DATA VISUALIZATION

```
import seaborn as sns
sns.get_dataset_names()
dataset=sns.load_dataset('car_crashes')
dataset
```

⇥

19	15.1	5.738	4.530	13.137	12.684	661.88	96.57
20	12.5	4.250	4.000	8.875	12.375	1048.78	192.70
21	8.2	1.886	2.870	7.134	6.560	1011.14	135.63
22	14.1	3.384	3.948	13.395	10.857	1110.61	152.26
23	9.6	2.208	2.784	8.448	8.448	777.18	133.35
24	17.6	2.640	5.456	1.760	17.600	896.07	155.77
25	16.1	6.923	5.474	14.812	13.524	790.32	144.45
26	21.4	8.346	9.416	17.976	18.190	816.21	85.15
27	14.9	1.937	5.215	13.857	13.410	732.28	114.82

1.BARPLOT

| 29 | 11.6 | 4.060 | 3.480 | 10.092 | 9.628 | 746.54 | 120.21 |

```python
import matplotlib.pyplot as plt
plt.figure(figsize=(50,15))
sns.barplot(x=dataset['abbrev'],y=dataset['total'],data=dataset)
```
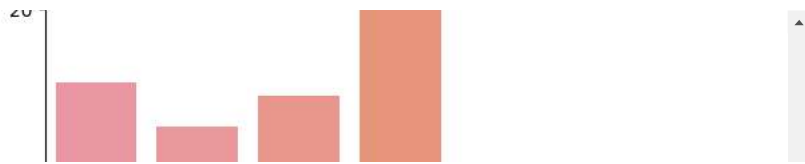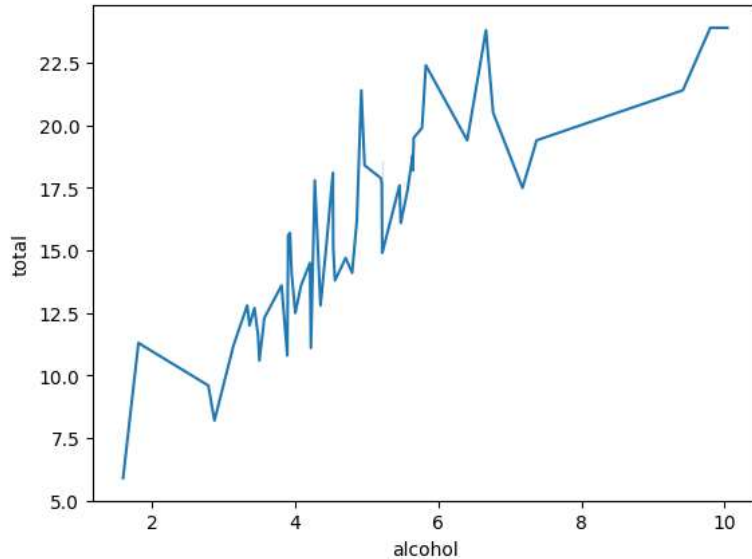
Inference: From the above graph we can see that maximum carcrashes are from city with abbrevation ND

2.LinePlot

```
sns.lineplot(x='alcohol',y='total',data=dataset)
```
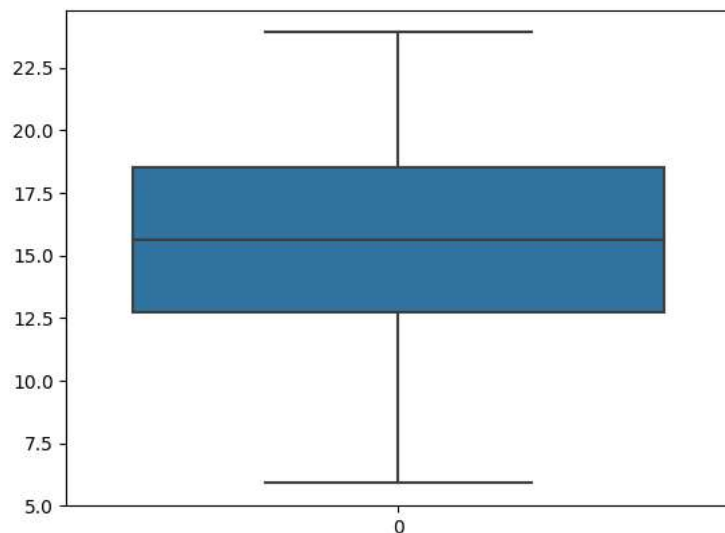
```
<Axes: xlabel='alcohol', ylabel='total'>
```


Inference: From the above graph we can see that in the cities with high total crashes, alcohol is also high

3. BoxPlot

```
sns.boxplot(dataset['total'])
```
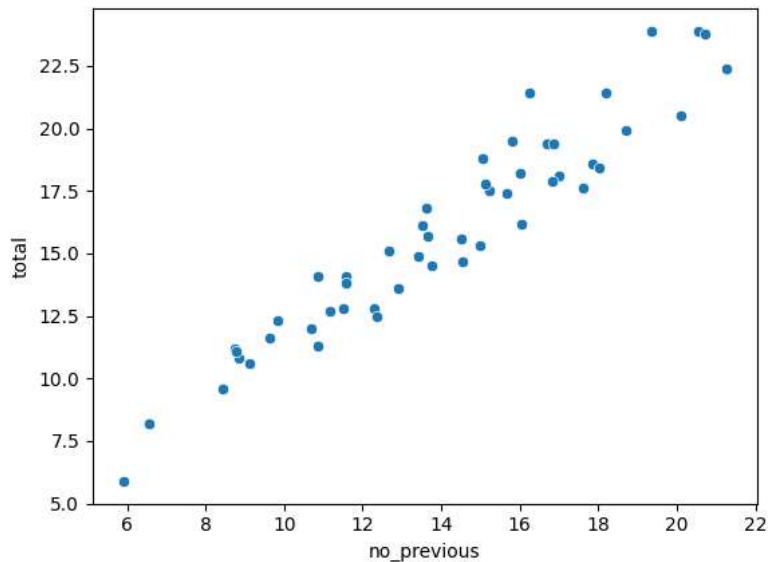
```
<Axes: >
```


Inference: In the above graph there are no outliers, From this we can say that there are no cities with extremely less car crashes as wells as extremely high car crashes

4. Scatter Plot

```
sns.scatterplot(x='no_previous',y='total',data=dataset)
```

```
<Axes: xlabel='no_previous', ylabel='total'>
```
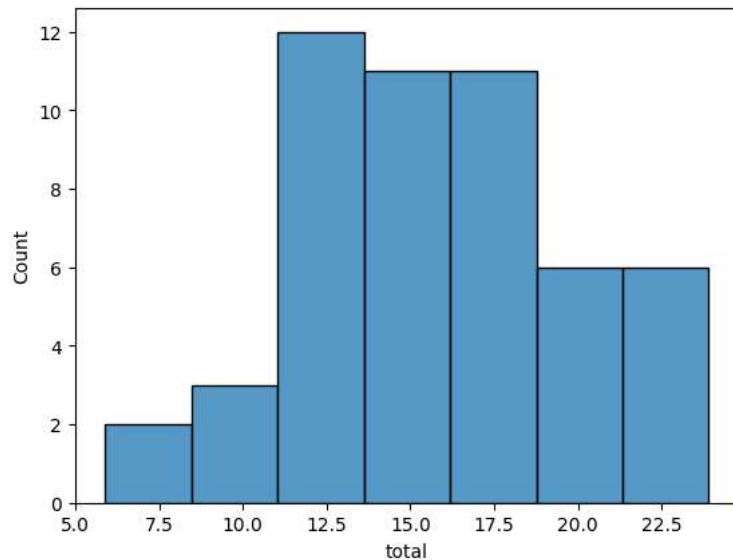

Inference: From the above graph we can say that in every city , car crashes are higher for the first time i.e, in cities with higher car crashes, the crashes with no previuos experience or first time crashes are also high¶

5.Histogram

```
sns.histplot(data=dataset,x='total')
```
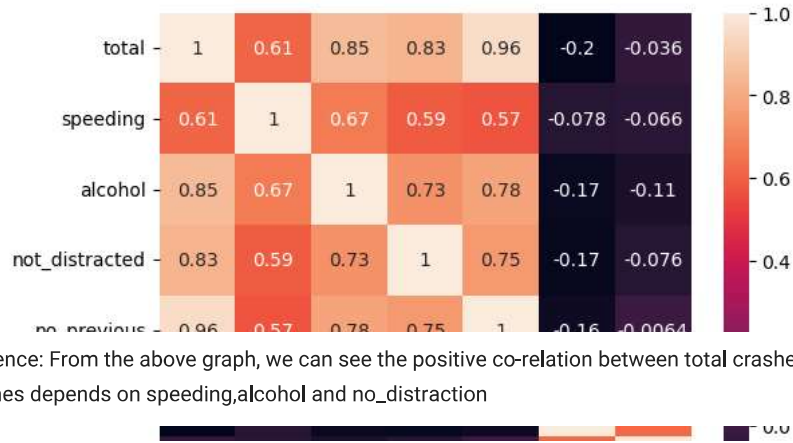
```
<Axes: xlabel='total', ylabel='Count'>
```


Inference: From the above graph we can say that in most of the cities the total car crashes are in range of (10-12)

6. HeatMap

```
correlation_matrix = dataset.corr(numeric_only=True)
sns.heatmap(correlation_matrix,annot=True)
```
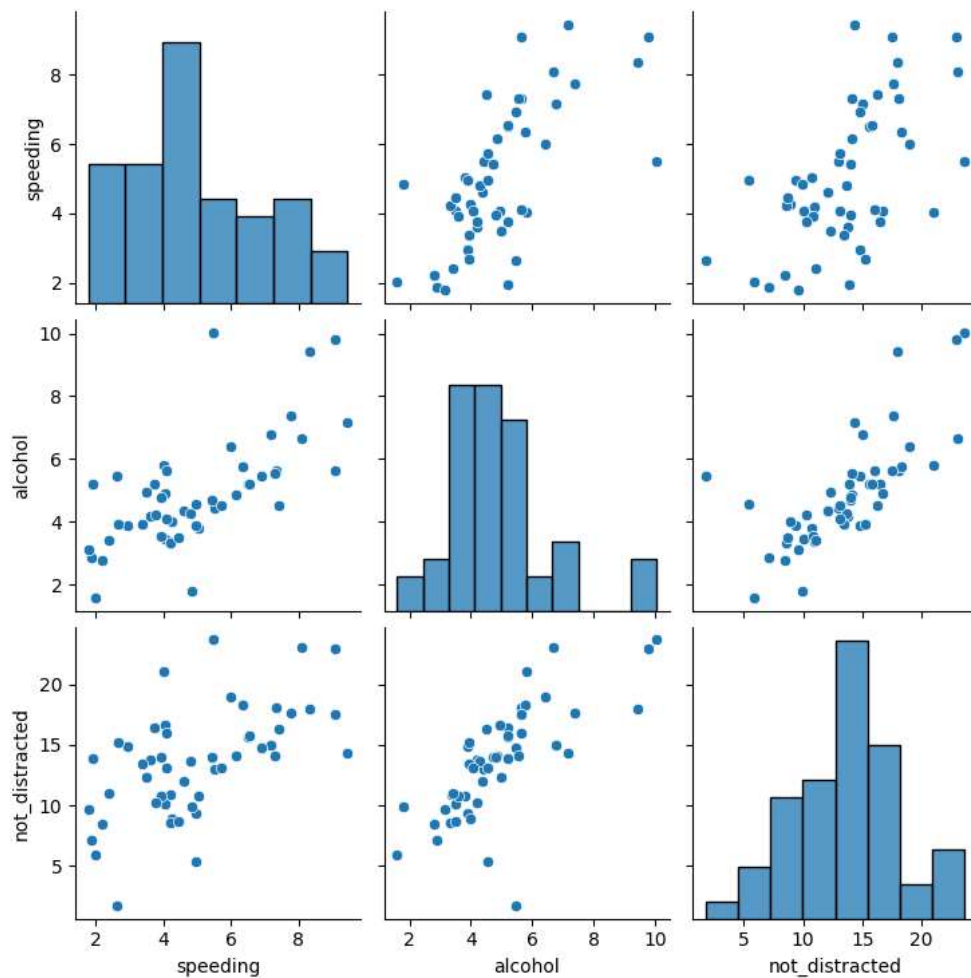
```
<Axes: >
```


Inference: From the above graph, we can see the positive co-relation between total crashes and speeding,alcohol and not_distracted. so total crashes depends on speeding,alcohol and no_distraction

7. Pair Plot

```
sns.pairplot(dataset, vars=['speeding', 'alcohol', 'not_distracted'])
```

```
<seaborn.axisgrid.PairGrid at 0x7acfced53e20>
```


Inference: From the above graphs we can say that all three features alcohol,speeding,not_distracted are proportional to each other
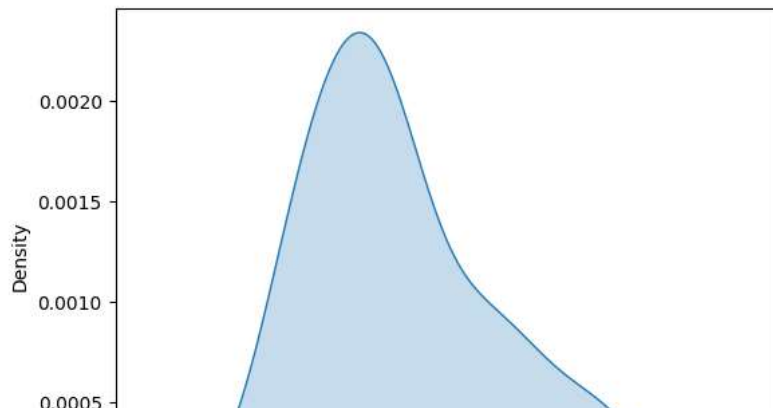
8.Kernel Density Estimation Plot

```
sns.kdeplot(data=dataset['ins_premium'], shade=True)
plt.show()
```

```
<ipython-input-12-70ff29c9e8d0>:1: FutureWarning:

`shade` is now deprecated in favor of `fill`; setting `fill=True`.
This will become an error in seaborn v0.14.0; please update your code.

  sns.kdeplot(data=dataset['ins_premium'], shade=True)
```

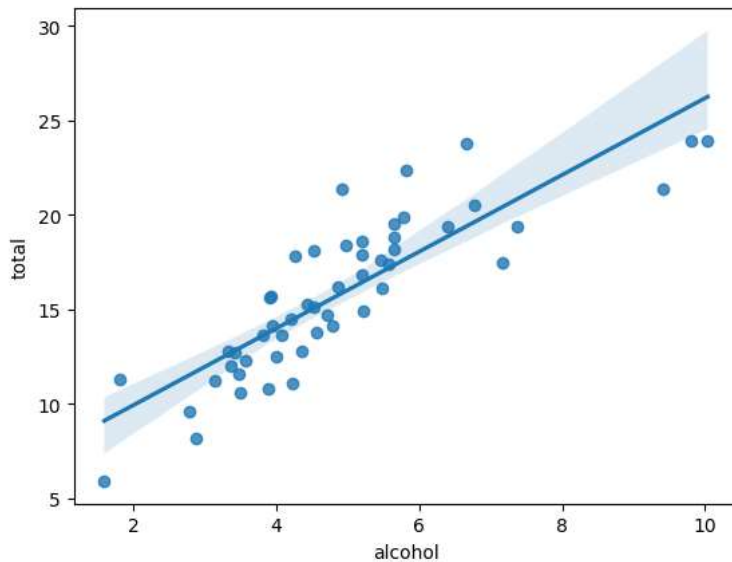Inference: In most of the cities the insurance premium is around 800 and 1000

9.Regression Plot

```
sns.regplot(data=dataset, x='alcohol', y='total')
```

```
<Axes: xlabel='alcohol', ylabel='total'>
```

Inference: From the above graph we can say that there is a linear relationship between alcohol and the total carcrashes

10. Violin Plot

```
plt.figure(figsize=(40,15))
sns.violinplot (x=dataset['ins_premium'], y=dataset['total'],data=dataset)
```