- Name: Prathamesh Mahesh Bhangale
- Email: prathamesh.mahesh2021@vitstudent.ac.in
- Phone_no : +91 9373419818
- Campus: VIT Vellore

```python
import seaborn as sns
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import warnings

# Ignore FutureWarnings
warnings.simplefilter(action='ignore', category=FutureWarning)
```

```python
print(sns.get_dataset_names())
```

    ['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes', 'diamonds', 'dots', 'dowjones', 'exercise', 'flights', 'fmri

```python
df = pd.read_csv('car_crashes.csv')
```

```python
df.info()
```

    <class 'pandas.core.frame.DataFrame'>
    RangeIndex: 51 entries, 0 to 50
    Data columns (total 8 columns):
     #   Column   Non-Null Count Dtype
    --- ------   -------------- -----
     0   total   51 non-null   float64
     1   speeding   51 non-null   float64
     2   alcohol   51 non-null   float64
     3   not_distracted 51 non-null   float64
     4   no_previous   51 non-null   float64
     5   ins_premium   51 non-null   float64
     6   ins_losses   51 non-null   float64
     7   abbrev   51 non-null   object
    dtypes: float64(7), object(1)
    memory usage: 3.3+ KB

▾ Inferences

We have No Null values, and all the columns are Floating Number

```python
df.describe()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_
count	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000	51.
mean	15.790196	4.998196	4.886784	13.573176	14.004882	886.957647	134.
std	4.122002	2.017747	1.729133	4.508977	3.764672	178.296285	24.
min	5.900000	1.792000	1.593000	1.760000	5.900000	641.960000	82.
25%	12.750000	3.766500	3.894000	10.478000	11.348000	768.430000	114.
50%	15.600000	4.608000	4.554000	13.857000	13.775000	858.970000	136.
75%	18.500000	6.439000	5.604000	16.140000	16.755000	1007.945000	151.
max	23.900000	9.450000	10.038000	23.661000	21.280000	1301.520000	194.

```python
df.head()
```

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

Inference

- Since we have all the of the data as Numerical vs Numerical we may use Scatter Plot

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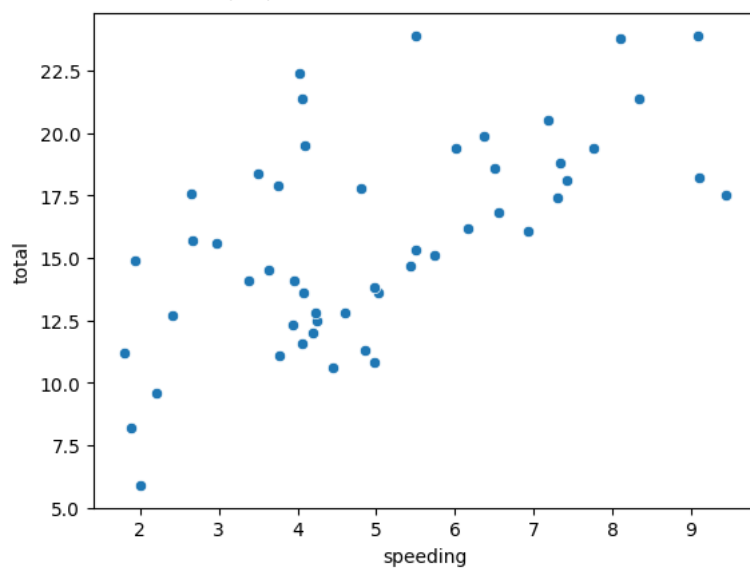
```
<Axes: xlabel='speeding', ylabel='total'>
```
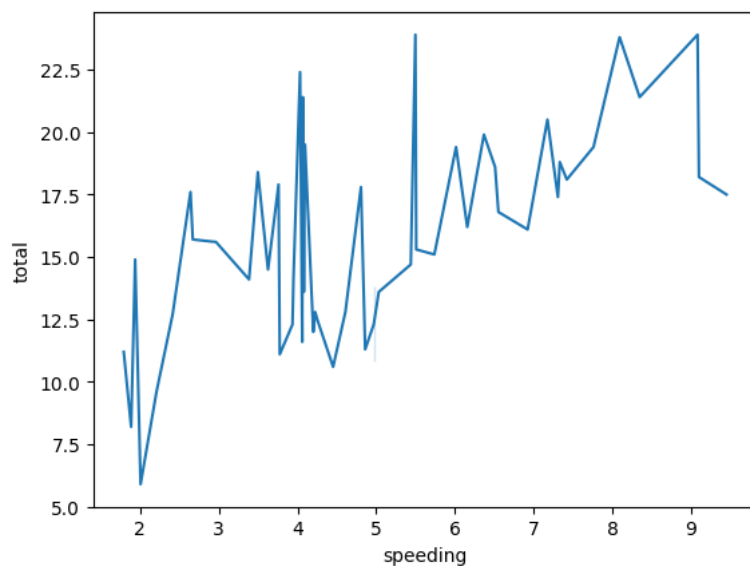


```
sns.scatterplot(y='total',x='speeding',data=df)
```
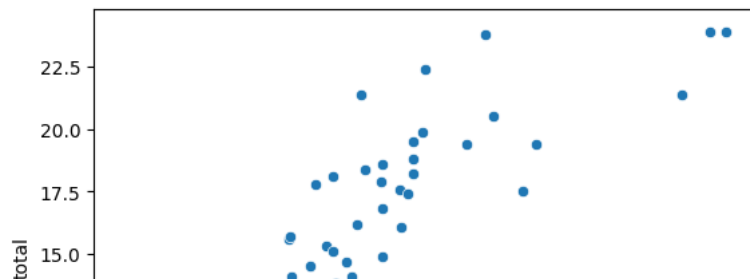
```
<Axes: xlabel='speeding', ylabel='total'>
```


Infrences From Above Plot:

- Here we Can see as the Speeding is increased the total number of Crashes is also increasing
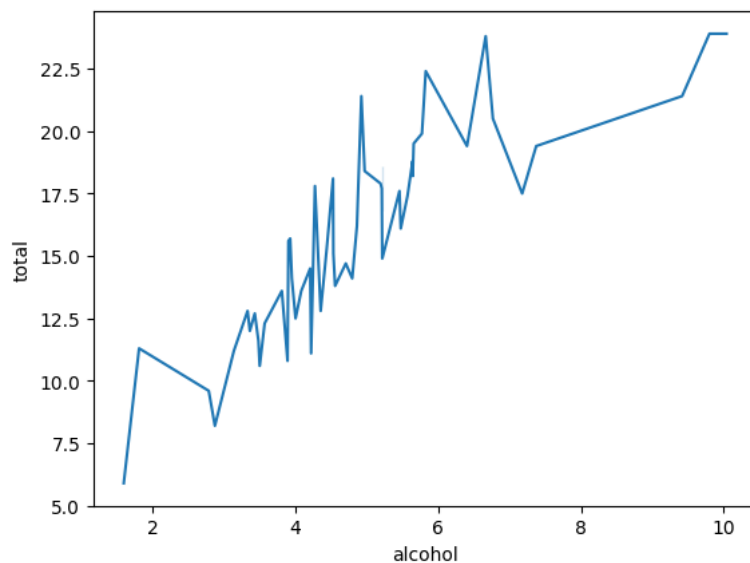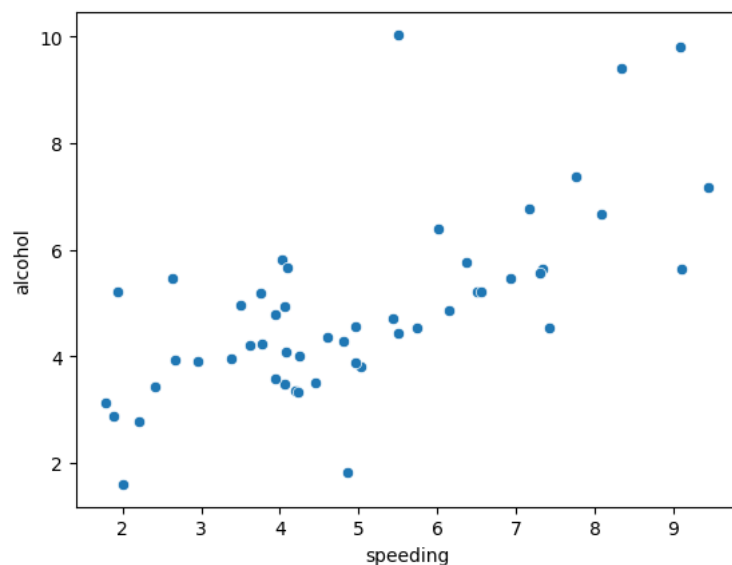
```
sns.scatterplot(y='total',x='alcohol',data=df)
```

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


▾ Infrences From Above Plot:(+ve Correlation)

- Here we Can see as the Speeding is increased the total number of Crashes is also increasing

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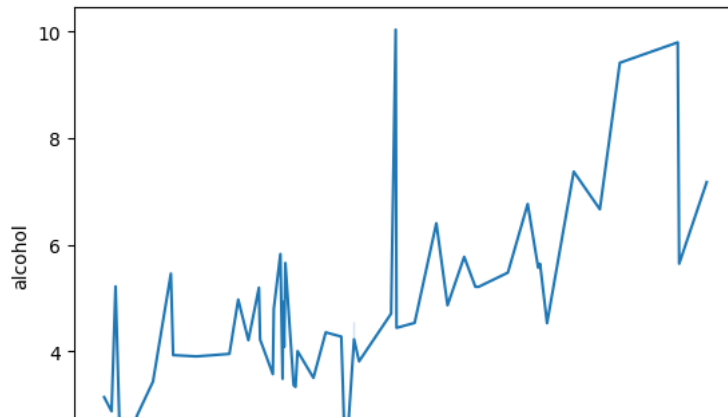
```
<Axes: xlabel='alcohol', ylabel='total'>
```



```
sns.scatterplot(y='alcohol',x='speeding',data=df)
```

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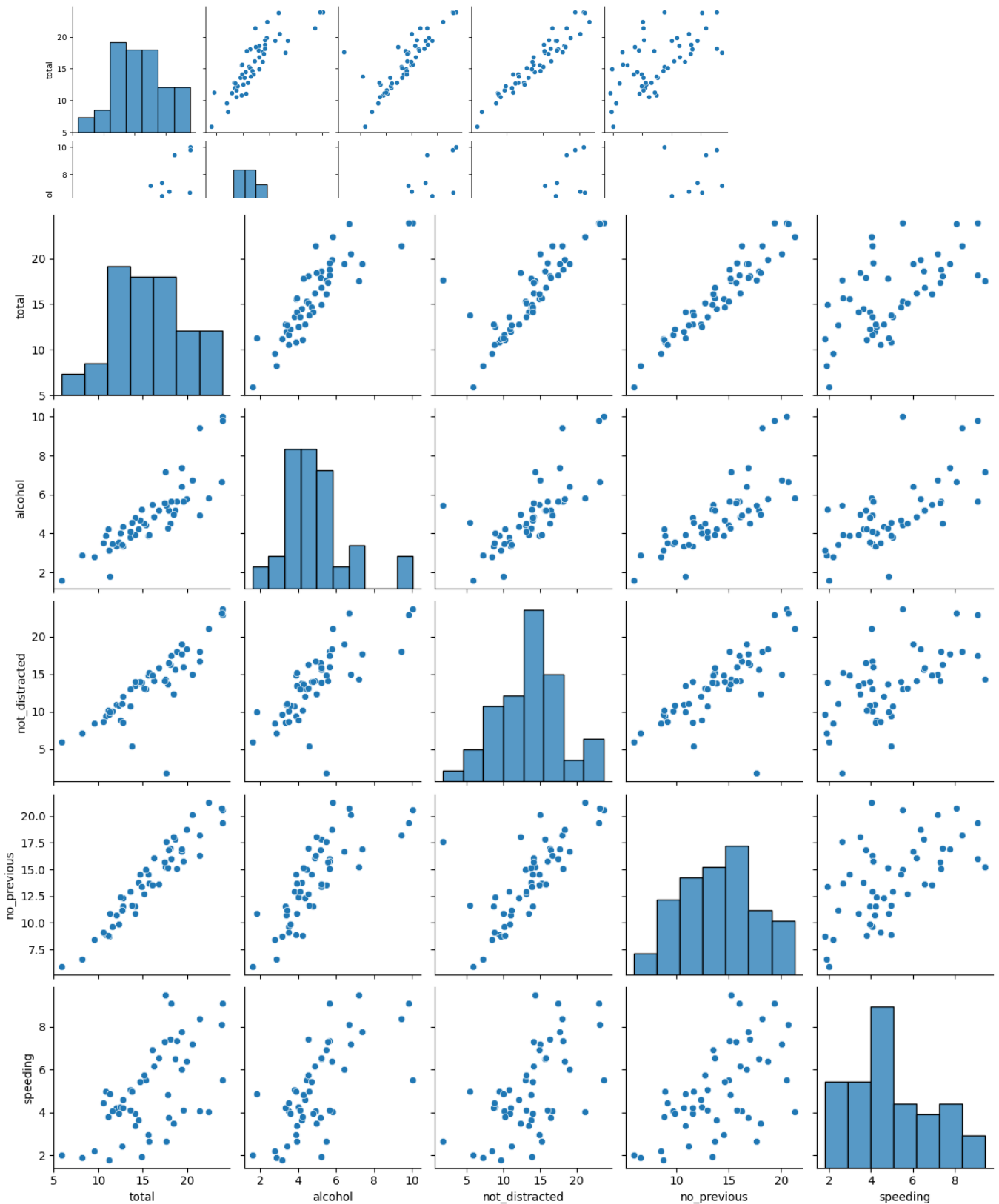


```
sns.lineplot(y='alcohol',x='speeding',data=df)
```

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


▾ Inferences From Above Graphs:

- here we see that generally, as we drink more alchohol more chances of speeding

```
sns.pairplot(df[['total','alcohol','not_distracted','no_previous','speeding']])
plt.show()
```

```
c:\Python311\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The figure layout has c
  self._figure.tight_layout(*args, **kwargs)
```


Inferences

From the Above Plots we can see that:

> Indented block

```
-Realations According to the Total:(Positively related)
    1:- total vs Alcohol
    2:- total vs not_distracted
    3:- total vs no_previous
    This Symbolises that the number of acident increasing as the Y-axis values increases;
```

-Relations According to the Speeding:

    1:- speeding vs alcohol

    This shows that speeding is done by the ones who drank more alcohol.

-Relations According to the Alcohol:

    1:- Alcohol vs no_previous

    2:- Alcholo vs no_distracted

    From the Graphs we can see a relationship that persons who drank the most are less likely to be distracted.

   The persons who drinks more had no_previous crashes

---

-Relations According to the not_distracted:

    1:- not_distracted vs no_previous

        This says that persons who are not distracted , have had no previous accident.

Realtionship between other plot are hard to draw because of does not follow any Good pattern.

```
correlation_matrix = df.corr(numeric_only=True)
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.show()
```
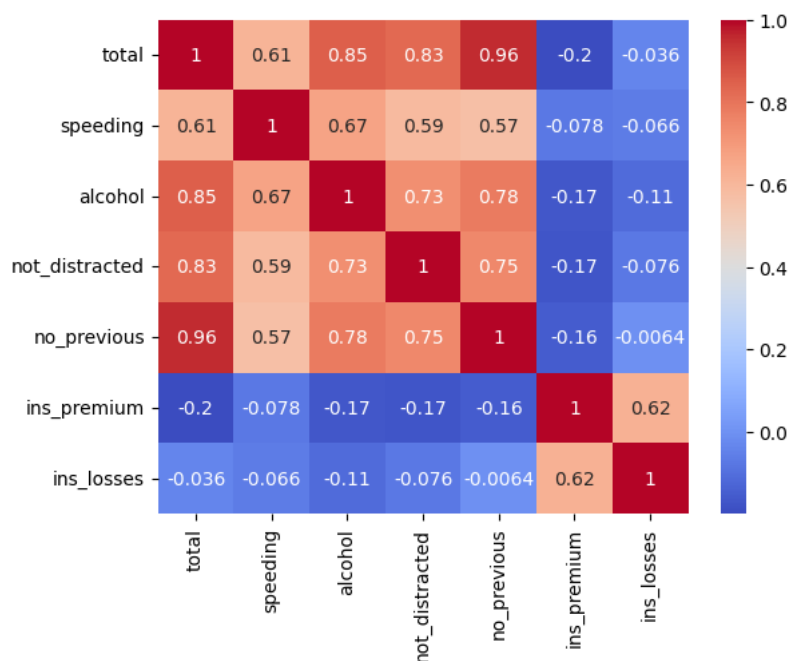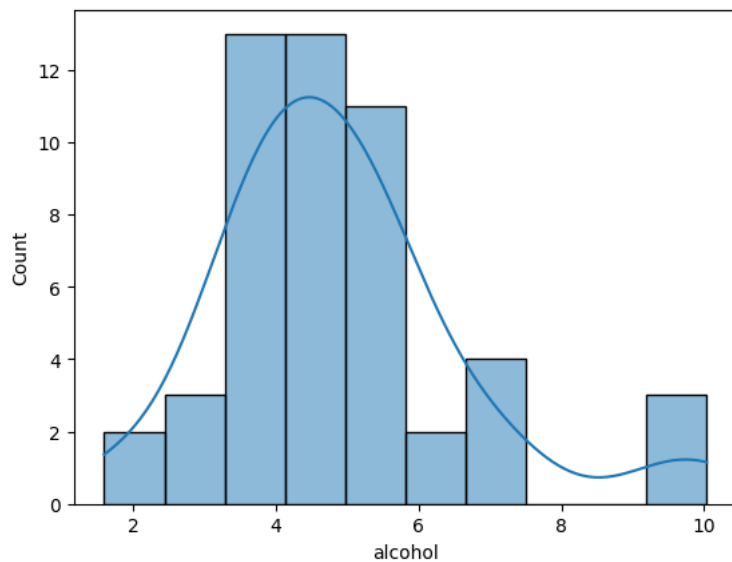


```
sns.histplot(df['speeding'], kde=True)
plt.show()
```

▾ Inference

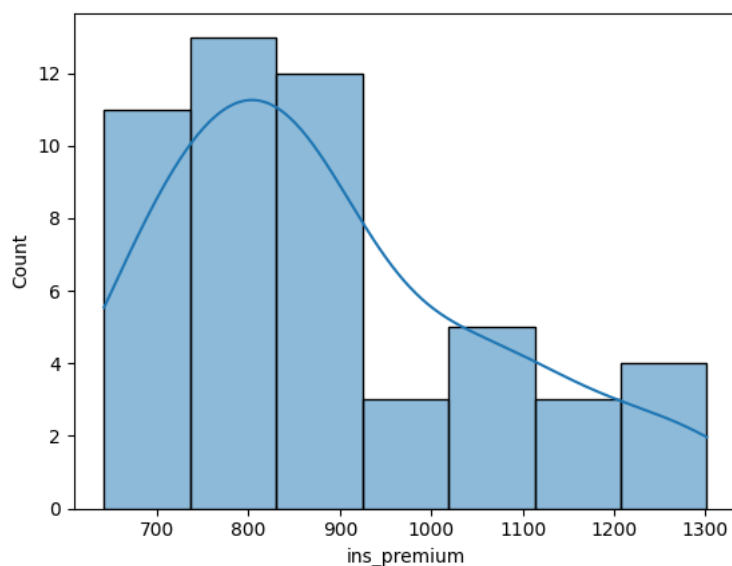- Most of the people drive in range of 2-5 speeding

```
sns.histplot(df['alcohol'], kde=True)
plt.show()
```


▾ inference

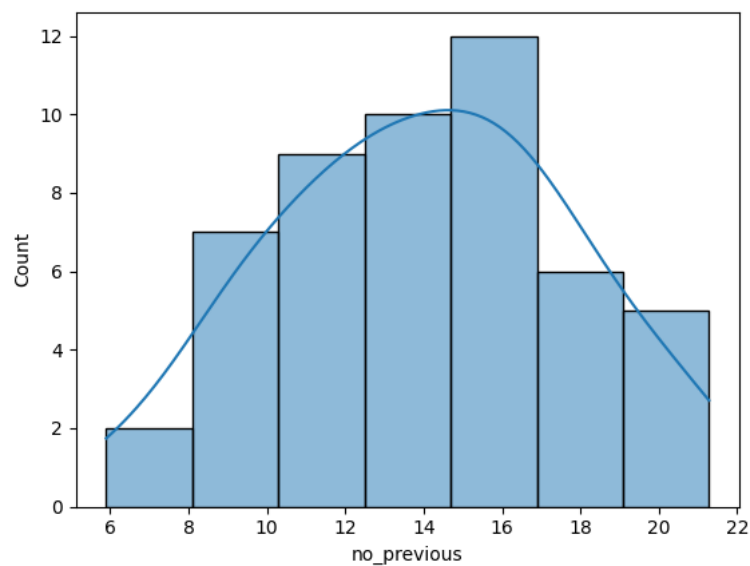- most of the driver consume alchols between 3-6

```
sns.histplot(df['ins_premium'], kde=True)
plt.show()
```


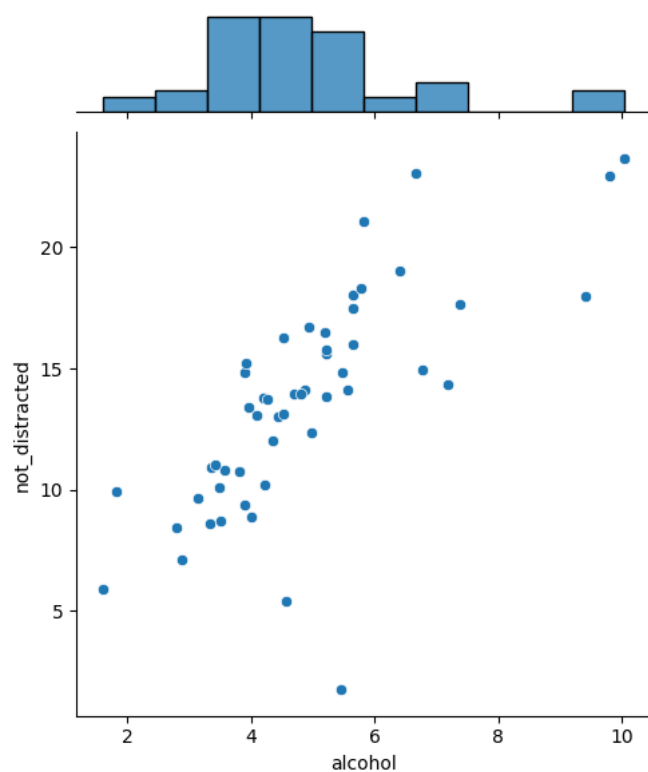## ▾ inference

- most of the driver have insurance from 700-900

```
sns.histplot(df['no_previous'], kde=True)
plt.show()
```
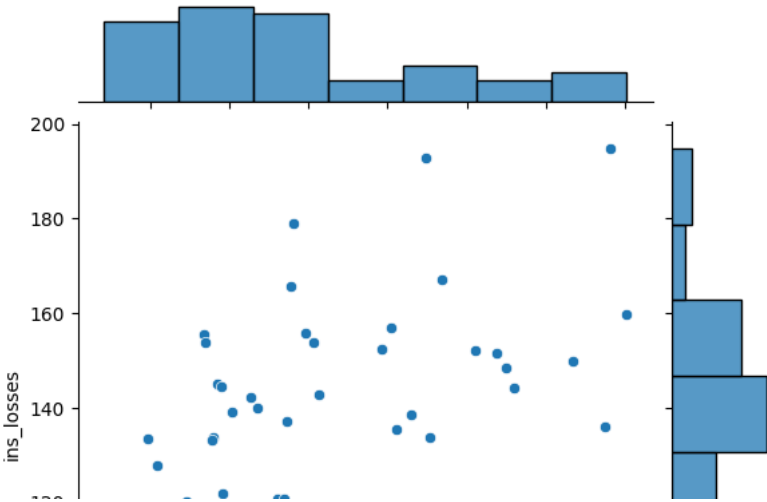
inference

- most of the driver had no_previous accidents

```
sns.jointplot(data=df, x='alcohol', y='not_distracted')
plt.show()
```
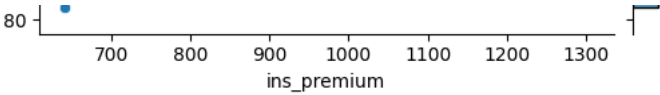


```
sns.jointplot(data=df, x='ins_premium', y='ins_losses')
plt.show()
```

INFERENCE

- Insurances Losses are more for the ones who have more Insurance Premiums.

⚠ 0s   completed at 12:29 PM   ● ✕