Name - Amaan khan
Reg no - 21BCI0351
Campus - VIT vellore
Morning batch

Assignment 2

Question

1. Take car crashes dataset from seaborn library
2. load the dataset
3 .Perform Data Visualization
4.Inference is must for each and every graph

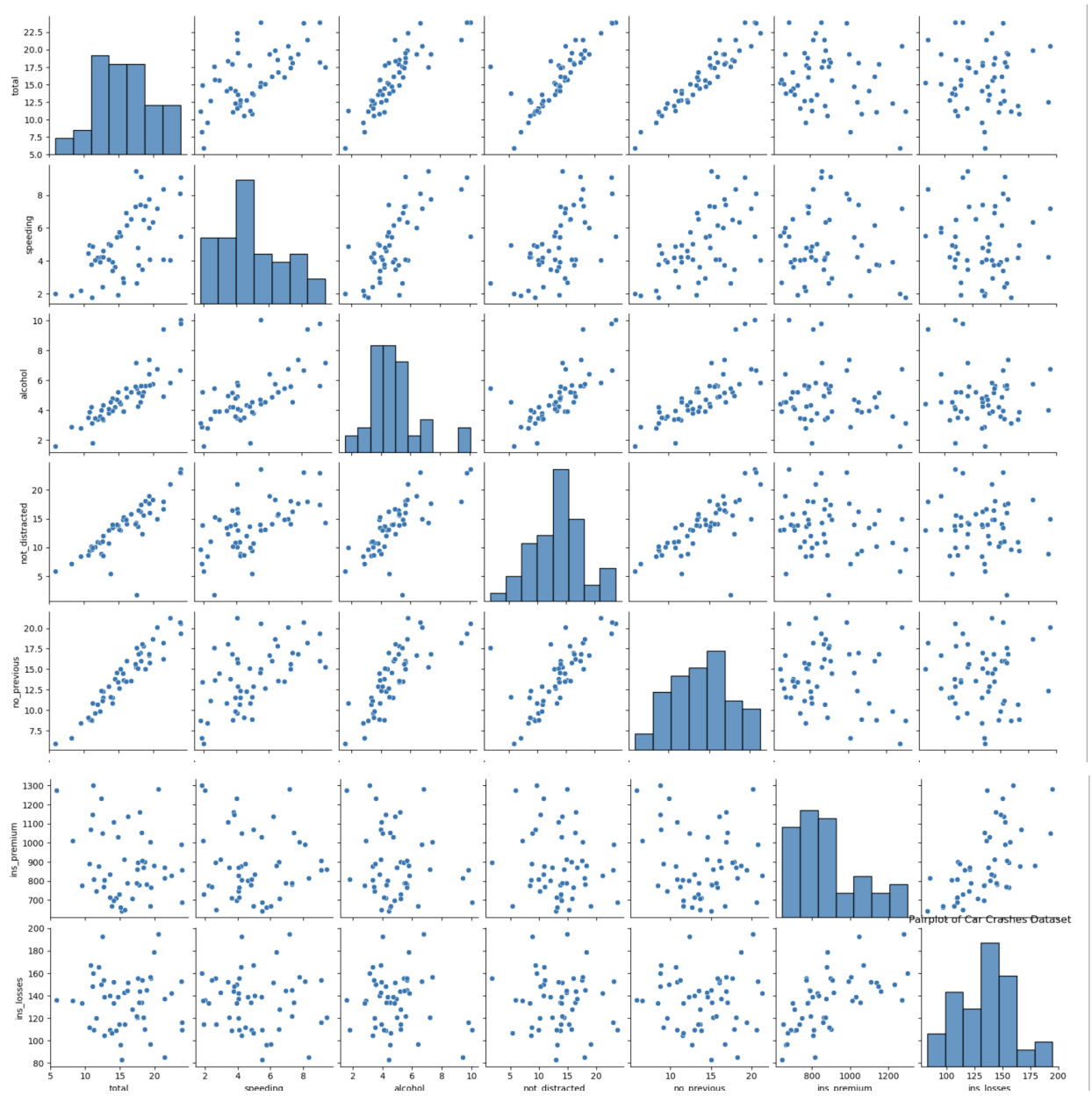
Solution

1. First we will import necessary libraries
2. We will load the car crashes dataset
3. As seaborn comes with a sample dataset “car_crashes” , so we can load directly from it.
4. After that we will perform data visualization
5. After every code of the data visualization we have also written the inference for the better understanding of the graph.

Code

```
import seaborn as sns
import matplotlib.pyplot as plt
car_crashes = sns.load_dataset("car_crashes")

    # Visualization 1: Pairplot
    # A pairplot allows you to visualize relationships between numerical variables.
sns.pairplot(car_crashes)
plt.title("Pairplot of Car Crashes Dataset")
plt.show()
```

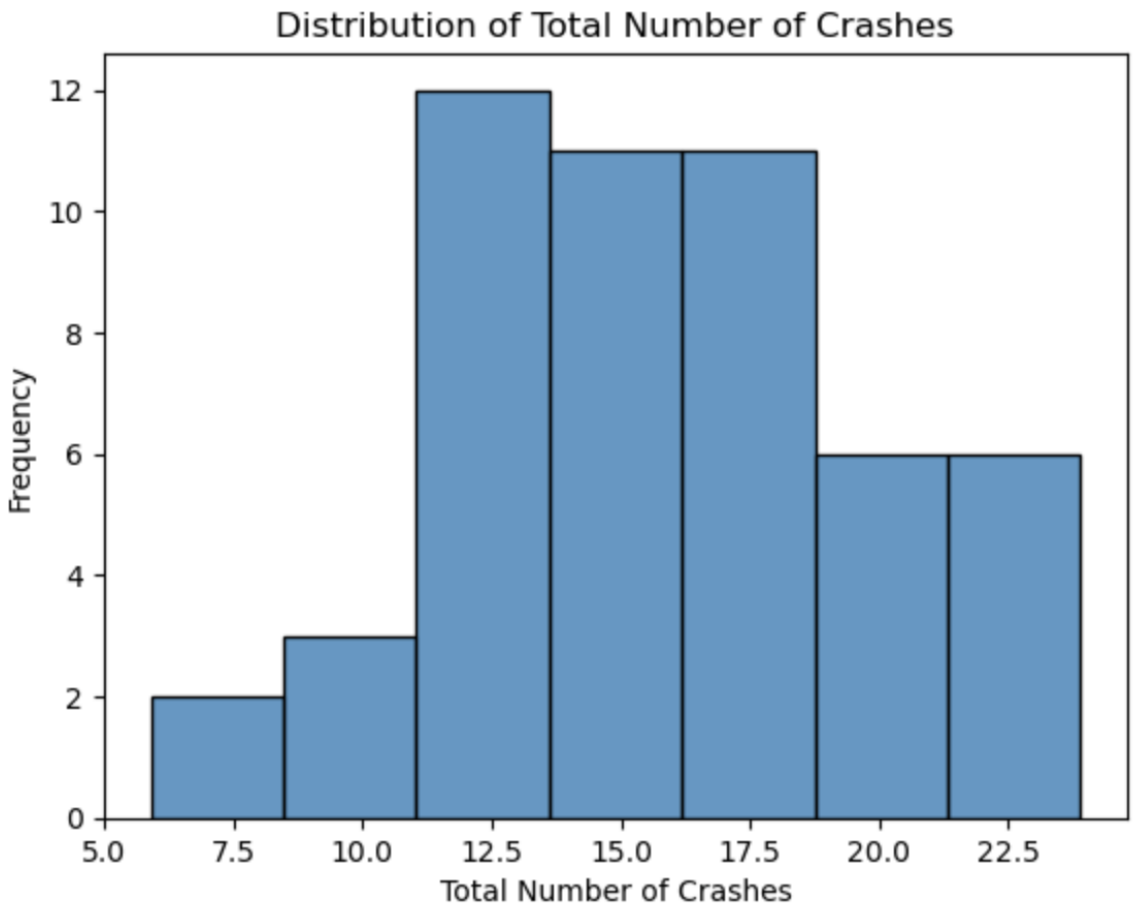


#* Inference: Use this plot to identify any patterns or relationships between numerical variables.

#Visualization 2: Histogram
A histogram helps you visualize the distribution of a single numerical variable.

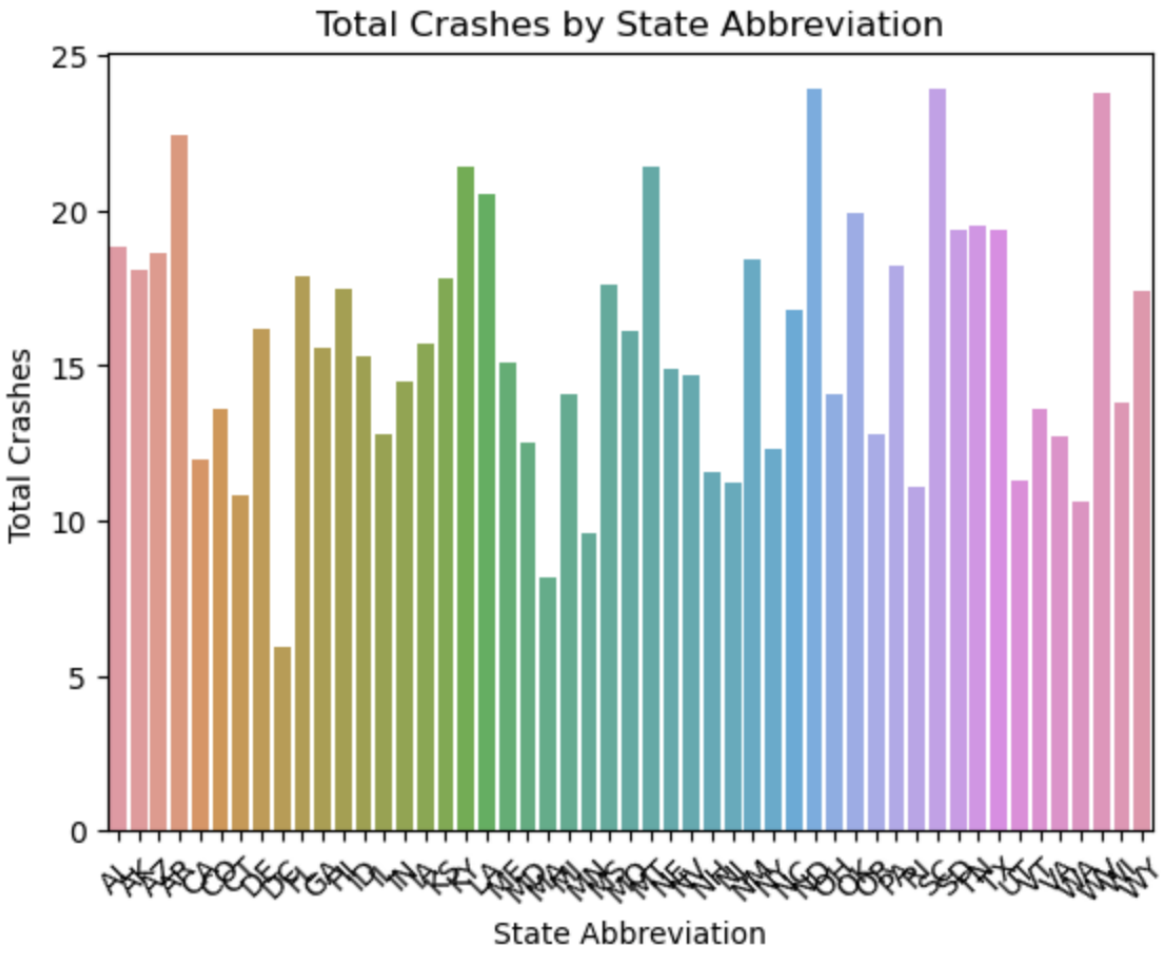
```
sns.histplot(car_crashes["total"])
plt.title("Distribution of Total Number of Crashes")
plt.xlabel("Total Number of Crashes")
plt.ylabel("Frequency")
```

plt.show()

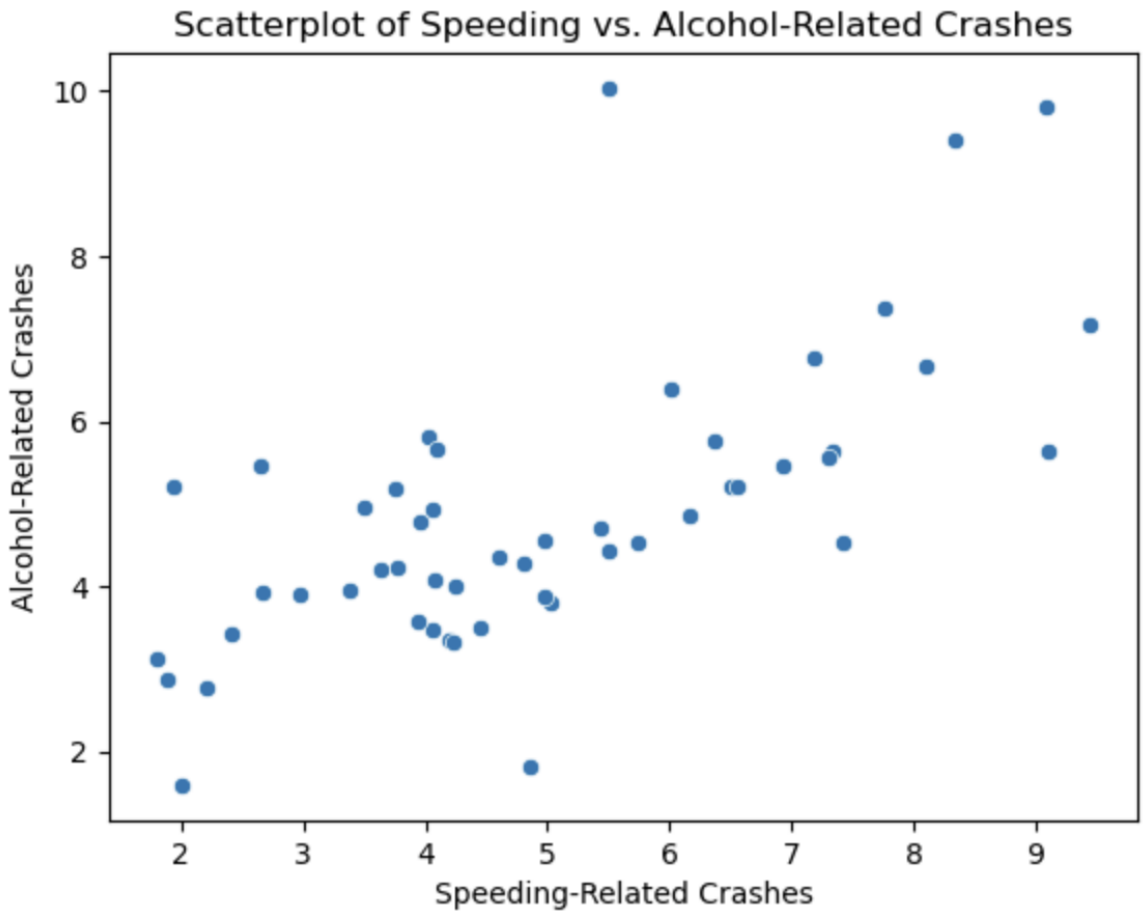


#* Inference: This histogram shows the distribution of total crashes, and you can see if it's skewed or normal.

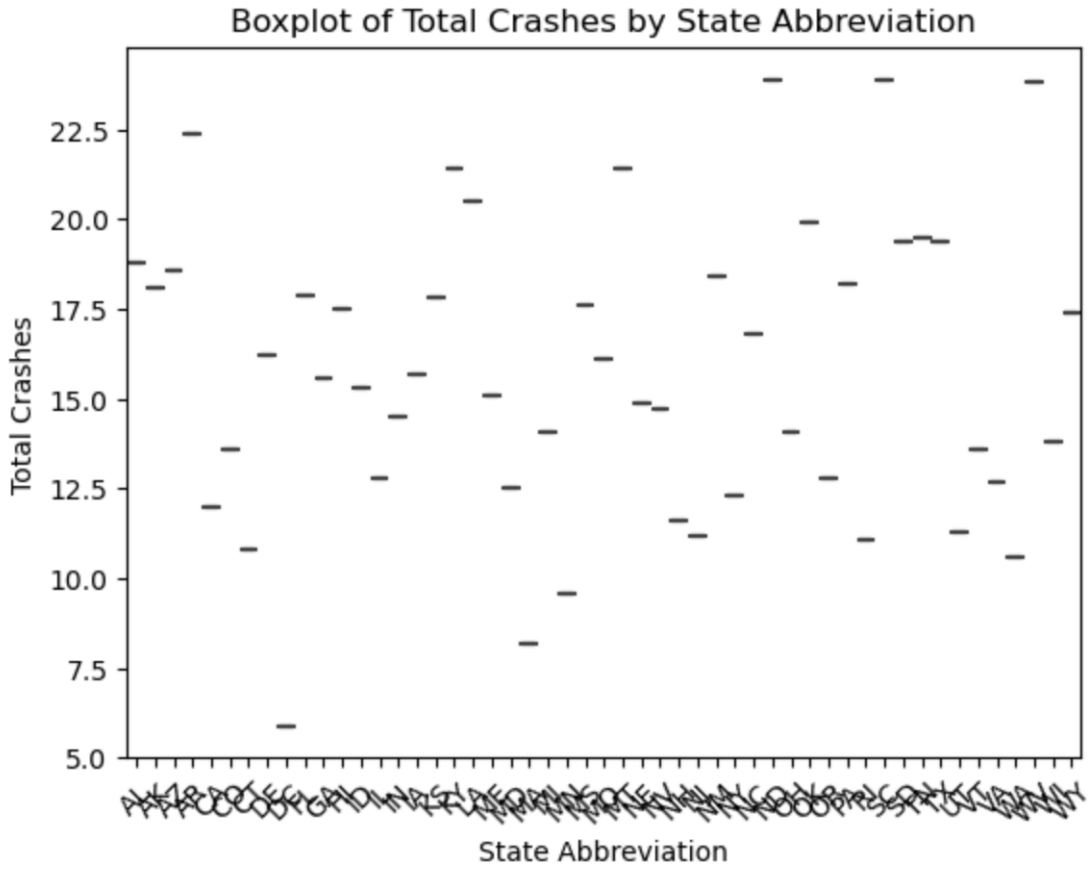
# Visualization 3: Barplot
# A barplot can help visualize categorical data, such as "abbrev."
sns.barplot(x="abbrev", y="total", data=car_crashes)
plt.title("Total Crashes by State Abbreviation")
plt.xlabel("State Abbreviation")
plt.ylabel("Total Crashes")
plt.xticks(rotation=45)
plt.show()



#* Inference: This barplot displays the total number of crashes for each state, allowing you to compare them.

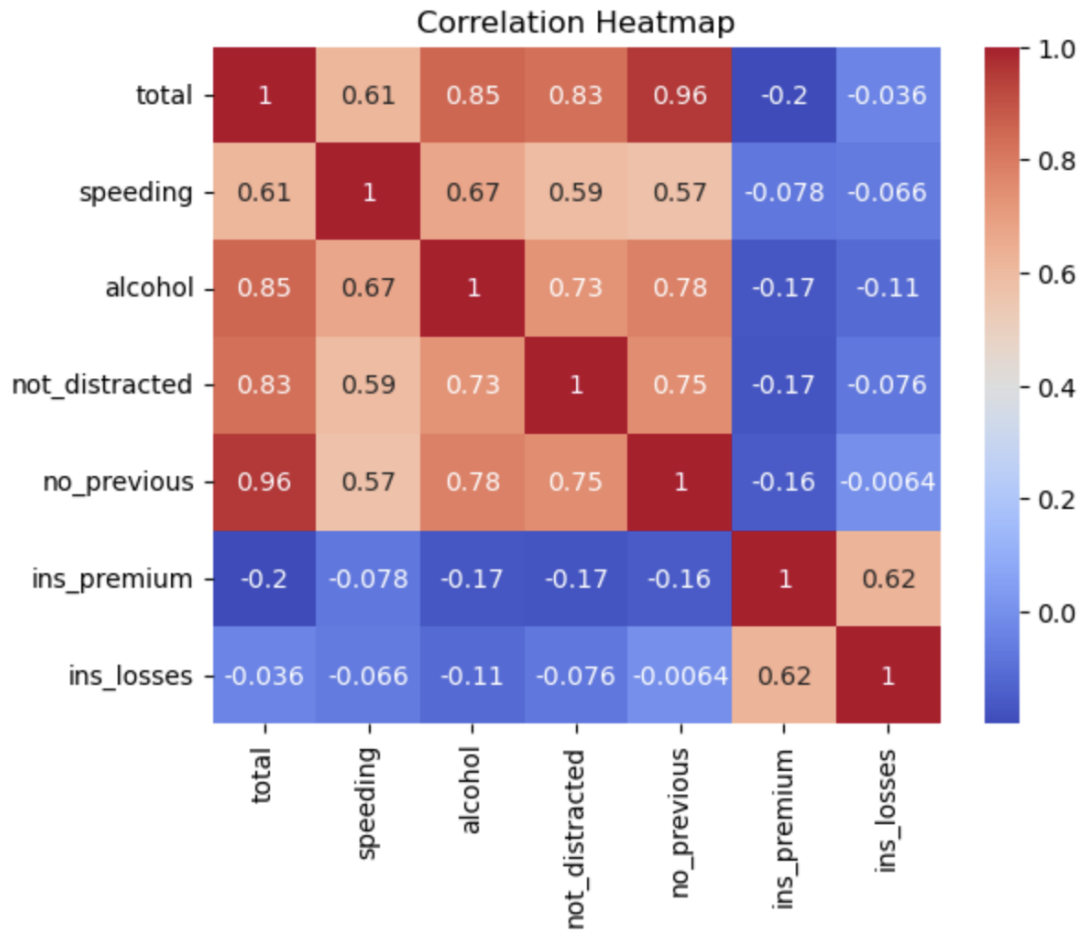


#* Inference: Use this scatterplot to observe if there is a correlation between speeding-related and alcohol-related



#* Inference: Boxplots can help identify any outliers and compare the spread of total crashes by state

# Visualization 6: Heatmap (Correlation)
# A heatmap helps visualize the correlation between numerical variables.
correlation_matrix = car_crashes.corr()
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm")
plt.title("Correlation Heatmap")
plt.show()



#* Inference: The heatmap shows the correlation between different numerical variables in the dataset.

21BCI0351
Amaan khan