Experiment no 4

Aim :
Create basic charts using R programming language on dataset Crime or Police / Law and Order

- Basic - Bar chart, Pie chart, Histogram, Time line chart, Scatter plot, Bubble plot
- Write observations from each chart

Example

https://app.powerbigov.us/view?r=eyJrIjoiYmU4MDhiYWItYjEwOS00ZDg5LTk1OTUtNzNlMmU0MDFjNTk5IiwidCI6IjI5MzU3MDllLWMxMGMtNDgwOS1hMzAyLTg1MmQzNjlmODcwMCJ9

Objectives:

- To understand and apply basic data visualization techniques in R.
- To create various types of charts (Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot) using a crime-related dataset.
- To interpret and analyze the data through visual representations.

Theory:

Data visualization is an essential skill in data analysis that helps in understanding trends, patterns, and relationships within a dataset. R, a powerful statistical programming language, provides a wide range of tools for creating visually appealing and informative charts. In this experiment, we will use basic chart types to analyze crime data and derive insights.

Chart Types:

1. **Bar Chart:** A bar chart is used to display categorical data with rectangular bars representing the frequency or count of each category.
2. **Pie Chart:** A pie chart shows the proportion of categories as slices of a pie, useful for comparing parts of a whole.
3. **Histogram:** A histogram is used to represent the distribution of numerical data by grouping it into bins.
4. **Timeline Chart:** A timeline chart visualizes data points in chronological order, often used to show trends over time.
5. **Scatter Plot:** A scatter plot displays the relationship between two numerical variables using points in a Cartesian plane.
6. **Bubble Plot:** A bubble plot is an extension of a scatter plot where the size of the points (bubbles) represents an additional variable.

Steps to Perform in R:

1. **Set Up the Environment:**
   - Install and load necessary libraries.

R
Copy code
```r
install.packages("ggplot2")
install.packages("dplyr")
library(ggplot2)
library(dplyr)
```

2. **Load the Dataset:**
   - Load the crime dataset (replace `crime_data.csv` with your dataset's file name).

R
Copy code
```r
crime_data <- read.csv("crime_data.csv")
```

3. **Data Preprocessing:**
   - Inspect and clean the data if necessary (handle missing values, filter relevant columns, etc.).

R
Copy code
```r
crime_data <- crime_data %>% na.omit()
```

4. **Create Visualizations:**

Bar Chart:
R
Copy code
```r
ggplot(crime_data, aes(x = Crime_Type)) +
  geom_bar() +
  ggtitle("Number of Crimes by Type") +
  xlab("Type of Crime") + ylab("Number of Incidents")
```

   - **Observation:** This chart shows the frequency of each crime type, helping identify the most and least common crimes.

Pie Chart:

R

Copy code

```R
crime_summary <- crime_data %>% group_by(Crime_Type) %>%
summarise(Count = n())
ggplot(crime_summary, aes(x = "", y = Count, fill = Crime_Type)) +
  geom_bar(width = 1, stat = "identity") +
  coord_polar("y") +
  ggtitle("Crime Type Distribution")
```

  ○ **Observation:** The pie chart illustrates the proportion of each crime type within
     the dataset, revealing dominant categories.

Histogram:

R

Copy code

```R
ggplot(crime_data, aes(x = Crime_Severity)) +
  geom_histogram(binwidth = 1) +
  ggtitle("Distribution of Crime Severity") +
  xlab("Crime Severity") + ylab("Frequency")
```

  ○ **Observation:** The histogram shows the distribution of crime severity, identifying
     common severity levels.

Timeline Chart:

R

Copy code

```R
ggplot(crime_data, aes(x = Date, y = Crime_Count)) +
  geom_line() +
  ggtitle("Trend of Crimes Over Time") +
  xlab("Date") + ylab("Number of Crimes")
```

  ○ **Observation:** The timeline chart highlights trends in crime frequency over time,
     revealing periods of high and low crime rates.

Scatter Plot:

R

Copy code

```R
ggplot(crime_data, aes(x = Crime_Severity, y = Police_Response_Time))
+
  geom_point() +
  ggtitle("Crime Severity vs Police Response Time") +
  xlab("Crime Severity") + ylab("Response Time (minutes)")
```

  ○ **Observation:** The scatter plot examines the relationship between crime severity

and police response time, identifying any correlation.

Bubble Plot:
R
Copy code
```
ggplot(crime_data, aes(x = Crime_Severity, y = Police_Response_Time,
size = Crime_Count)) +
  geom_point(alpha = 0.5) +
  ggtitle("Crime Severity, Response Time, and Frequency") +
  xlab("Crime Severity") + ylab("Response Time (minutes)")
```

   ○ **Observation:** The bubble plot adds another dimension by showing the frequency
     of crimes with varying severity and response times.

Outcomes:

   ● Successfully created multiple types of charts using R to visualize crime data.
   ● Gained insights into the distribution, frequency, and relationships within the crime
     dataset.
   ● Developed an understanding of how different chart types can be used to analyze and
     present data effectively.

Conclusion:

This experiment demonstrated the power of data visualization in uncovering patterns and trends
in a crime dataset. By using R, we efficiently created visual representations that allowed us to
explore the data from different perspectives, leading to better-informed conclusions.

Submission

https://docs.google.com/forms/d/e/1FAIpQLScrs2lYhqrJPz7M9IVvAqZn3M8cM5H1hc58ZpJRC
O2jxcWuxg/viewform?usp=sf_link