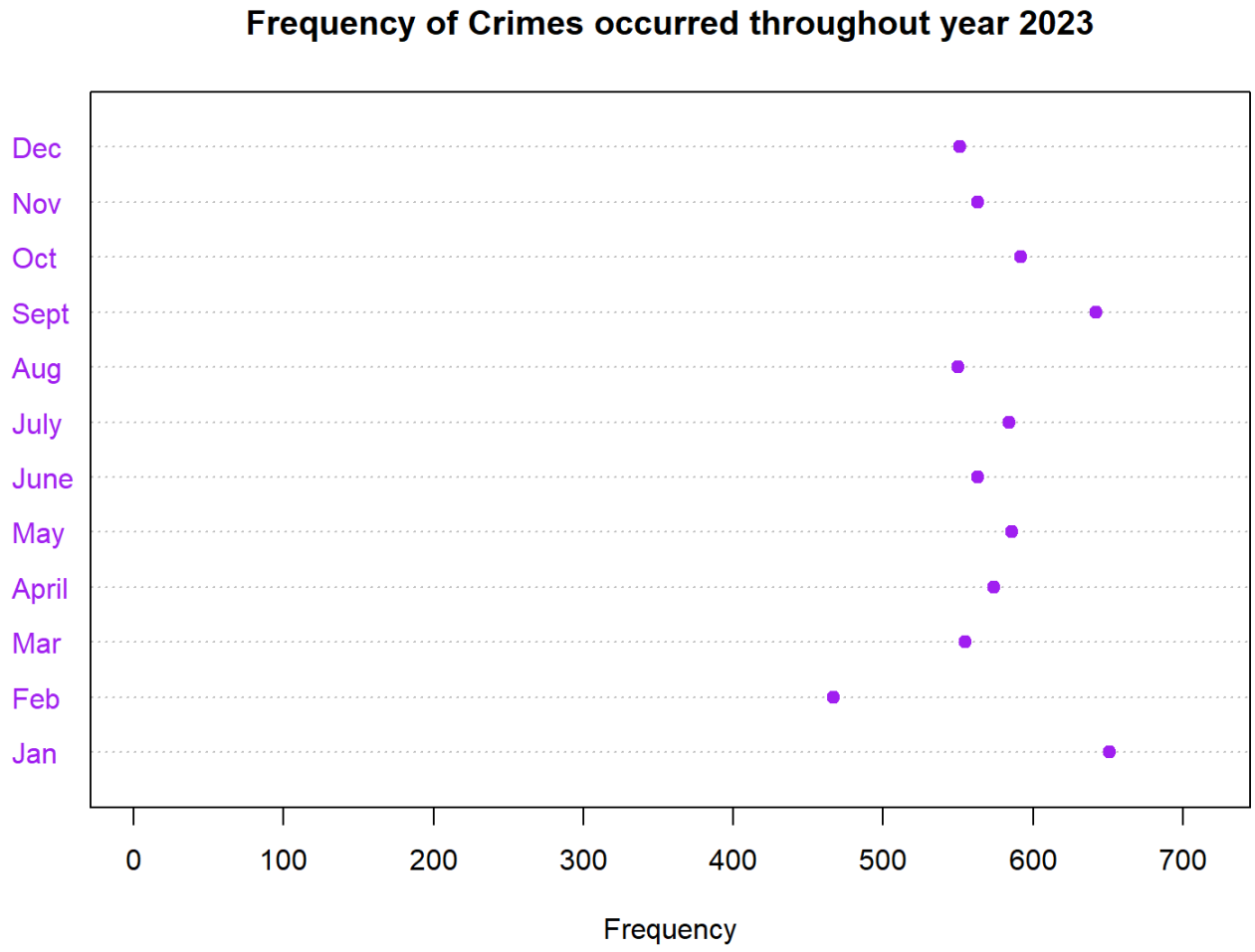
The above density plots are used to display the distribution of latitude and longitude coordinates linked to Colchester crime reports for the year 2023. The probability density function (PDF) of the geographic coordinates is revealed by each density plot, which also shows the variety and concentration of criminal episodes over various spatial regions.

The combined density plot allows for the simultaneous display and comparison of geographical distributions by integrating the latitude and longitude density charts using facets.

Density plot analysis can provide important information about the spatial features of crime episodes, such as the identification of hotspots, spatial trends, and regional differences in crime density.

G. DOT PLOT

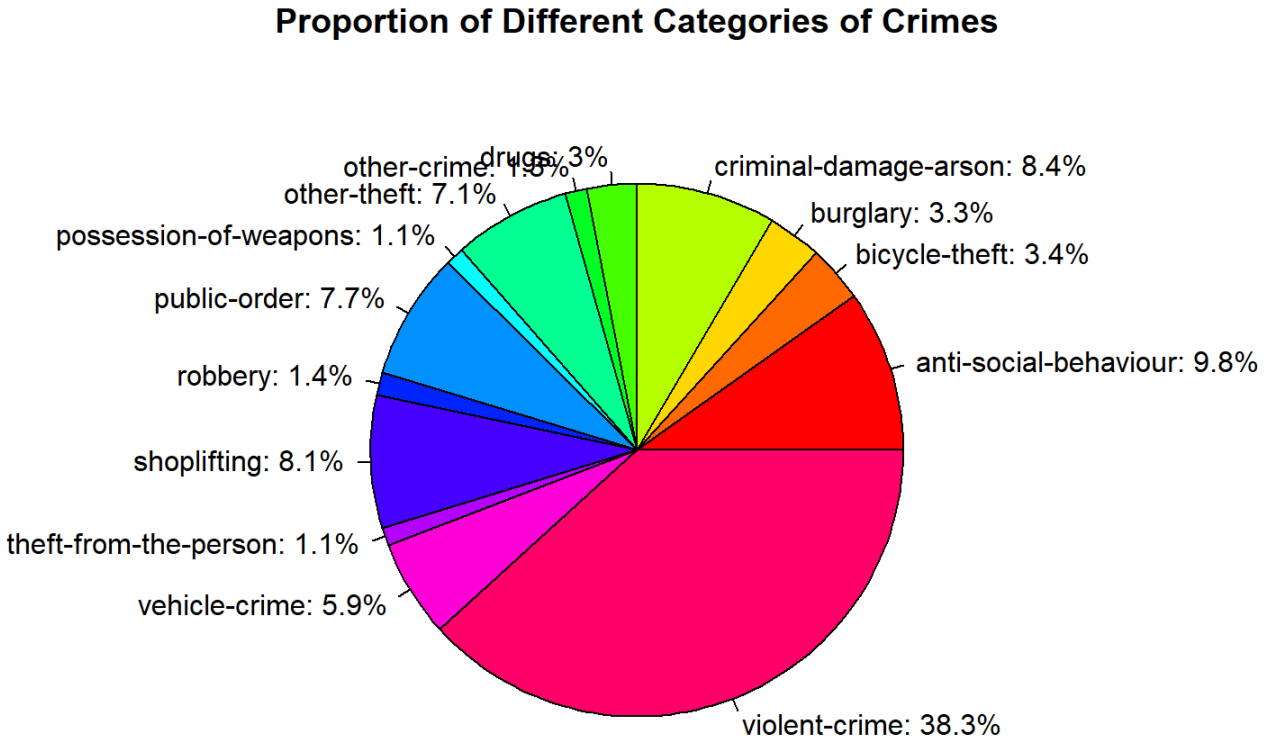
```
# Get frequency of crimes per month
crime_freq <- table(crimeData$date)

# Create dot plot
dotchart(crime_freq,
         labels = c("Jan", "Feb", "Mar", "April", "May", "June",
                    "July", "Aug", "Sept", "Oct", "Nov", "Dec"),
         main = "Frequency of Crimes occurred throughout year 2023",
         xlab = "Frequency",
         cex.names = 0.9,
         xlim = c(0, max(crime_freq) * 1.1),
         pch = 19,  # Use filled circles as points
         col = "purple",
         lcolor = "grey",
         lwd = 2)
```



The frequency of recorded crime events in Colchester every month for the year 2023 is depicted in the dot plot above. It is possible to quickly and intuitively comprehend seasonal trends and fluctuations in criminal activity by looking at the dots, each of which indicates the number of crimes reported in a certain month. The distribution of criminal occurrences from January to December of each year is shown in a dot plot. Every dot on the map represents the monthly frequency of recorded crimes; bigger dots signify months with more occurrences, while smaller dots indicate months with fewer reported crimes.

H. PIE CHART

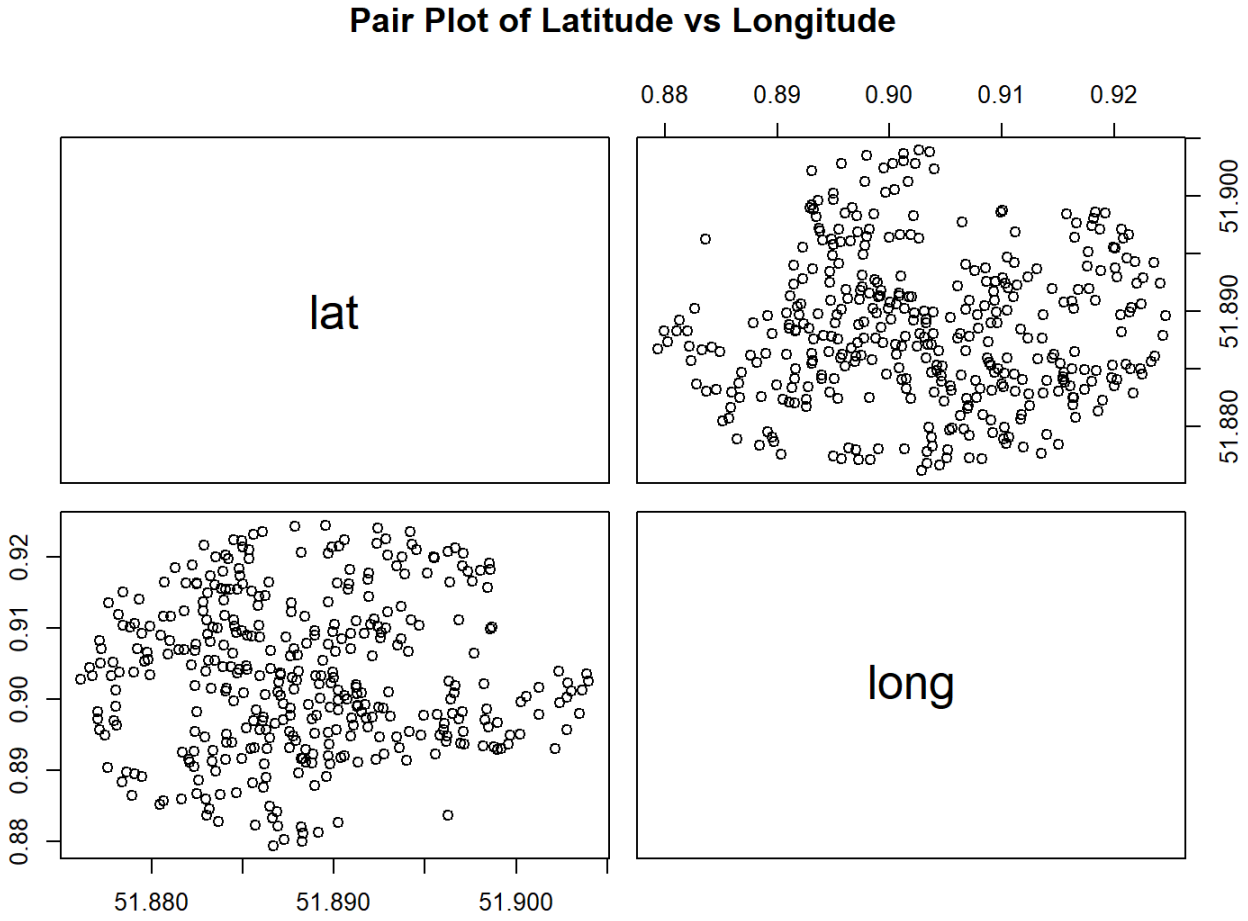
```
crime_counts <- table(crimeData$category)

# Plot pie chart
pie(crime_counts,
    main = "Proportion of Different Categories of Crimes",
    col = rainbow(length(crime_counts)),  # Color palette for the pie slices
    labels = paste(names(crime_counts), ": ", round(100 * crime_counts / sum(crime_counts), 1), "%", s
ep = ""))  # Add percentage labels
```



The proportional distribution of the various kinds of recorded crime events is depicted in the pie chart above. A particular crime category is represented by each segment of the pie, and the size of each segment indicates the percentage of all recorded offenses that fall into that category. The pie chart visually depicts the relative prevalence of various types of criminal activities within the dataset. Each segment of the pie corresponds to a specific crime category, such as anti-social behavior, burglary, theft, drugs, vandalism, and others. Segments with larger proportions indicate crime categories that are more prevalent or frequently reported, while smaller segments represent less common or infrequent crime types.

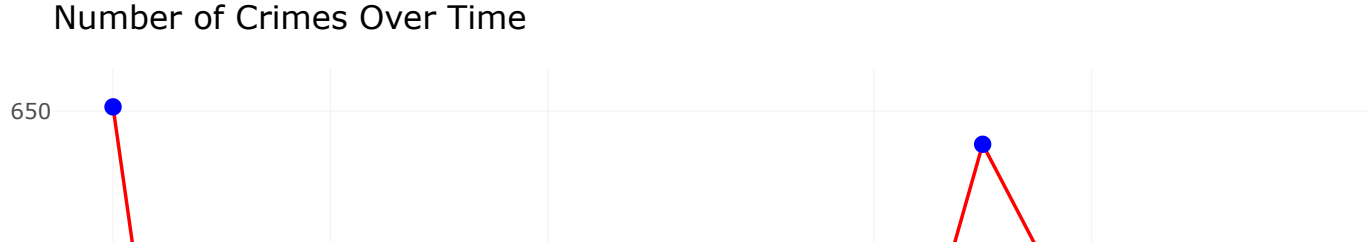
I. PAIR PLOT

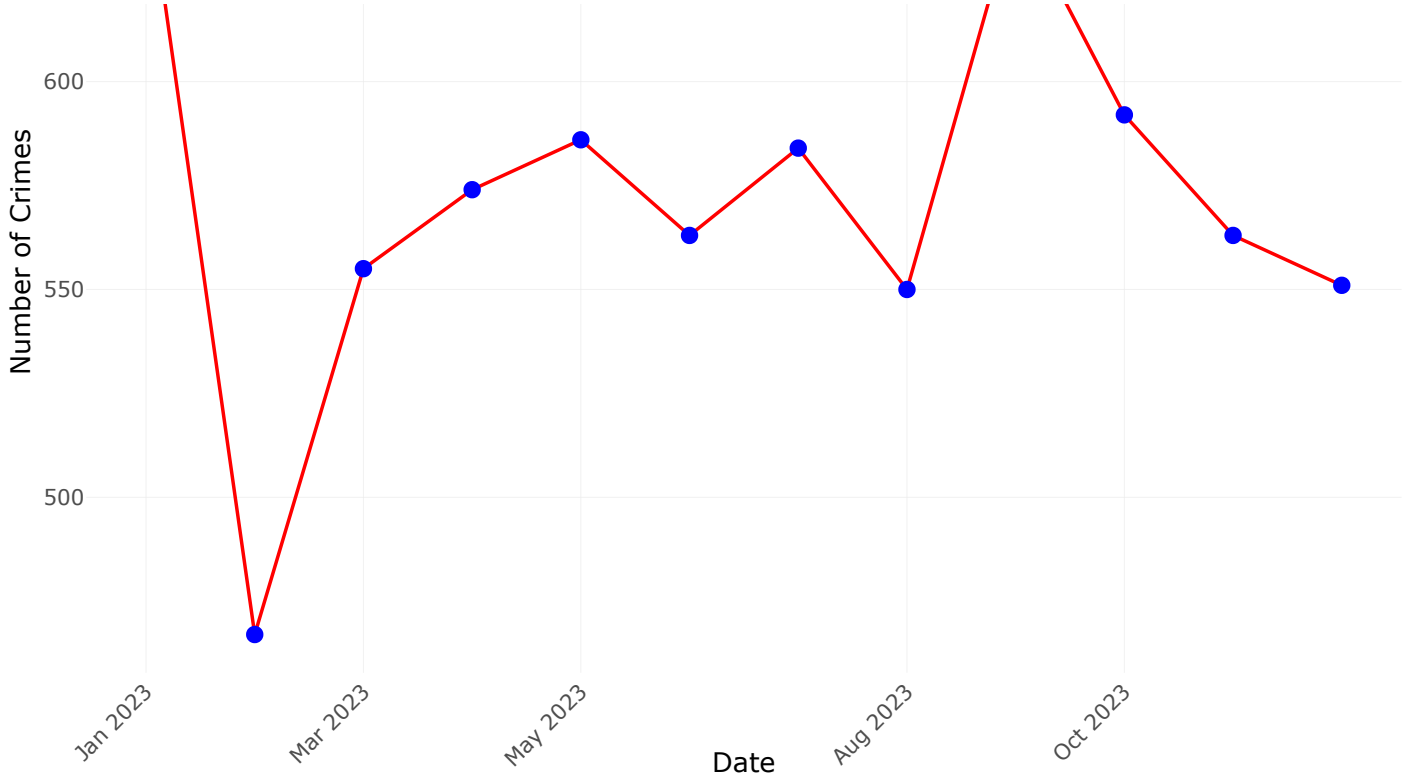
```
# Pair plot
pairs(~ lat + long, data = crimeData, main = "Pair Plot of Latitude vs Longitude")
```



J. i) TIME SERIES PLOT

```
# Time series plot
# Grouping number of crimes by date
crime_count <- crimeData %>%
  group_by(date) %>%
  summarise(crime_count = n())



# Creating the time series plot
gplot<-ggplot(crime_count, aes(x = date, y = crime_count, group=2)) +
  geom_line(color = "red") +
  geom_point(color = "blue", size = 2) +  # Adding points for each data point
  labs(title = "Number of Crimes Over Time",
       x = "Date",
       y = "Number of Crimes") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

ggplotly(gplot)
```





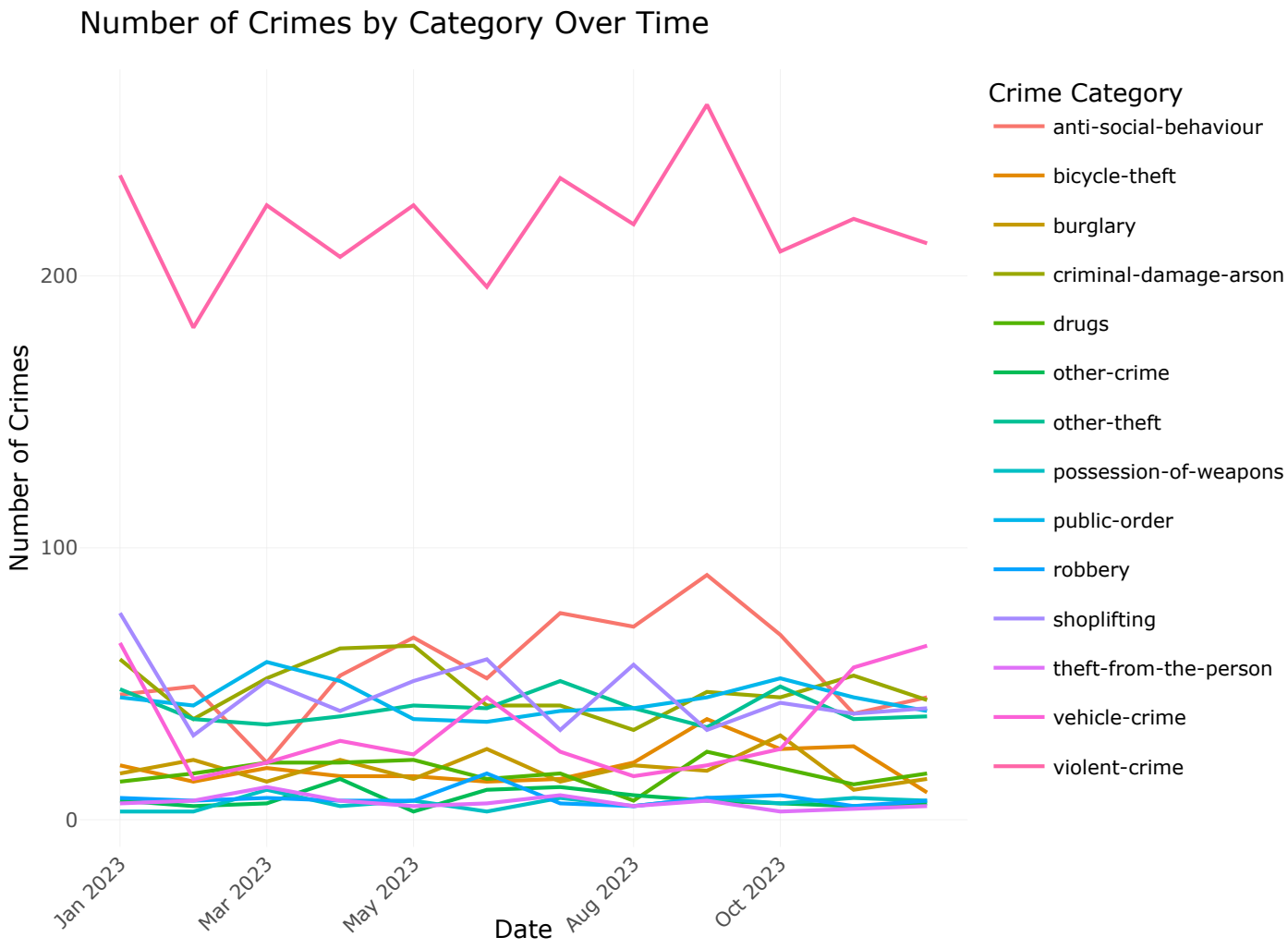
The periodic pattern of recorded crime events in Colchester during the course of 2023 is represented by the time series graphic. It displays the variation in crime rates across time, offering insights into trends, patterns, and possible seasonality in crime. The plot depicts the variation in the frequency of reported crimes across different time intervals, typically organized by date or month. The x-axis represents the timeline, with dates spanning the year 2023. The y-axis indicates the count of reported crimes, showcasing the volume of incidents recorded within each time period. As you can see January month has the highest number of crimes occurred followed by September month. The time series' highs and lows show times of increased or decreased criminal activity. The plot is interactive, allowing users to hover over data points to view specific information about crime counts at particular dates.

ii)

```
# Aggregating number of crimes by date and category
crime_count_category <- crimeData %>%
  group_by(date, category) %>%
  summarise(crime_count = n())
```

```
## `summarise()` has grouped output by 'date'. You can override using the
## `.groups` argument.
```

```
# Creating the time series plot
gplot1<-ggplot(crime_count_category, aes(x = date, y = crime_count, color = category)) +
  geom_line(size=0.60) +
  labs(title = "Number of Crimes by Category Over Time",
       x = "Date",
       y = "Number of Crimes",
       color = "Crime Category") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

ggplotly(gplot1)
```

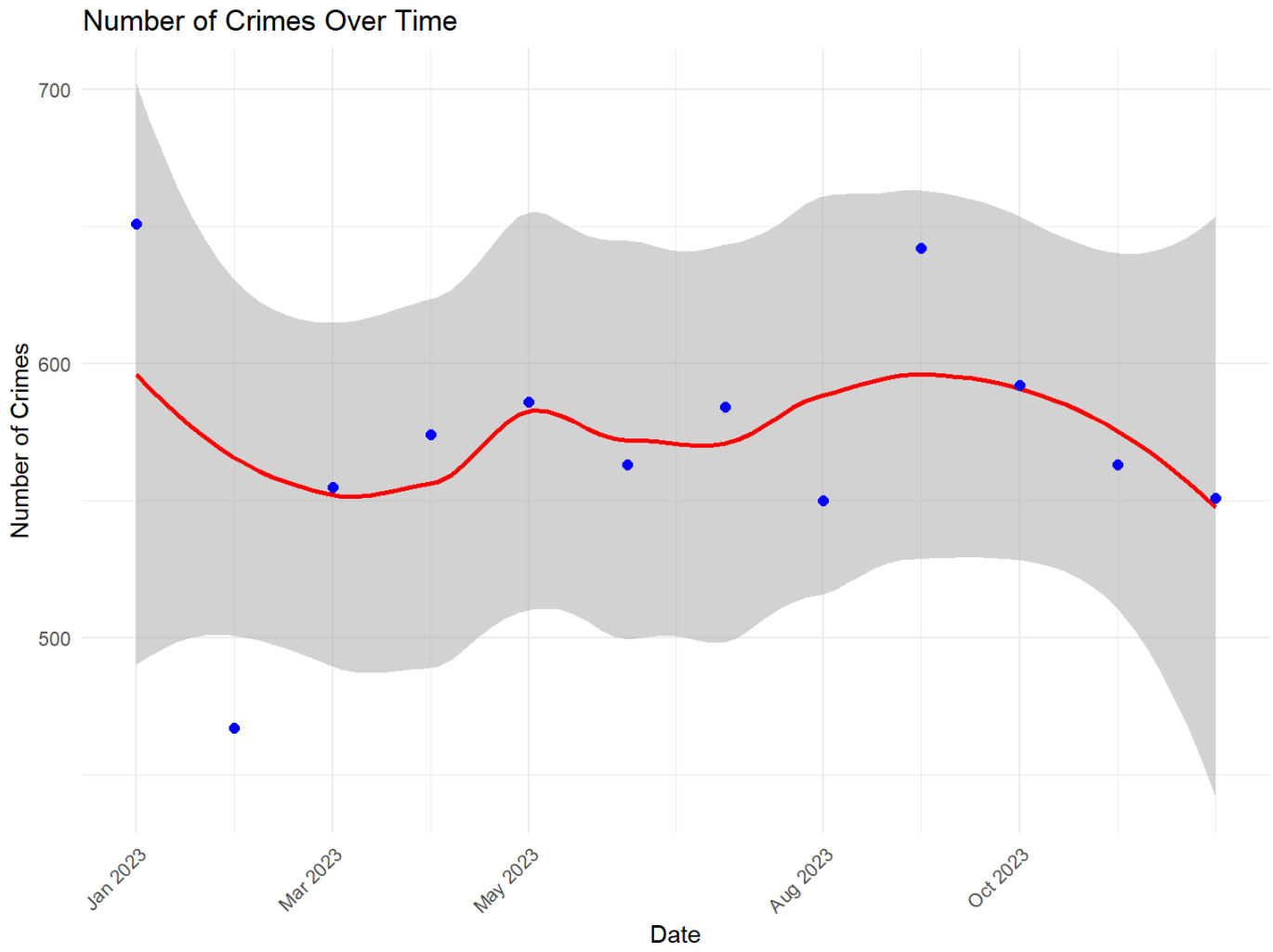


The time series plot depicts the temporal variance in recorded crime events in Colchester, classified by crime category, during the course of 2023. It gives a complete picture of how the frequency of different crime categories evolves over time. Each plot line depicts a different type of crime, such as criminal behavior, burglary, or theft. The x-axis represents the timeline until 2023, while the y-axis represents the number of recorded offenses for each category. Different colors are utilized to distinguish across crime categories, allowing for easier visual separation and comparison. The graph is interactive, allowing viewers to click over data points to get detailed information on crime numbers for each category on certain days.

K. Smoothning

```
ggplot(crime_count, aes(x = date, y = crime_count, group=2)) +
  geom_smooth(method="loess",color = "red") +
  geom_point(color = "blue", size = 2) +  # Add points for each data point
  labs(title = "Number of Crimes Over Time",
       x = "Date",
       y = "Number of Crimes") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



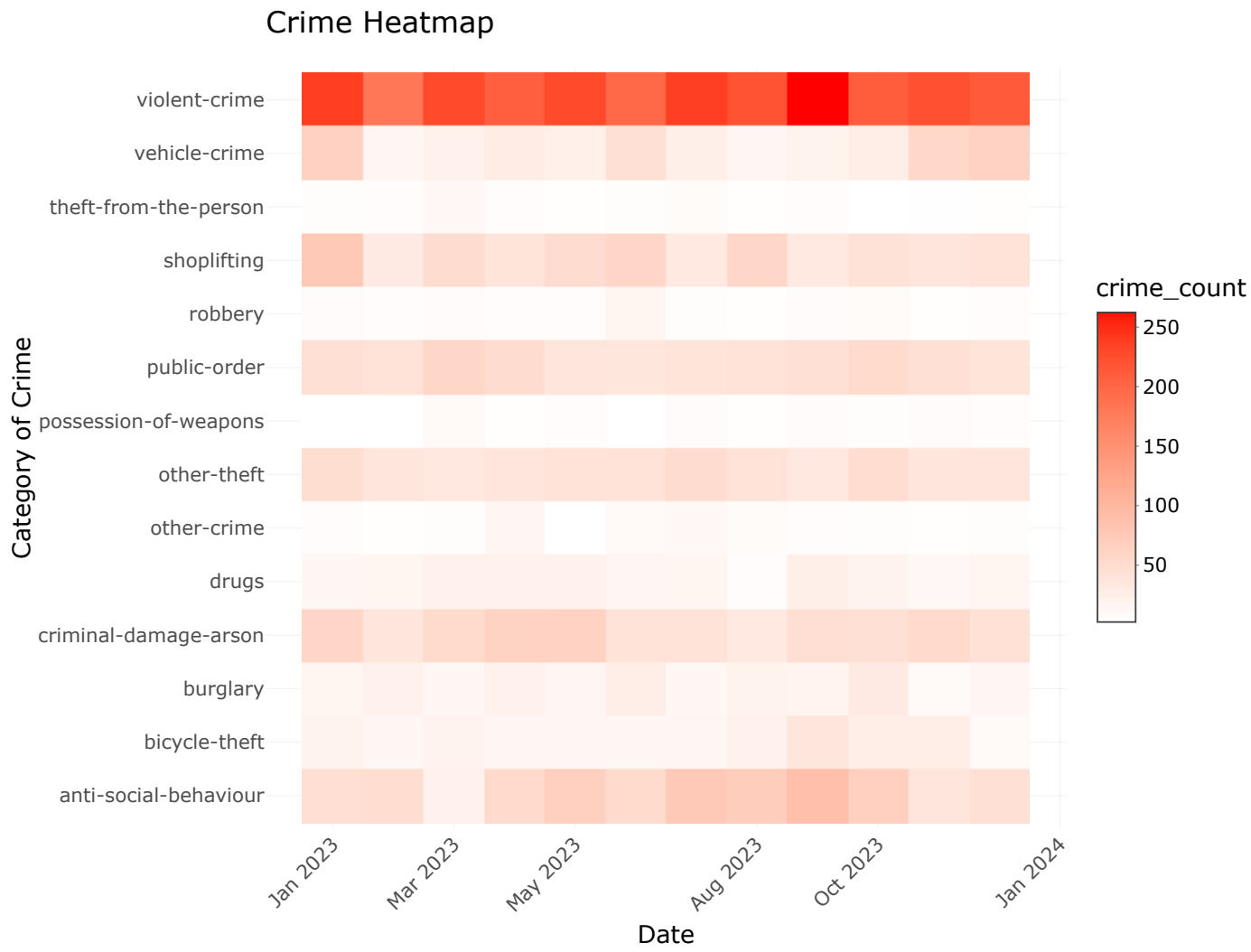
The above time series plot displays the trend in the number of recorded crimes in Colchester over the course of the year 2023, using an LOESS (Locally Estimated Scatterplot Smoothing) curve to highlight the underlying pattern within the data's inherent noise.

L. Heat Map

```
# Aggregate the data to get counts of crimes by category and date
crime_count_heatmap <- crimeData %>%
  group_by(category, date) %>%
  summarise(crime_count = n()) %>%
  ungroup()
```

```
## `summarise()` has grouped output by 'category'. You can override using the
## `.groups` argument.
```

```
# Create the heatmap
gplot2<-ggplot(crime_count_heatmap, aes(x = date, y = category, fill = crime_count)) +
  geom_tile() +
  scale_fill_gradient(low = "white", high = "red") +  # Choose a color gradient
  labs(title = "Crime Heatmap",
       x = "Date",
       y = "Category of Crime") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        legend.position = "right")  # Adjust legend position as needed

ggplotly(gplot2)
```



The crime heatmap depicts the frequency with which various types of crimes are recorded over time, offering a thorough picture of crime distribution and temporal patterns in Colchester in 2023. The heatmap uses a color gradient to indicate the frequency of recorded crimes, with warmer colors signifying greater crime counts and colder hues denoting lower counts. Each cell in the heatmap represents a distinct combination of crime category and date, providing a thorough representation of crime distribution across categories and time periods.

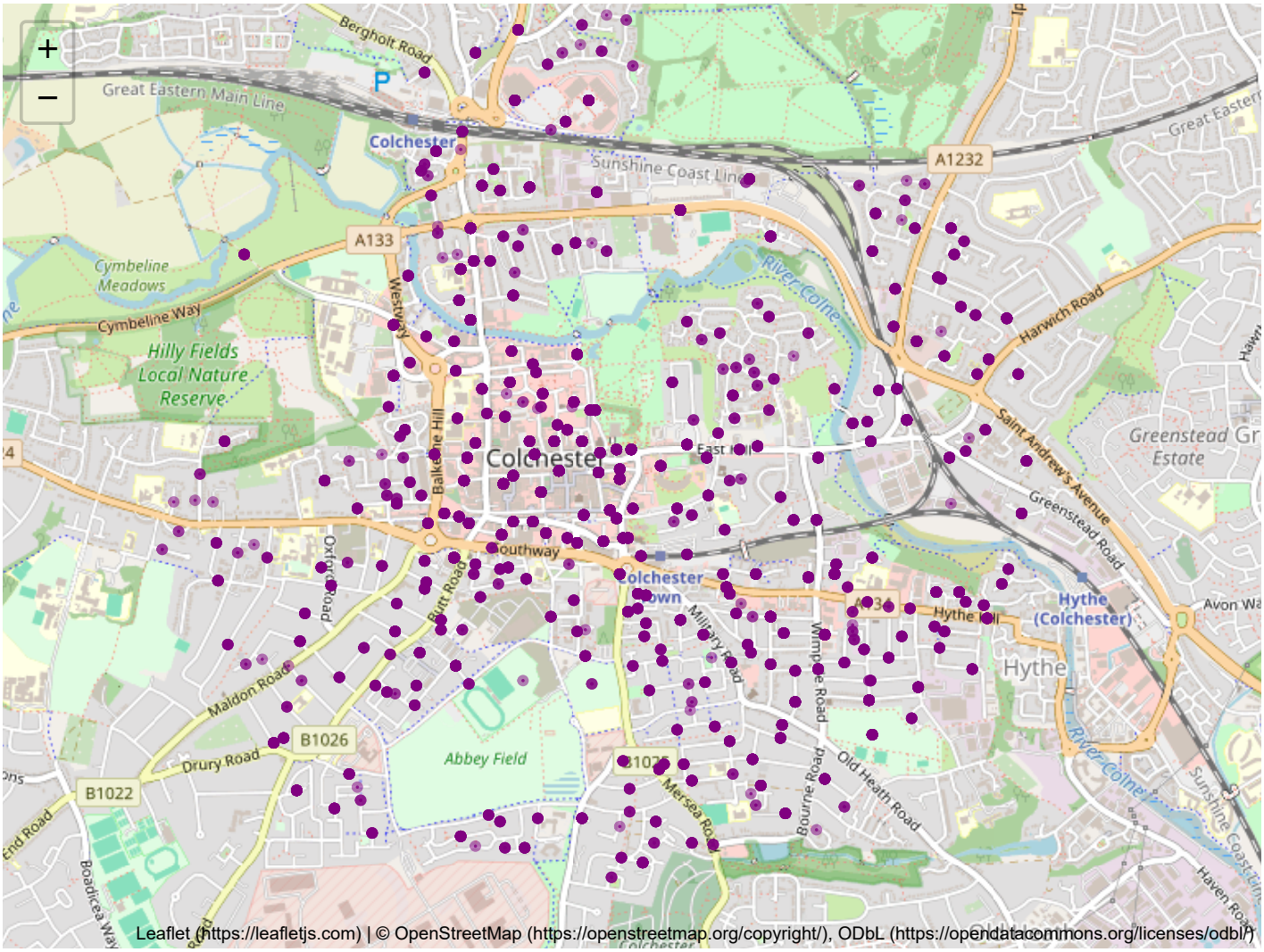
M. Map

i)

```
# Creating a new data frame having only neccessary columns.
location_map <- data.frame(
  inc_type = crimeData$category,
  street = crimeData$street_name,
  street_id = crimeData$street_id,
  latitude = as.numeric(crimeData$lat),
  longitude = as.numeric(crimeData$long)
)

# Creating a leaflet map with circle markers for each incident location in Colchester city.
map_colchester <- leaflet(location_map) %>%
  addTiles() %>%
  addCircleMarkers(radius = 0.5, color = "purple", fillOpacity = 0.8,
                   popup = paste0("Street: ", location_map$street,
                                  "<br>Incident Type: ", location_map$inc_type)) %>%
  setView(lng = 0.9040, lat = 51.8891, zoom = 13.5)
```

```
## Assuming "longitude" and "latitude" are longitude and latitude, respectively
```

```
map_colchester
```



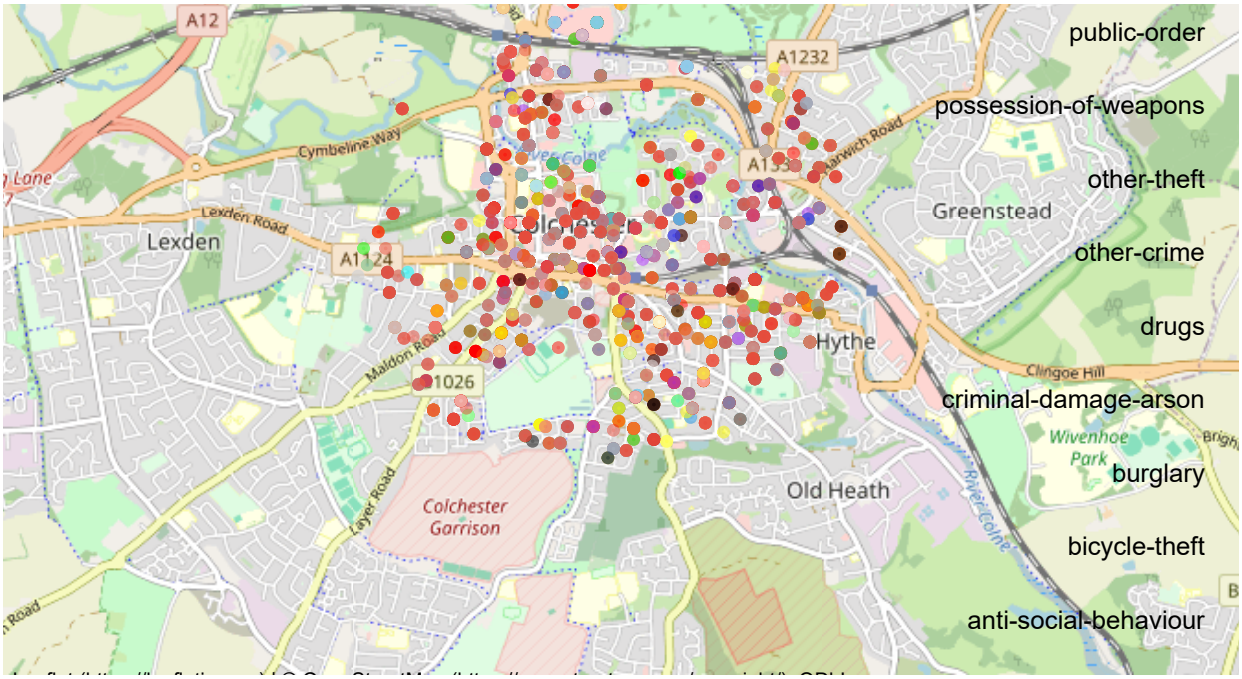
The above Leaflet map depicts the geographic distribution of crime events in Colchester, providing geographical insights on where certain types of crimes happened in 2023. The map has an interactive interface that allows users to zoom in/out and pan across it to explore different parts of Colchester. The color of the markers (purple) enhances visibility and distinguishes them from the map background.

ii)

```
colors <- c(
  "anti-social-behaviour" = "red",
  "bicycle-theft" = "blue",
  "burglary" = "green",
  "criminal-damage-arson" = "yellow",
  "drugs" = "purple",
  "other-crime" = "black",
  "other-theft" = "white",
  "possession-of-weapons" = "turquoise",
  "public-order" = "#5DADE2",
  "robbery" = "orange",
  "shoplifting" = "skyblue",
  "theft-from-the-person" = "cyan",
  "vehicle-crime" = "grey",
  "violent-crime" = "#E74C3C"
)

# Creating a factor variable for category with corresponding colors
location_map$inc_type <- factor(location_map$inc_type, levels = names(colors))
palette <- colorFactor(palette = colors, domain = location_map$inc_type)

# Creating a leaflet map with circle markers for each incident location in Colchester.
map_colchester <- leaflet(location_map) %>%
  addTiles() %>%
  addCircleMarkers(radius = 1,
                   color = ~palette(inc_type),  # Assign color based on category
                   popup = paste0("Street: ", location_map$street,
                                  "<br>Incident Type: ", location_map$inc_type)) %>%
  setView(lng = 0.9040, lat = 51.8891, zoom = 13)
```

```
## Assuming "longitude" and "latitude" are longitude and latitude, respectively
```

```
# Define legend labels
legend_labels <- names(colors)

# Creating legend
for (i in seq_along(legend_labels)) {
  map_colchester <- map_colchester %>%
    addLegend("bottomright", colors = colors[legend_labels[i]],
              labels = legend_labels[i], opacity = 0.5)
}

# Displaying the map
map_colchester
```

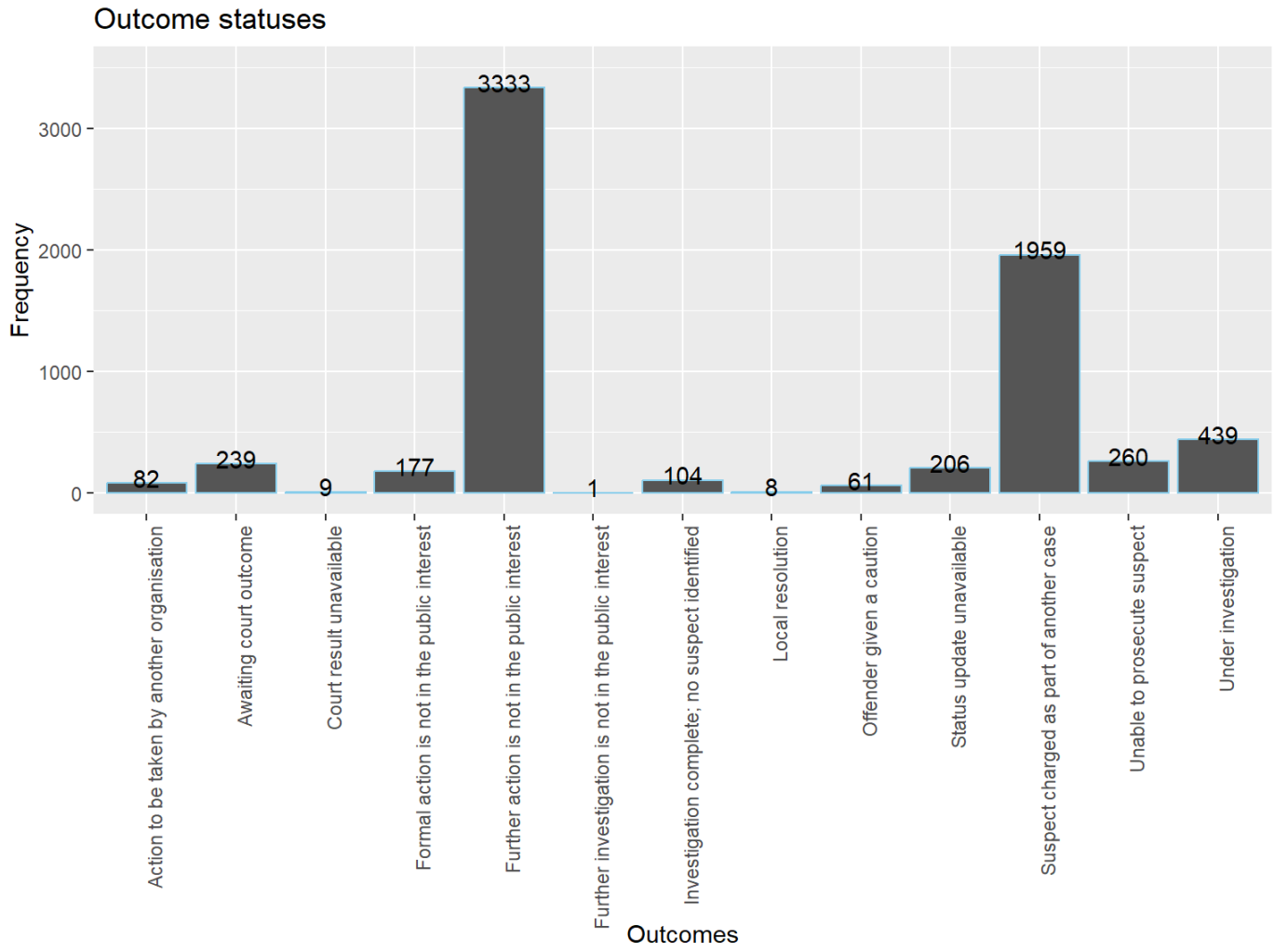


We're using a leaflet map to visualize criminal incidences in Colchester in this bit of code. A circular marker is used to indicate each occurrence, and the color of the marker identifies the sort of crime that occurred. Each criminal category has a predetermined set of colors, which ensures visual clarity and distinguishes between various occurrence kinds. I've also included a legend on the map to serve as a guide for understanding the colors.

This graphic provides a geographical overview of how various crime categories are distributed in Colchester. It facilitates the identification of trends, hotspots, and problem areas by stakeholders, which helps with well-informed decision-making for law enforcement and community safety initiatives. Furthermore, color coding improves communication's efficacy and clarity when delivering complicated details concerning criminal situations.

N. Other Plots

```
counts <- as.numeric(table(crimeData$outcome_status))
outcome <- unique(crimeData$outcome_status)
df1 <- data.frame(outcome=outcome,counts=counts)


ggplot(data = df1) +
  aes(x = outcome, y = counts) +
  geom_bar(stat = "identity", color = "skyblue") +
  labs(title = "Outcome statuses",x='Outcomes', y = "Frequency") +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  coord_cartesian(ylim = c(0, 3500)) +
  geom_text(aes(label = counts, hjust = "center", vjust = 0.2))
```



The bar plot depicts the frequency distribution of various outcome statuses resulting from recorded crime incidences. The x-axis represents various outcome statuses, indicating the actions taken or the status of each reported crime incident, such as “Awaiting court outcome,” “Court result unavailable,” “Formal action is not in the public interest,” and “Further action is not in the public interest.” The y-axis displays the frequency of each outcome status, representing the number of crime incidents associated with each outcome. Numeric labels are added above each bar, displaying the exact frequency count for better clarity and interpretation.

CONCLUSION

Throughout the course of the year 2023, the examination of the policing dataset from Colchester yields significant findings regarding the characteristics and trends of reported criminal occurrences. We gained a thorough understanding of key aspects such as the frequency and distribution of different crime categories, the spatial distribution of incidents across the city, temporal trends in crime rates, and the outcomes associated with reported incidents by employing a variety of data visualization techniques. The examination of crime frequency over time highlighted fluctuations in criminal activities across different months, allowing us to identify potential seasonal variations or trends. Furthermore, the visualization of crime categories provided a clear overview of the types of offenses most prevalent in Colchester, aiding in resource allocation and strategic planning for law enforcement efforts. Spatial analysis using maps and spatial visualizations revealed the location of crime hotspots around the city, allowing for targeted interventions and proactive enforcement techniques in high-risk regions. We can establish evidence-based plans for reducing crime, improving community involvement, and creating a safer environment for both residents and tourists by employing data-driven techniques and visualizations.