Name	Ayush Kedare
UID	2021300058
Experiment	4

Aim : Create basic charts using R programming language on dataset Crime.

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

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.

1. Bar plot :

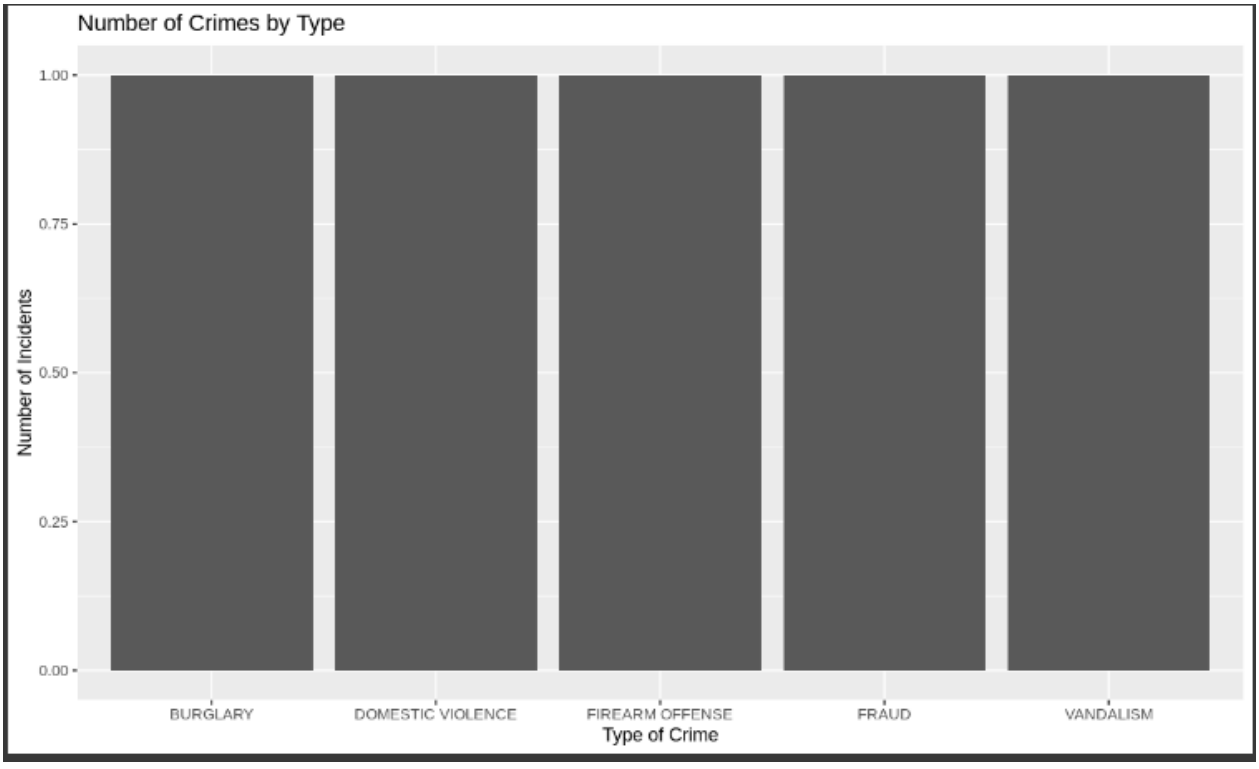
Code :

```
crime_type = crime_data %>% group_by(Crime.Description)   %>%
                    summarise(count = n(),
                              .groups = 'drop')
top_5_crimes <- crime_type %>%
  arrange(desc(count)) %>%
  slice_head(n = 5)

options(repr.plot.width = 10, repr.plot.height = 6)  # Adjust width and
height as needed

ggplot(top_5_crimes, aes(x = Crime.Description)) +
geom_bar() +
ggtitle("Number of Crimes by Type") +
xlab("Type of Crime") + ylab("Number of Incidents")
```

Output :



Observation : We can see that the top 5 crimes are burglary, domestic violence, firearm offense, fraud and vandalism.

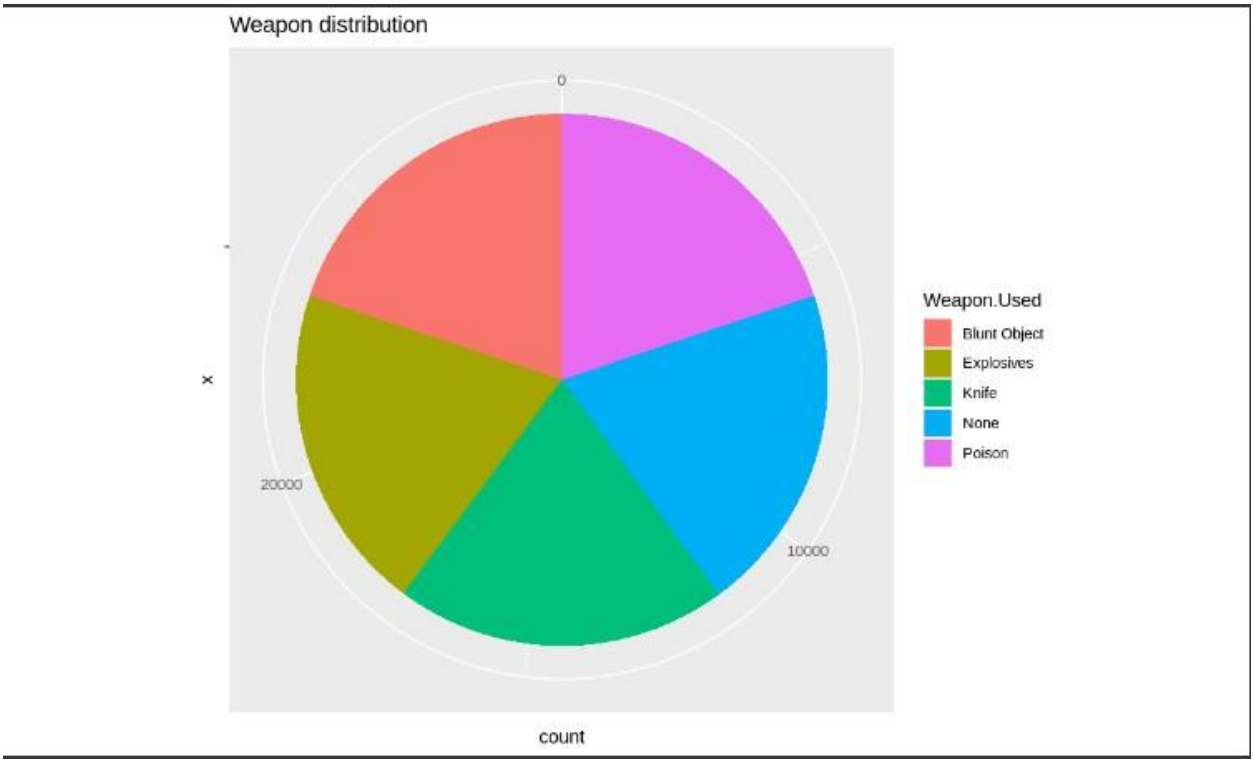
2. Pie plot :

Code :

```
weapon_type = crime_data %>% group_by(Weapon.Used)   %>%
                    summarise(count = n(),
                              .groups = 'drop')
top_5_weapon_type <- weapon_type %>%
  arrange(desc(count)) %>%
  slice_head(n = 5)
options(repr.plot.width = 10, repr.plot.height = 6)  # Adjust width and
height as needed

ggplot(top_5_weapon_type, aes(x = "", y = count, fill = Weapon.Used)) +
geom_bar(width = 1, stat = "identity") +
coord_polar("y") +
ggtitle("Weapon distribution")
```

Output :



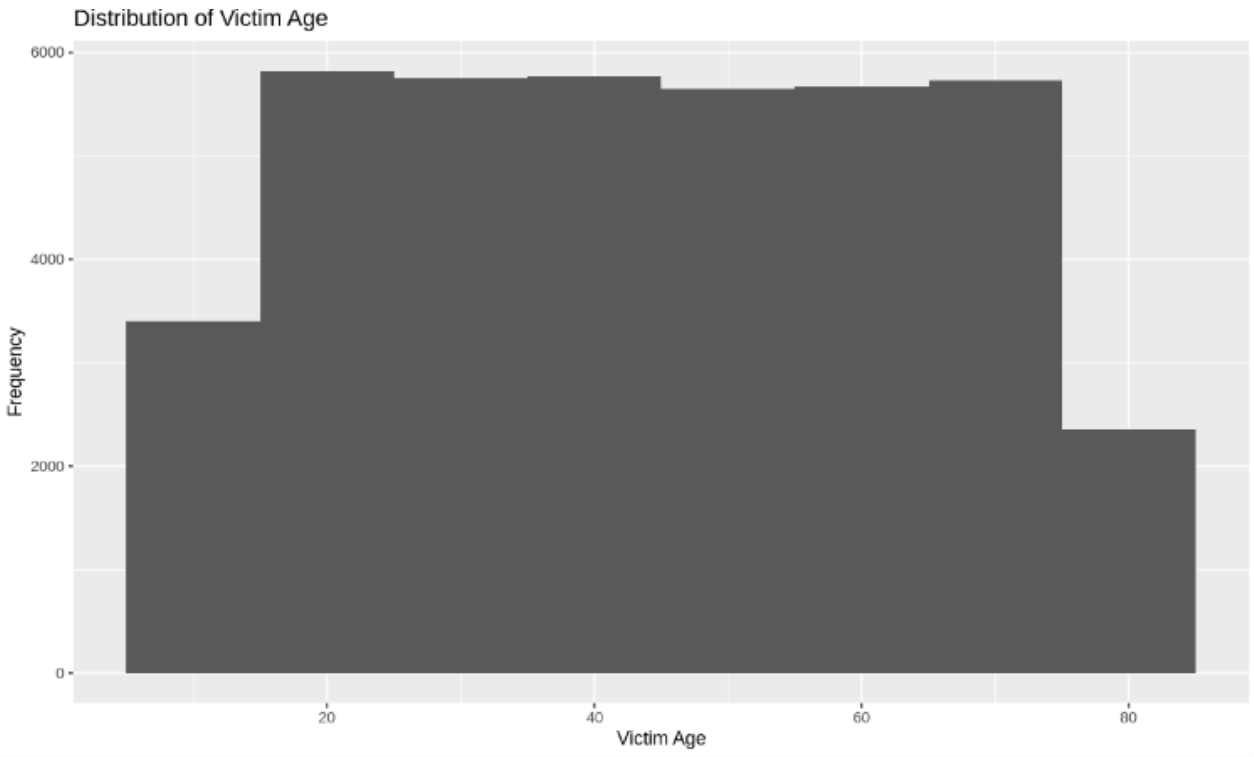
Observation : Top weapons used in crimes are blunt objects, explosives, knives and poison.

3.Histogram plot :

Code :

```
ggplot(crime_data, aes(x = Victim.Age)) +
geom_histogram(binwidth = 10) +
ggtitle("Distribution of Victim Age") +
xlab("Victim Age") + ylab("Frequency")
```

Output :



Observation : Most of the victims are of age from 20 to 75.

4 .Timeseries plot :

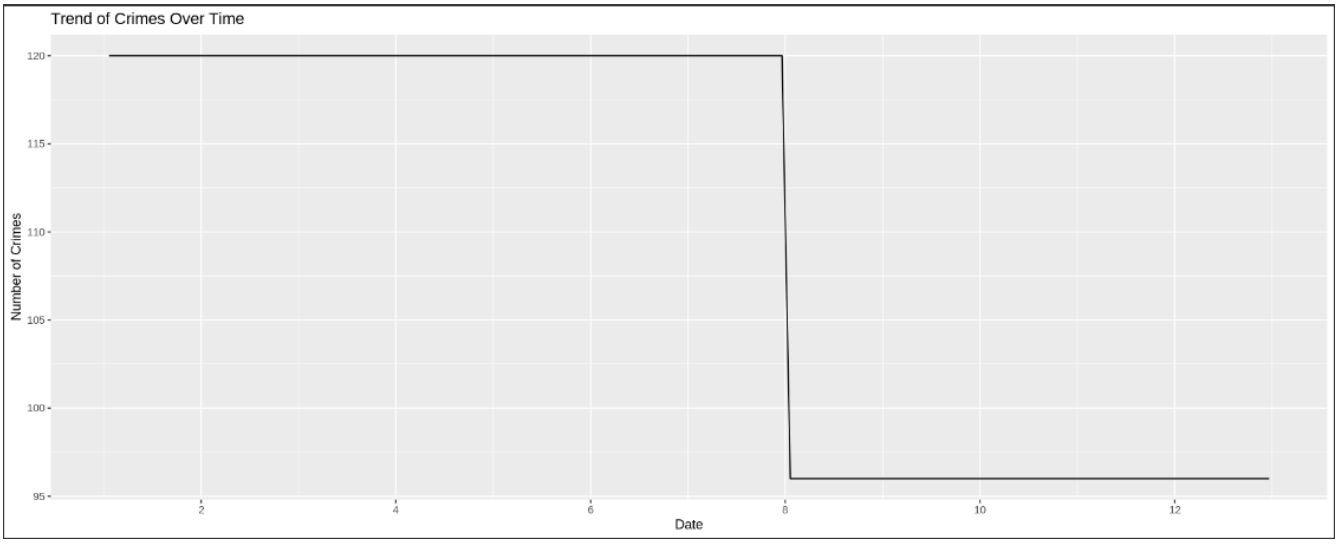
Code :

```
crime_data$date_only <- as.Date(crime_data$Date.of.Occurrence)
crimes_according_date = crime_data %>% group_by(date_only)  %>%
                        summarise(count = n(),
                        .groups = 'drop')
crimes_according_date <- crimes_according_date %>% na.omit()
```

```
options(repr.plot.width = 15, repr.plot.height = 6)  # Adjust width and
height as needed

ggplot(crimes_according_date, aes(x = date_only, y = count)) +
geom_line() +
ggtitle("Total crimes committed over time in 2020") +
xlab("Month") + ylab("Number of Crimes")
```

Output :



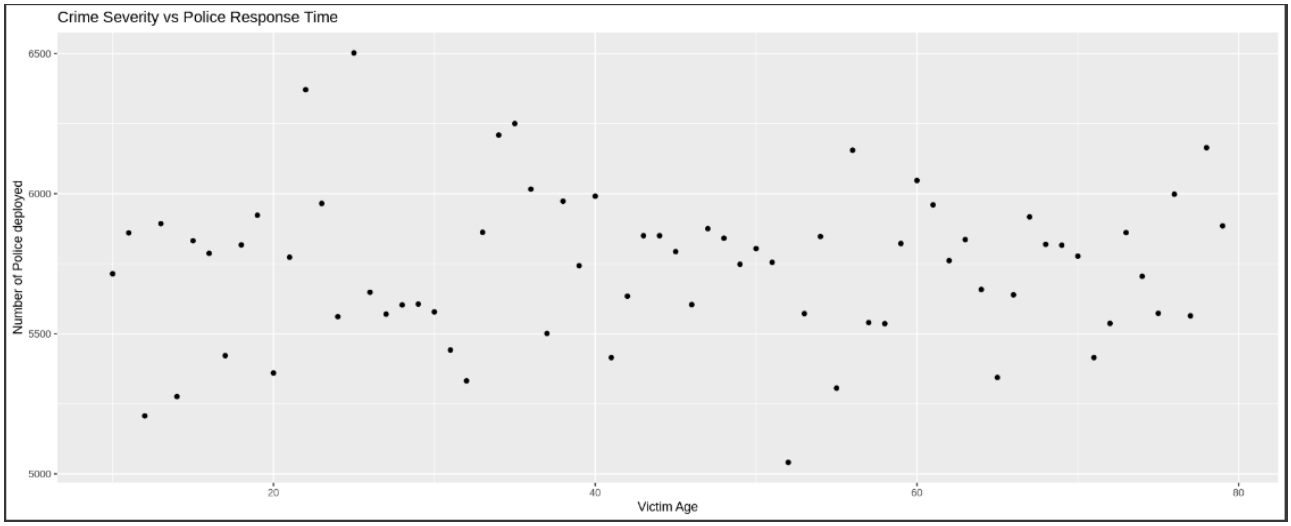
Observation : Most of the crimes took place from January to July in 2020. After July there was a sudden drop in the number of crimes.

5. Scatter plot :

Code :

```
police_count = crime_data %>% group_by(Victim.Age)   %>%
                          summarise(police_deployed = sum(Police.Deployed),
                          .groups = 'drop')
ggplot(police_count, aes(x = Victim.Age, y = police_deployed))
+geom_point() + ggtitle("Victim age Vs Total police deployed") +
xlab("Victim Age") + ylab("Number of Police deployed")
```

Output :



Observation : There seems to be no relation between the victim's age and number of police deployed.

Conclusion:

- 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.