

ADV EXP 4

Name : Krutika Rupchand Shinde

Batch : B

UID : 2021300119

```
!apt-get install r-base
```

```
➞ Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
r-base is already the newest version (4.4.1-1.2204.0).
0 upgraded, 0 newly installed, 0 to remove and 49 not upgraded.
```

```
!pip install rpy2
```

```
➞ Requirement already satisfied: rpy2 in /usr/local/lib/python3.10/dist-packages (3.4.2)
Requirement already satisfied: cffi>=1.10.0 in /usr/local/lib/python3.10/dist-packages (from rpy2) (1.17.1)
Requirement already satisfied: Jinja2 in /usr/local/lib/python3.10/dist-packages (from rpy2) (3.1.4)
Requirement already satisfied: pytz in /usr/local/lib/python3.10/dist-packages (from rpy2) (2024.1)
Requirement already satisfied: tzlocal in /usr/local/lib/python3.10/dist-packages (from rpy2) (5.2)
Requirement already satisfied: pycparser in /usr/local/lib/python3.10/dist-packages (from cffi>=1.10.0->rpy2) (2.22)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from Jinja2->rpy2) (2.1.5)
```

```
# Load rpy2 to enable R code execution
%load_ext rpy2.ipython
```

```
%%R
# Install necessary packages
install.packages("ggplot2")
install.packages("dplyr")

# Loading the libraries
library(ggplot2)
library(dplyr)

# Load the dataset
crime_data <- read.csv("crime_data.csv")

# Display the structure of the dataset
str(crime_data)
```



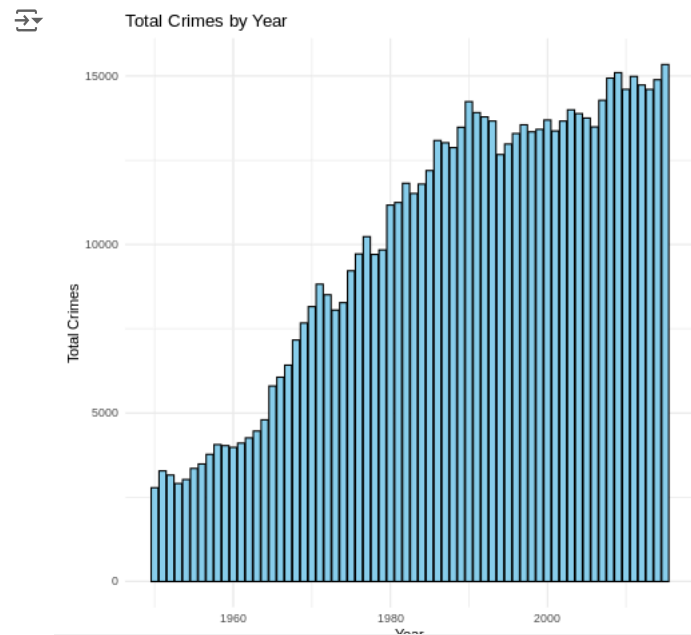
WARNING: rpy2.rinterface.lib.callbacks.R[write to console]: The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
'data.frame': 66 obs. of 21 variables:
 $ Year      : int  1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 ...
 $ crimes.total : int  2784 3284 3160 2909 3028 3357 3488 3774 4064 4033 ...
 $ crimes.penal.code : int  2306 2754 2608 2689 2791 3101 3215 3520 3791 3733 ...
 $ crimes.person : int  120 125 119 119 126 135 133 133 127 125 ...
 $ murder      : int  1 1 1 1 1 1 1 1 1 1 ...
 $ assault     : int  105 109 104 105 107 118 116 116 113 110 ...
 $ sexual.offenses : int  40 45 39 45 41 44 38 36 40 47 ...
 $ rape        : int  5 6 4 5 5 5 5 5 5 6 ...
 $ stealing.general : int  1578 1899 1846 1929 1981 2254 2363 2635 2880 2793 ...
 $ burglary    : int  295 342 372 361 393 459 470 580 724 715 ...
 $ house.theft : int  NA NA NA NA NA NA NA NA NA NA ...
 $ vehicle.theft : int  NA NA NA NA NA NA NA 245 279 238 ...
 $ out.of.vehicle.theft : int  NA NA NA NA NA NA NA NA NA NA ...
 $ shop.theft   : int  NA NA NA NA NA NA NA NA NA NA ...
 $ robbery      : int  3 3 3 4 4 4 6 6 6 6 ...
 $ fraud        : int  209 310 217 209 236 236 234 254 254 251 ...
 $ criminal.damage : int  72 73 82 88 101 111 133 155 167 179 ...
 $ other.penal.crimes : int  477 530 553 220 237 255 273 255 273 299 ...
 $ narcotics    : int  0 0 0 0 0 0 0 0 1 1 ...
 $ drunk.driving : int  49 66 78 91 103 125 160 163 166 181 ...
 $ population   : int  7014000 7073000 7125000 7171000 7213000 7262000 7315000 7364000 7409000 7446000 ...
```

%%R

```
# Bar Chart: Total crimes per year
ggplot(crime_data, aes(x = Year, y = crimes.total)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black") +
  labs(title = "Total Crimes by Year", x = "Year", y = "Total Crimes") +
  theme_minimal()
```



Observation : The bar chart reveals fluctuations in total crime rates over the years. Some periods show significant increases, potentially indicating changes in societal or law enforcement factors.

%%R

```
# Select data for the year 1950
year_data <- crime_data %>% filter(Year == 1950)

# Crime categories for the pie chart
crime_types <- c("murder", "assault", "sexual.offenses", "rape", "burglary", "robbery")
crime_counts <- year_data[1, crime_types]

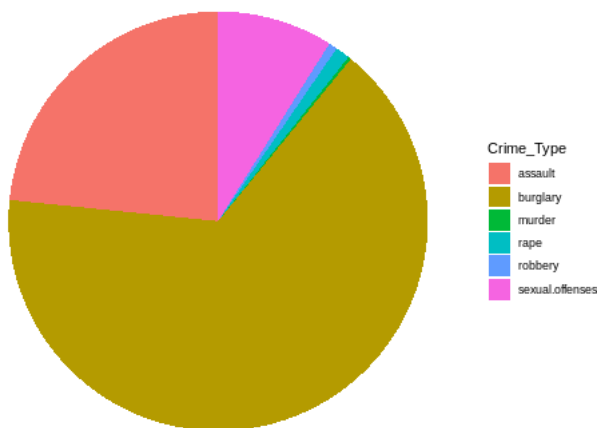
# Create a pie chart for crimes in 1950
```

```
crime_pie_data <- data.frame(
  Crime_Type = crime_types,
  Count = as.numeric(crime_counts)
)

ggplot(crime_pie_data, aes(x = "", y = Count, fill = Crime_Type)) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y") +
  labs(title = "Crime Distribution in 1950") +
  theme_void()
```



Crime Distribution in 1950

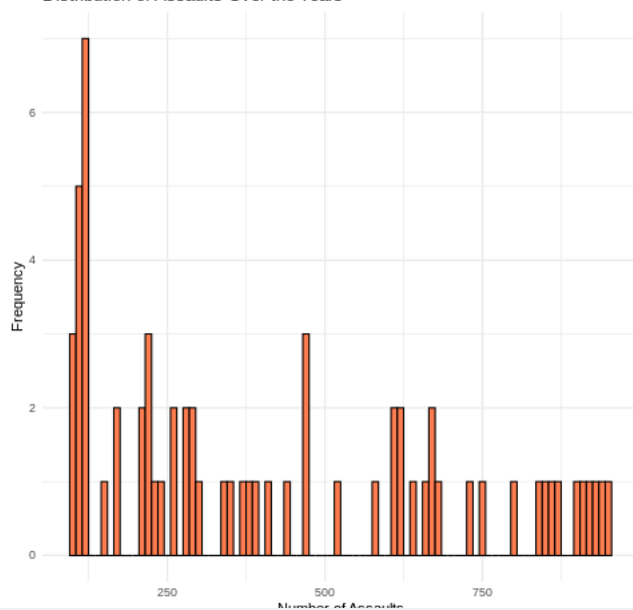


Observation : In 1950, crimes like stealing and burglary make up a large proportion of the total crimes, while violent crimes such as murder and sexual offenses are less frequent.

```
%%R
# Histogram of assaults over the years
ggplot(crime_data, aes(x = assault)) +
  geom_histogram(binwidth = 10, fill = "coral", color = "black") +
  labs(title = "Distribution of Assaults Over the Years", x = "Number of Assaults", y = "Frequency") +
  theme_minimal()
```

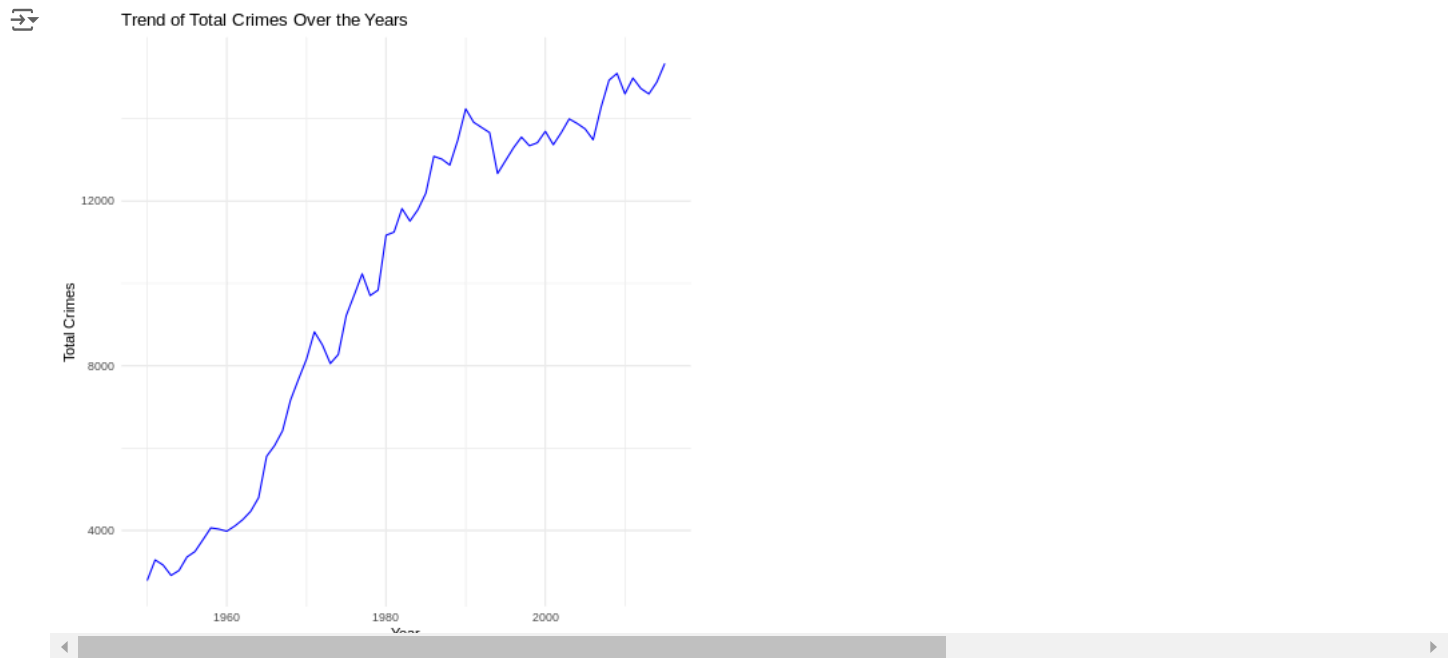


Distribution of Assaults Over the Years



Observation : The histogram indicates that assaults tend to cluster within a specific range, suggesting that most years have a consistent number of assaults with fewer extreme outliers.

```
%%R
# Time Line chart: Crime trend over the years
ggplot(crime_data, aes(x = Year, y = crimes.total)) +
  geom_line(color = "blue") +
  labs(title = "Trend of Total Crimes Over the Years", x = "Year", y = "Total Crimes") +
  theme_minimal()
```



Observation : The time line chart shows an upward trend in total crimes, with some years experiencing spikes, which could indicate socio-political changes or population growth during those periods.

```
%%R
# Scatter plot: Total crimes vs Population
ggplot(crime_data, aes(x = population, y = crimes.total)) +
  geom_point(color = "darkgreen") +
  labs(title = "Total Crimes vs Population", x = "Population", y = "Total Crimes") +
  theme_minimal()
```

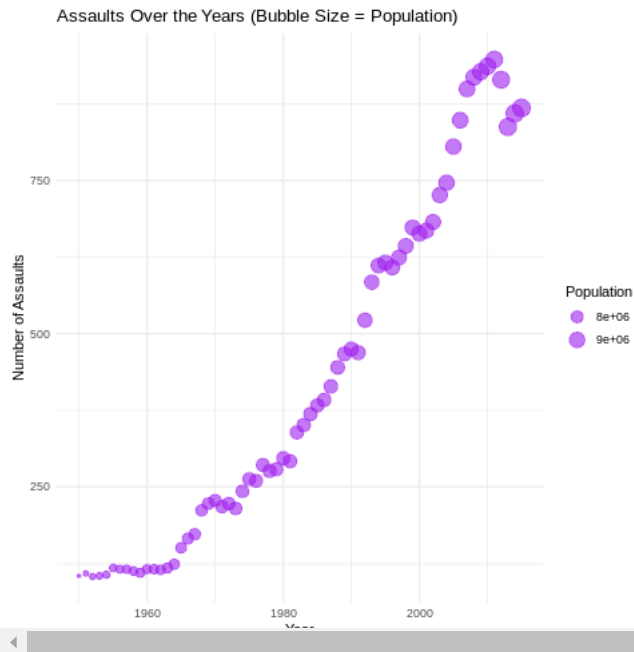


Total Crimes vs Population

Observation : The scatter plot shows a positive correlation between population size and the number of total crimes, implying that larger populations tend to experience more crimes.

%%R

```
# Bubble plot: Crimes with population size
ggplot(crime_data, aes(x = Year, y = assault, size = population)) +
  geom_point(alpha = 0.6, color = "purple") +
  labs(title = "Assaults Over the Years (Bubble Size = Population)",
       x = "Year", y = "Number of Assaults", size = "Population") +
  theme_minimal()
```



Observation : The bubble plot demonstrates that higher populations generally experience more assaults, with some larger bubbles indicating significant crime rates in years with growing populations.