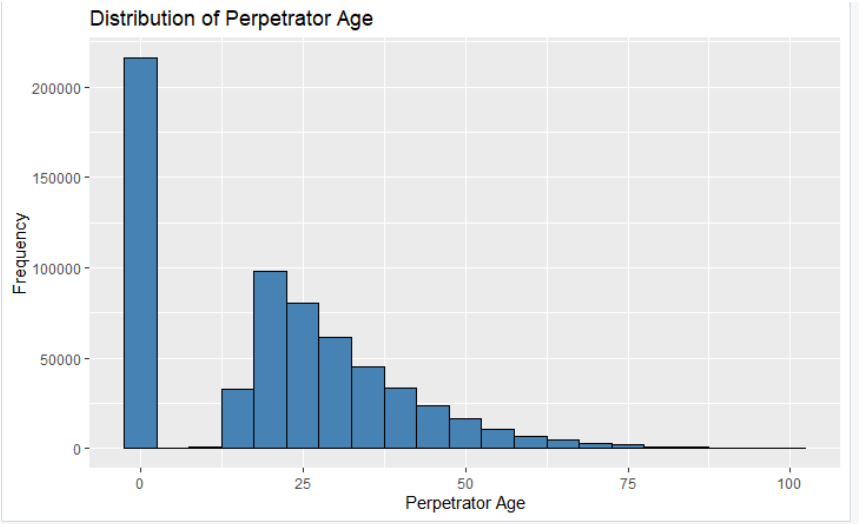
Name : Chinmay Jadhav

UID : 2021300046

Exp : 4

Dataset link : https://www.kaggle.com/datasets/mrayushagrawal/us-crime-dataset

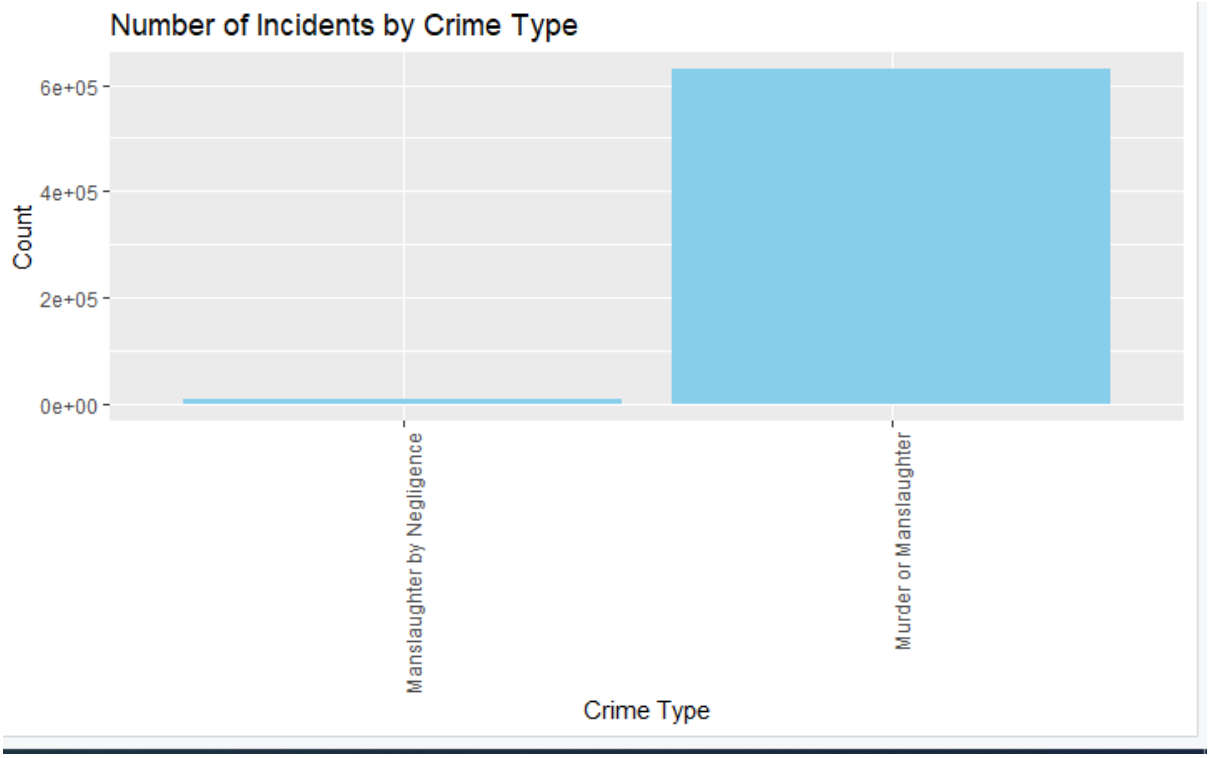


Purpose:

The histogram visualizes the distribution of perpetrator ages, showing how frequently different age groups are involved in criminal activities.

Analysis:

- The **X-axis** represents the age of perpetrators, ranging from 0 to 100 years.
- The **Y-axis** indicates the frequency, or the number of incidents, involving perpetrators in each age group.
- There is a sharp spike at age 0, indicating either missing or erroneous data for a large number of records.
- After excluding the age 0 spike, the highest frequency of incidents is observed for perpetrators between the ages of 15 to 30, with the frequency decreasing as age increases.
- The distribution shows a right-skewed pattern, with the majority of perpetrators concentrated in younger age groups, and fewer incidents involving older perpetrators.
- The data implies that younger individuals are more frequently involved in incidents, and the likelihood of involvement decreases with age.

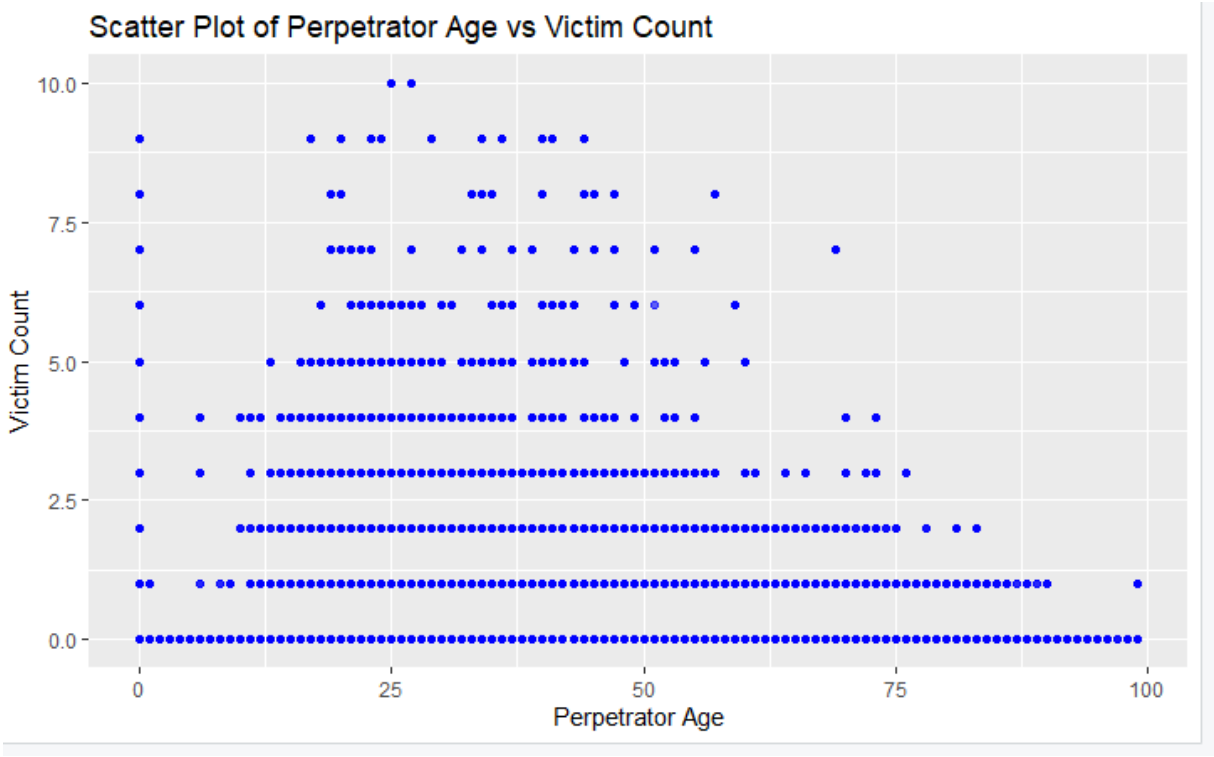


Purpose:

The bar plot visualizes the number of incidents for two types of violent crimes: "Manslaughter by Negligence" and "Murder or Manslaughter," highlighting the disparity between the two categories.

Analysis:

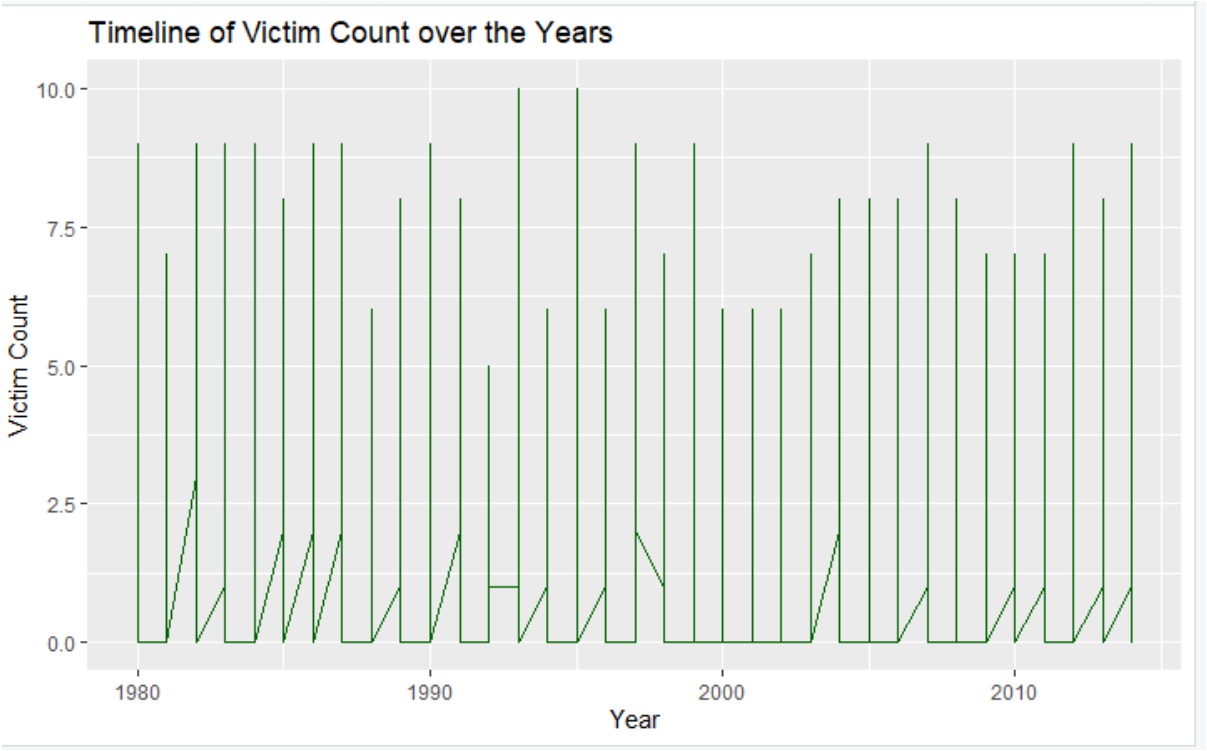
- The **X-axis** represents the two crime types: "Manslaughter by Negligence" and "Murder or Manslaughter."
- The **Y-axis** shows the count of incidents for each type, ranging from 0 to over 600,000.
- The number of incidents for "Murder or Manslaughter" is significantly higher, with close to 600,000 cases.
- "Manslaughter by Negligence" has a far lower count, close to zero, in comparison.
- This plot emphasizes the substantial difference in the occurrence of these two crime types, where "Murder or Manslaughter" is much more prevalent than "Manslaughter by Negligence."



- **Purpose**: The scatter plot visualizes the relationship between the age of perpetrators and the number of victims involved in the incidents.

- **Analysis**:

  - Each point on the graph represents an individual incident, where the x-axis corresponds to the perpetrator's age and the y-axis to the number of victims.
  - If points are widely dispersed across the plot, it indicates no clear correlation between perpetrator age and victim count.
  - Clusters of points in specific areas might suggest patterns, such as certain age groups being more likely to have a higher or lower victim count.
  - If a line pattern appears (positive or negative), it could indicate a trend where victim count increases or decreases with the perpetrator's age.
  - **Outliers**: Any points far removed from the rest of the data might indicate unusual cases, such as very young or old perpetrators with an exceptionally high number of victims.

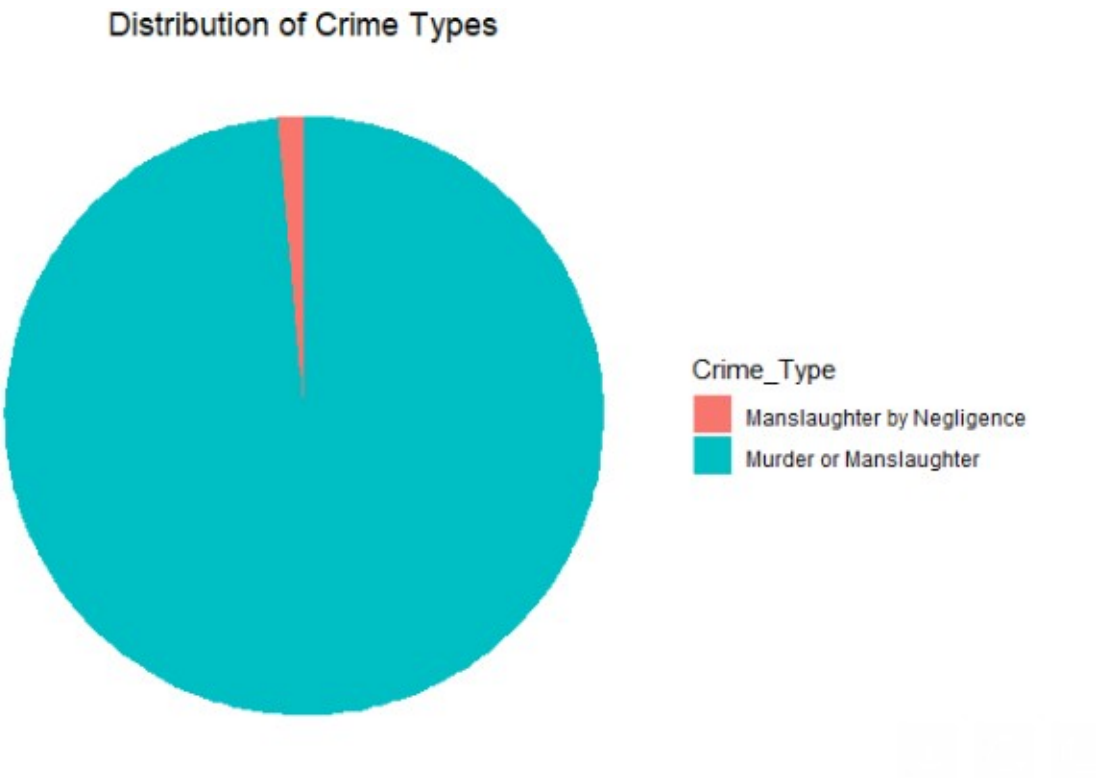


Purpose:

The line plot illustrates the timeline of victim count per incident over several years, from around 1980 to 2015, to observe changes and patterns in the number of victims over time.

Analysis:

- The **X-axis** represents the year, ranging from approximately 1975 to 2015.
- The **Y-axis** shows the number of victims per incident, ranging from 0 to 10.
- The graph displays peaks and valleys, with many incidents reaching a victim count of 10 in some years.
- There are recurring fluctuations over the timeline, with periods of higher victim counts followed by drops to lower values or zero.
- The data suggests no clear increasing or decreasing trend, but rather oscillations in victim count over time. This could imply that the number of victims per incident varies year by year without a consistent upward or downward pattern.



The above Pie chart is used to show the distribution between the 2 types of murders.

Conclusion : Hence, we were able to plot basic graphs using R for the given Criminal/Law dataset in R studios.