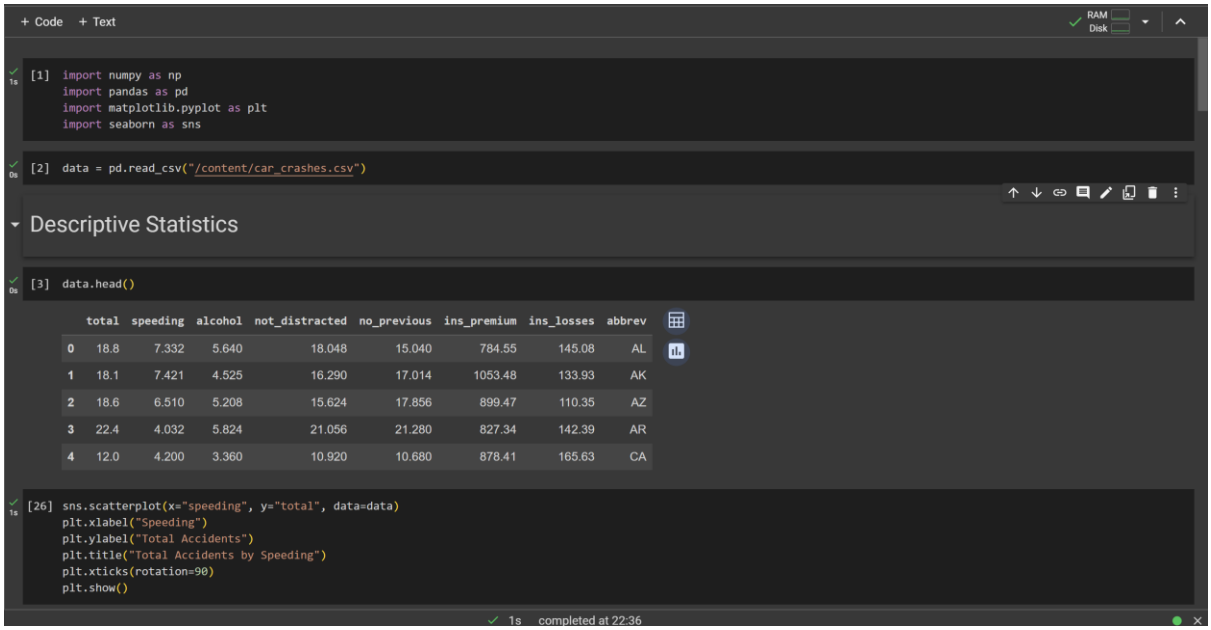
SmartInternz

AIML – Assignment 3

Rahul Palanivel

21BCE7828

Assignment 8 th September:
1.Take car_crashes dataset from seaborn library
2.load the dataset
3.Perfrom Data Visualization
4.Inference is must for each and every graph
5.Submit it by wednesday in html format



```
[1] import numpy as np
    import pandas as pd
    import matplotlib.pyplot as plt
    import seaborn as sns
```

```
[2] data = pd.read_csv("/content/car_crashes.csv")
```

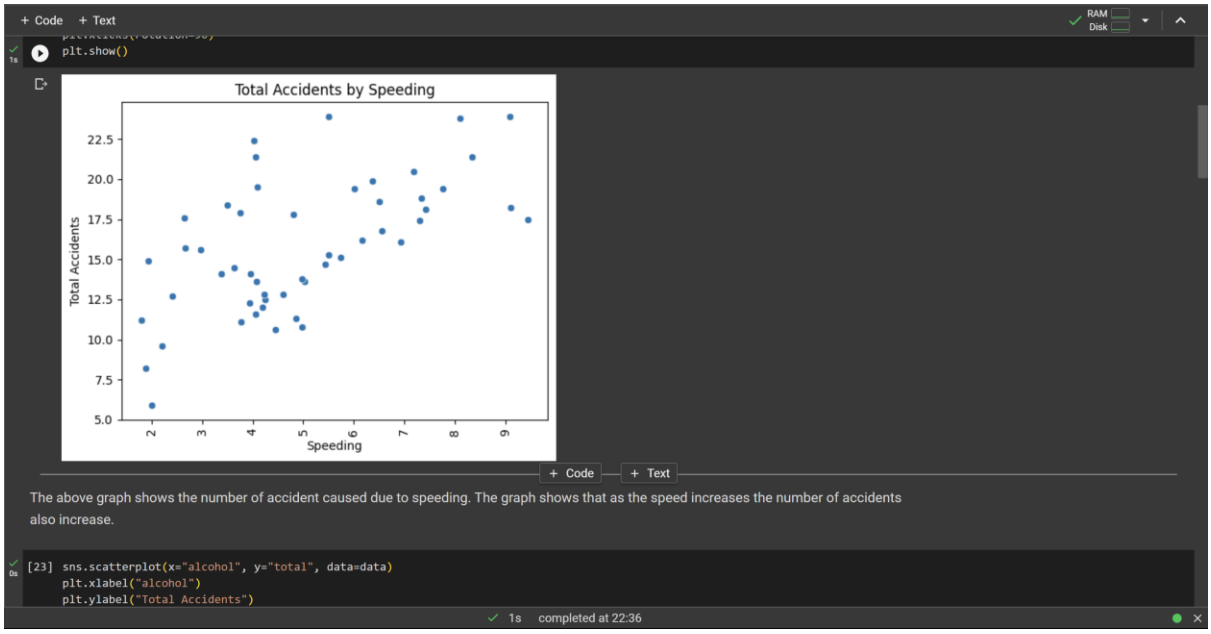
Descriptive Statistics

```
[3] data.head()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
0	18.8	7.332	5.640	18.048	15.040	784.55	145.08	AL
1	18.1	7.421	4.525	16.290	17.014	1053.48	133.93	AK
2	18.6	6.510	5.208	15.624	17.856	899.47	110.35	AZ
3	22.4	4.032	5.824	21.056	21.280	827.34	142.39	AR
4	12.0	4.200	3.360	10.920	10.680	878.41	165.63	CA

```
[26] sns.scatterplot(x="speeding", y="total", data=data)
     plt.xlabel("Speeding")
     plt.ylabel("Total Accidents")
     plt.title("Total Accidents by Speeding")
     plt.xticks(rotation=90)
     plt.show()
```

1s completed at 22:36



```
plt.show()
```


+ Code + Text

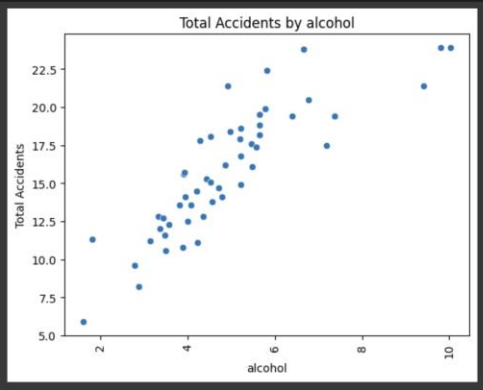
The above graph shows the number of accident caused due to speeding. The graph shows that as the speed increases the number of accidents also increase.

```
[23] sns.scatterplot(x="alcohol", y="total", data=data)
     plt.xlabel("alcohol")
     plt.ylabel("Total Accidents")
```

1s completed at 22:36

+ Code + Text

```
[23] sns.scatterplot(x="alcohol", y='total', data=data)
     plt.xlabel("alcohol")
     plt.ylabel("Total Accidents")
     plt.title("Total Accidents by alcohol")
     plt.xticks(rotation=90)
     plt.show()
```

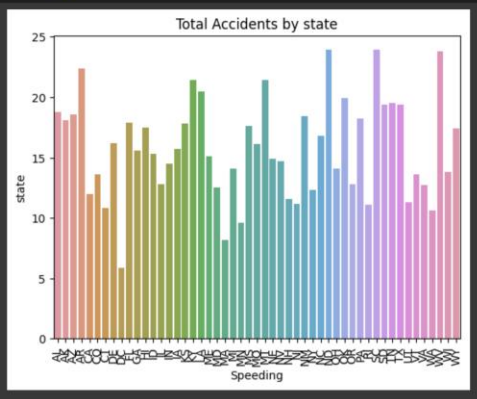


The above graph shows the number of accident caused due to alcohol. The graph shows that as the alcohol consumption increases the number of accidents also increase.

1s completed at 22:36

+ Code + Text

```
sns.barplot(x="abbrev", y="total", data=data)
plt.xlabel("Speeding")
plt.ylabel("state")
plt.title("Total Accidents by state")
plt.xticks(rotation=90)
plt.show()
```



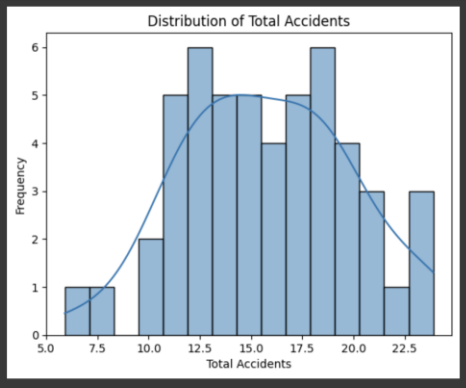
The above graph shows the number of accident caused in each state

1s completed at 22:36

+ Code + Text

The above graph shows the number of accident caused in each state

```
[28] sns.histplot(data['total'], bins=15, kde=True)
     plt.xlabel("Total Accidents")
     plt.ylabel("Frequency")
     plt.title("Distribution of Total Accidents")
     plt.show()
```




```
[34] sns.lineplot(x='speeding', y='total', data=data)
```

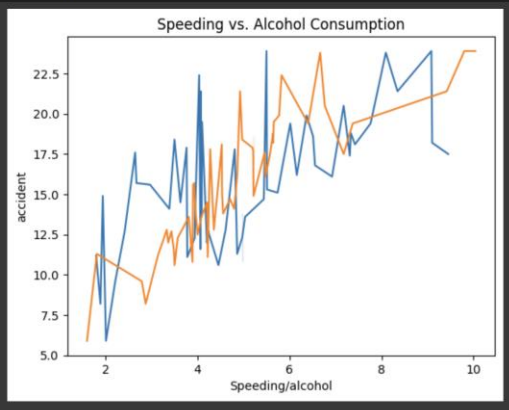
1s completed at 22:36

+ Code + Text

RAM Disk

```
sns.lineplot(x='speeding', y='total', data=data)
sns.lineplot(x='alcohol', y='total', data=data)

plt.xlabel("Speeding/alcohol")
plt.ylabel("accident")
plt.title("Speeding vs. Alcohol Consumption")
plt.show()
```



The above graph shows the comparision between accidents due to speed and accidents due to alcohol

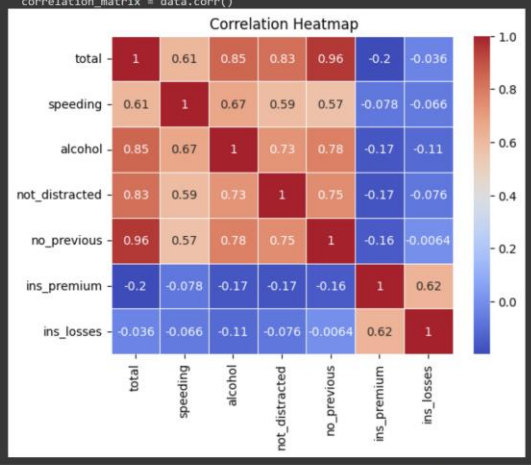
✓ 1s completed at 22:36

+ Code + Text

RAM Disk

```
correlation_matrix = data.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title("Correlation Heatmap")
plt.show()
```

```
<ipython-input-35-967ac05f8eb0>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select
  correlation_matrix = data.corr()
```



✓ 1s completed at 22:36