

Name: Sarthak Gharat

UID: 2021700025

Branch: CSE DS

## Expt 4: Crime Data

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

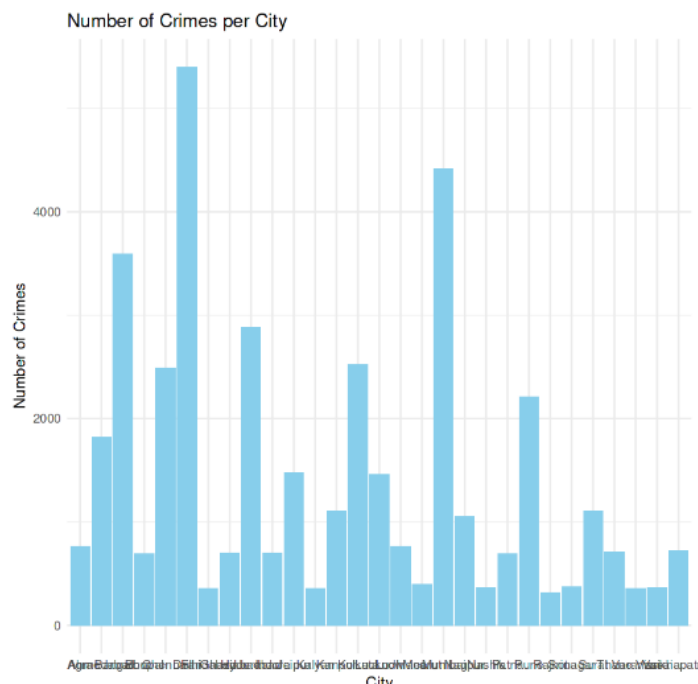
- Basic - Bar chart, Pie chart, Histogram, Scatter plot, Bubble plot
- Write observations from each chart

**Software Used:** R

**Dataset Used:** Indian Crimes Dataset

### 1. Bar Chart

```
# Bar chart for the number of crimes per city
ggplot(crime_data, aes(x = City)) +
  geom_bar(fill = "skyblue") +
  labs(title = "Number of Crimes per City", x = "City", y = "Number of Crimes") +
  theme_minimal()
```



Delhi and Mumbai have the highest crime rates.

### 2. Pie Chart

```
[9]: # Pie chart for the distribution of crimes by weapon used
weapon_distribution <- table(crime_data$Weapon.Used)

# Create the pie chart
pie(weapon_distribution,
    labels = names(weapon_distribution),
    main = "Distribution of Crimes by Weapon Used",
    col = rainbow(length(weapon_distribution)))
```

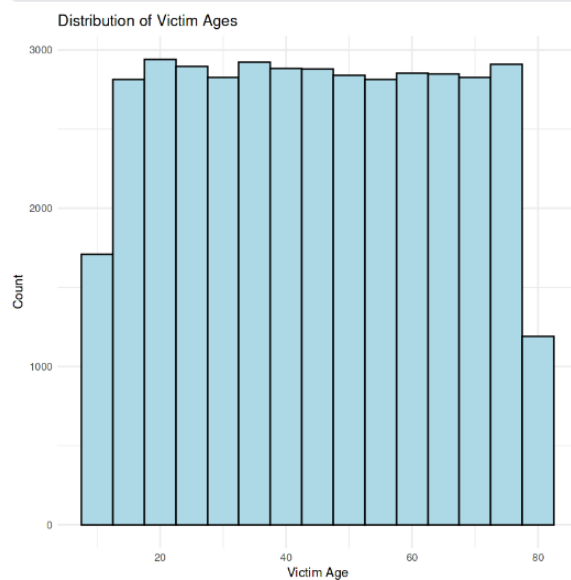
**Distribution of Crimes by Weapon Used**



There is almost equal distribution in the means used.

### 3. Histogram

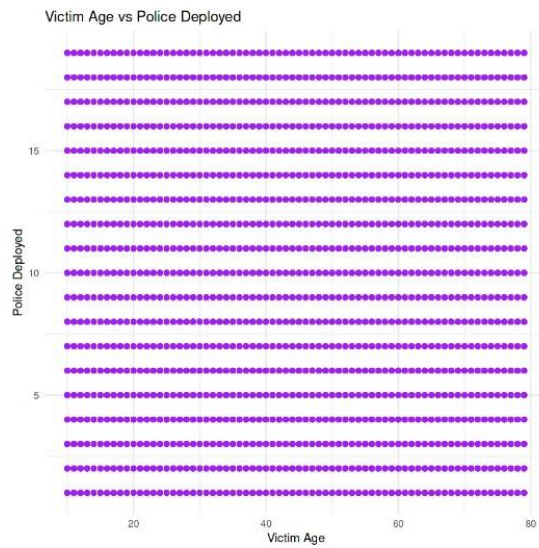
```
# Histogram for victim age
ggplot(crime_data, aes(x = Victim.Age)) +
  geom_histogram(binwidth = 5, fill = "lightblue", color = "black") +
  labs(title = "Distribution of Victim Ages", x = "Victim Age", y = "Count") +
  theme_minimal()
```



It's somewhat a flat normal distribution over age.

#### 4. Scatter Plot

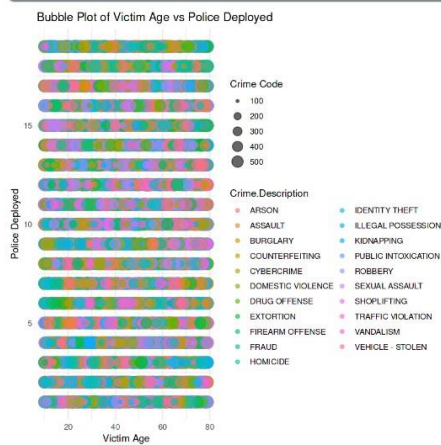
```
[12]: # Scatter plot for victim age vs police deployed
ggplot(crime_data, aes(x = Victim.Age, y = Police.Deployed)) +
  geom_point(color = "purple") +
  labs(title = "Victim Age vs Police Deployed", x = "Victim Age", y = "Police Deployed") +
  theme_minimal()
```



The number of police deployed is almost uniform with victim age.

#### 5. Bubble Plot

```
# Bubble plot for crime type, victim age, and police deployed
ggplot(crime_data, aes(x = Victim.Age, y = Police.Deployed, size = Crime.Code, color = Crime.Description)) +
  geom_point(alpha = 0.6) +
  labs(title = "Bubble Plot of Victim Age vs Police Deployed", x = "Victim Age", y = "Police Deployed", si
  theme_minimal()
```



The different categories can be observed.