Name :- Abhijeet Bhardwaj

Assignment - 2

```python
import seaborn as sns

dataset_names = sns.get_dataset_names()
print(dataset_names)
```

```
['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes',
'diamonds', 'dots', 'dowjones', 'exercise', 'flights', 'fmri',
'geyser', 'glue', 'healthexp', 'iris', 'mpg', 'penguins', 'planets',
'seaice', 'taxis', 'tips', 'titanic']
```

```python
#Load the data
car_crashes = sns.load_dataset('car_crashes')
car_crashes.head()
```

|	total	speeding	alcohol	not_distracted	no_previous	ins_premium
0	18.8	7.332	5.640	18.048	15.040	784.55
1	18.1	7.421	4.525	16.290	17.014	1053.48
2	18.6	6.510	5.208	15.624	17.856	899.47
3	22.4	4.032	5.824	21.056	21.280	827.34
4	12.0	4.200	3.360	10.920	10.680	878.41

	ins_losses	abbrev
0	145.08	AL
1	133.93	AK
2	110.35	AZ
3	142.39	AR
4	165.63	CA

```python
# Import matplotlib to plot data.
import matplotlib.pyplot as plt

#Let's plot location vs total accidents.
x_location = "abbrev"
y_total = "total"

sns.scatterplot(x=x_location, y=y_total, data=car_crashes)

plt.xlabel(x_location)
plt.ylabel(y_total)
plt.show()
```
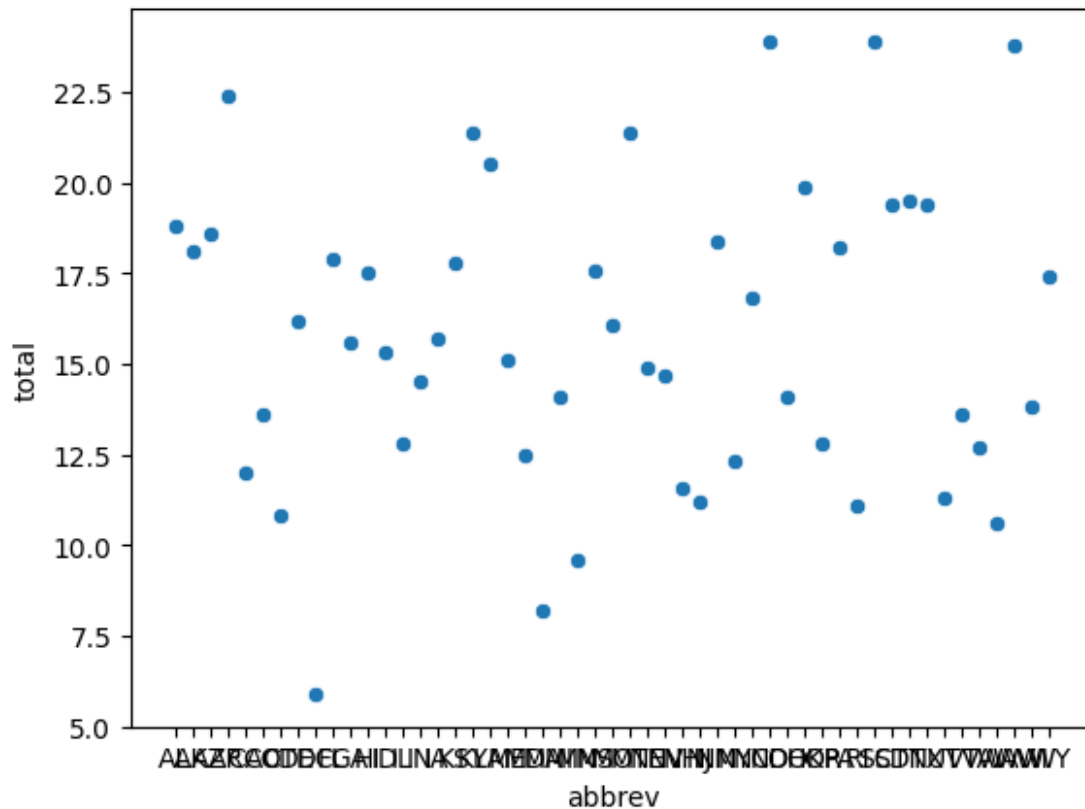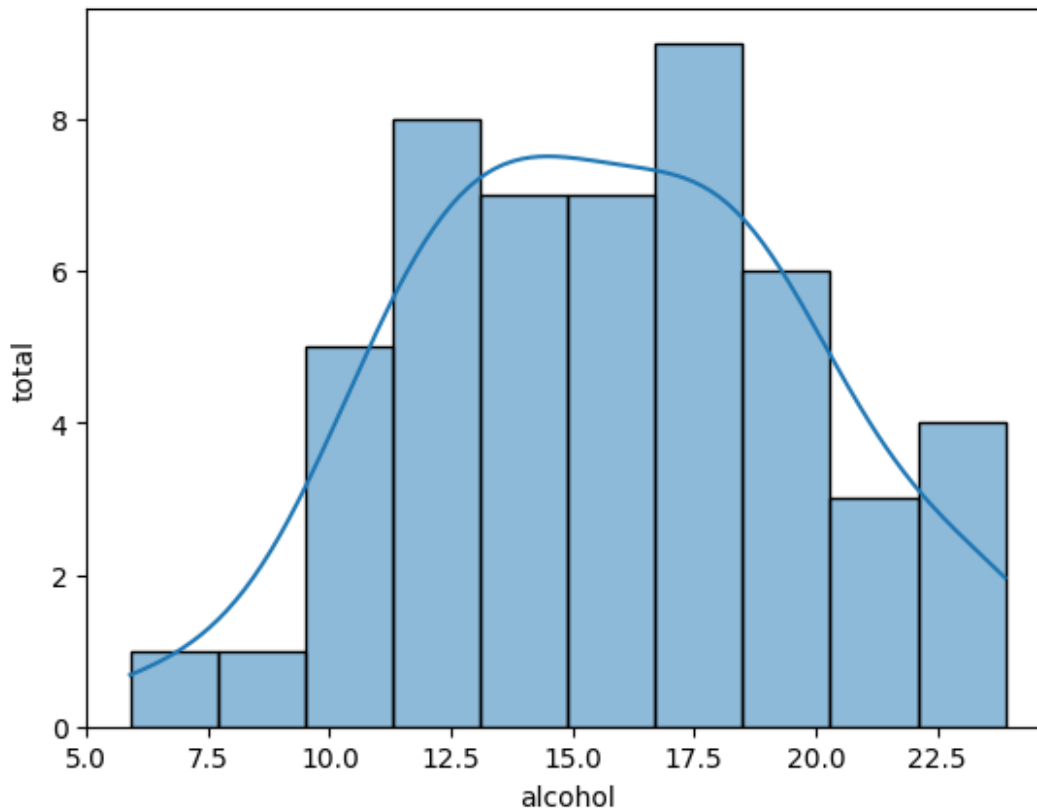
```
#All reasons are almost same with accidents

#Let's plot alcohol vs total accidents.
x_location = "alcohol"
y_total = "total"

sns.histplot(data=car_crashes, x=y_total, bins=10, kde=True)

plt.xlabel(x_location)
plt.ylabel(y_total)
plt.show()
```
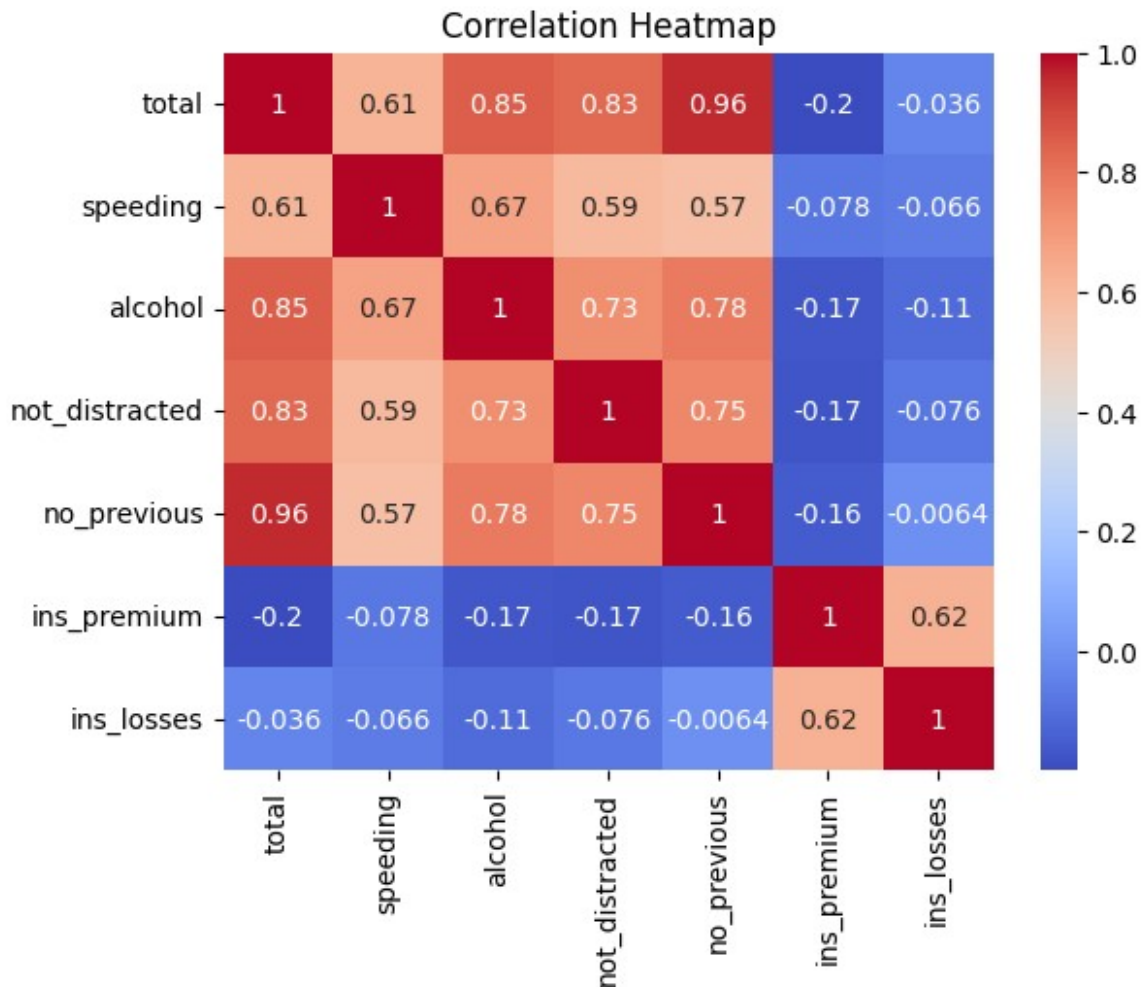
This Histogram shows,

1. that ammount of Acohol is related to total number accidents in bell shaped.
2. between 12.5 to 17.5 we can see maximum.

```python
#Let's plot heatmap

sliced_data = car_crashes.drop(columns=['abbrev'])
sliced_data.head()

corr_matrix = sliced_data.corr()

sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Heatmap')
plt.show()
```

This heat map shows

1. no_previous is highly related to accidents
2. not_distracted is also closely related to total number of accidents
3. drivers with alcohol is 0.85 related to total accidents

```
#Let's see relation between loss and total accidents
sns.boxplot(data=car_crashes, x='ins_losses', y='total')
plt.title('Box Plot of insurence loss by total accidents')
plt.show()
```

Box Plot of insurence loss by total accidents
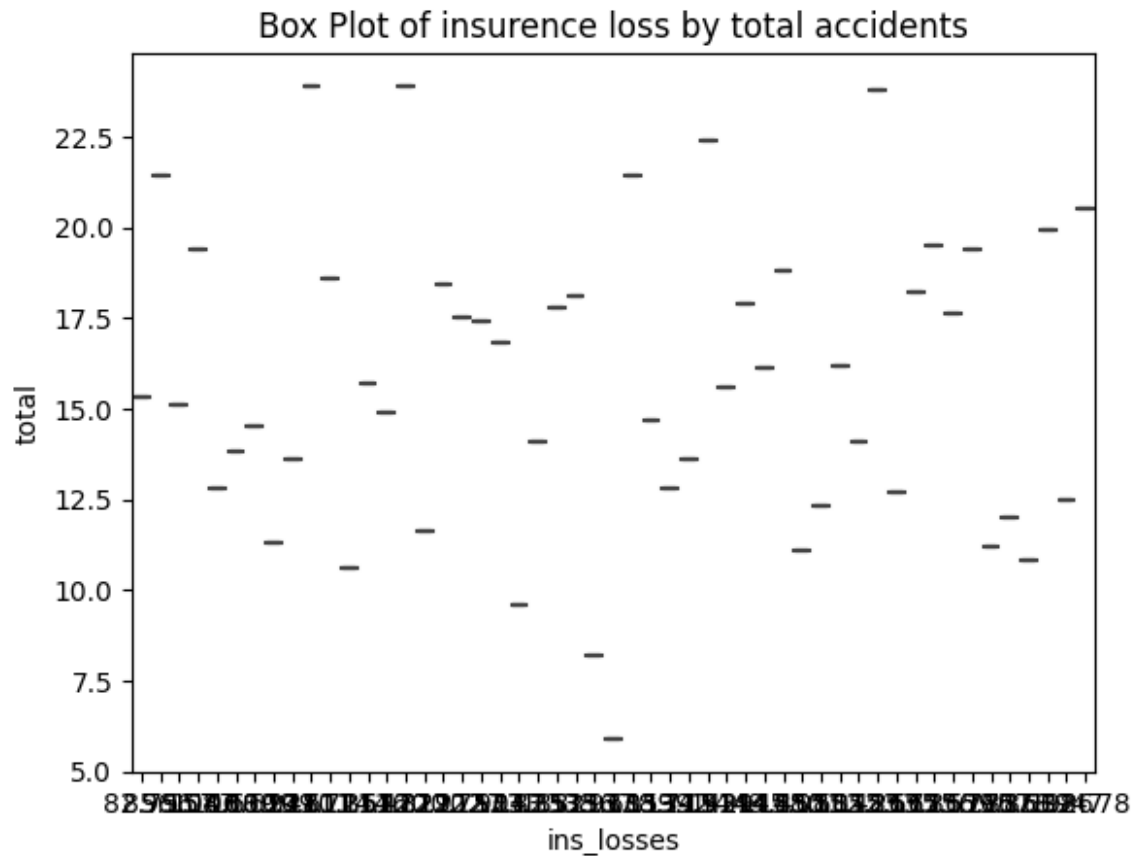
```
# Create a bar plot
sns.barplot(x='speeding', y='total', data=car_crashes)
plt.title('Bar Plot of Speeding Vs Total Accidents')
plt.show()
```

Bar Plot of Speeding Vs Total Accidents
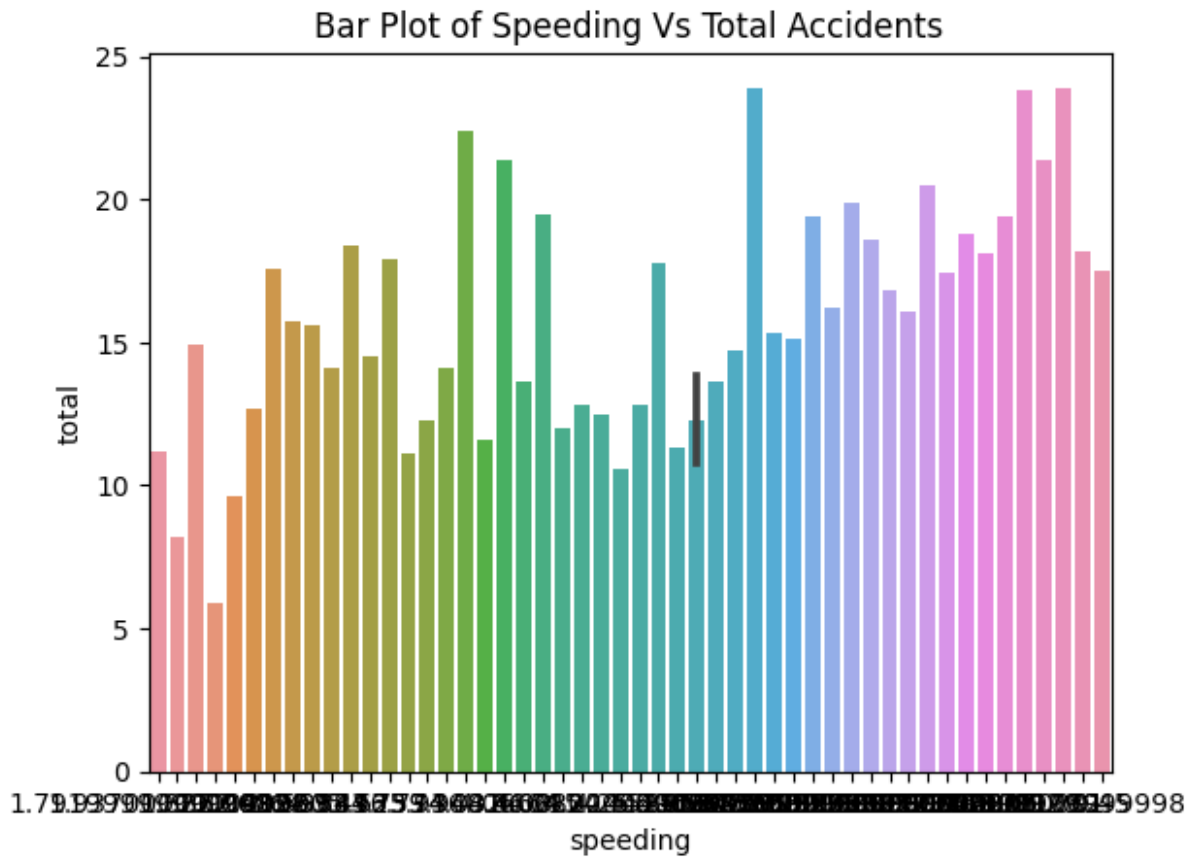
```
#Let's Count plot data
# Create a count plot
sns.countplot(x='total', data=car_crashes)
plt.title('Count Plot of total accident')
plt.show()
```

Count Plot of total accident

This Plot show:

- the counts are almost constant
- not closely related

```
# Create a regression plot for Speeding and accidents
sns.regplot(x='speeding', y='total', data=car_crashes)
plt.title('Regression Plot of Speeding vs. Total Accidents')
plt.show()
```

Regression Plot of Speeding vs. Total Accidents
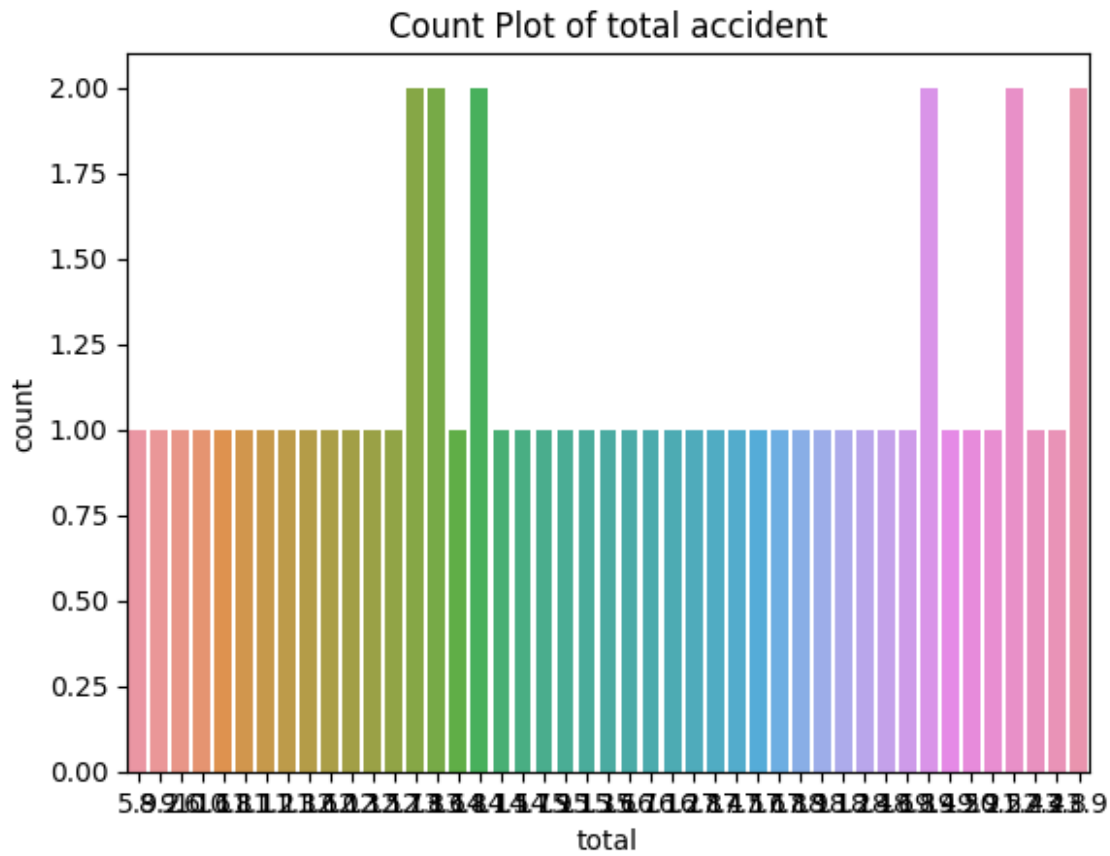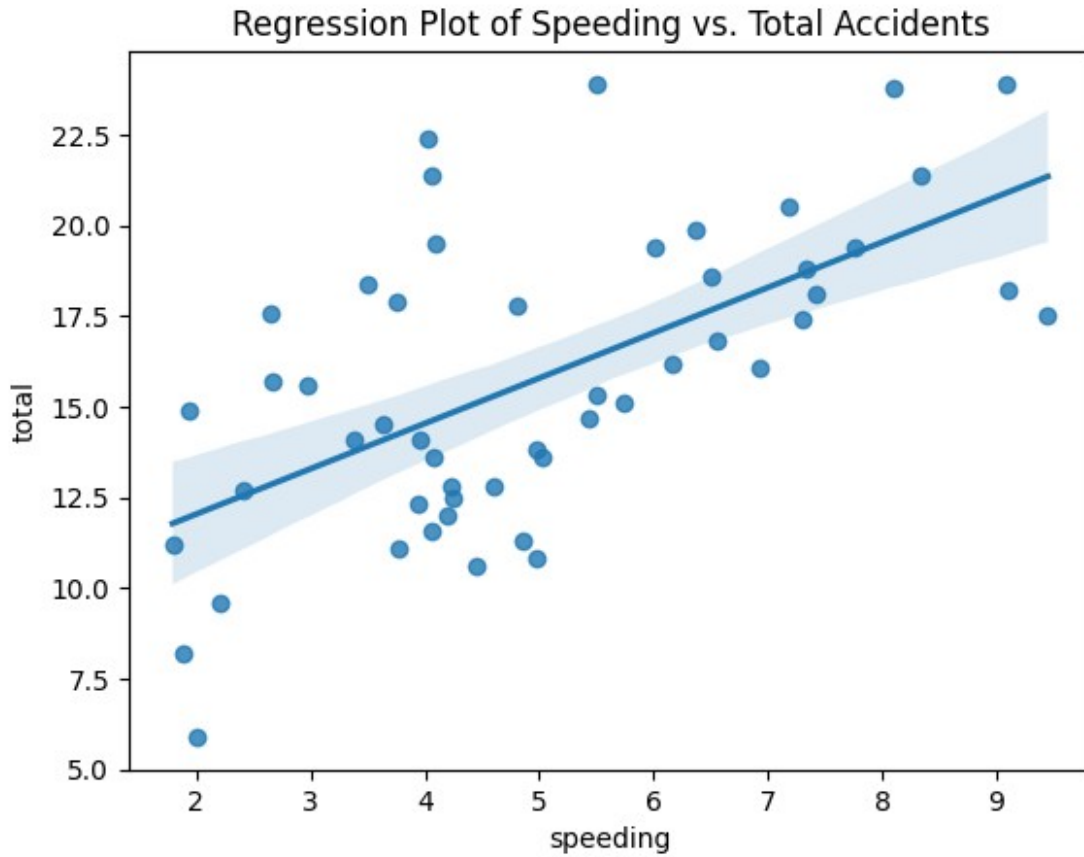
This Regression plot Shows:
1. With increse in speed the accidents also increses.
2. There is less chances of accidents for less speed.