

```
In [1]: import seaborn as sn
from matplotlib import pyplot as plt
```

loading the car_crashes dataset which is present in seaborn datasets

```
In [2]: df=sn.load_dataset('car_crashes')
df.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

getting information about the dataset

```
In [3]: 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
```

```
In [4]: fig, axes = plt.subplots(2, 3, figsize=(18, 10))
sn.scatterplot(ax=axes[0, 0], x='total', y='speeding', data=df)
sn.scatterplot(ax=axes[0, 1], x='total', y='alcohol', data=df)
sn.scatterplot(ax=axes[0, 2], x='total', y='not_distracted', data=df)
sn.scatterplot(ax=axes[1, 0], x='total', y='no_previous', data=df)
sn.scatterplot(ax=axes[1, 1], x='total', y='ins_premium', data=df)
sn.scatterplot(ax=axes[1, 2], x='total', y='ins_losses', data=df)
```

```
Out[4]: <AxesSubplot:xlabel='total', ylabel='ins_losses'>
```

inference from above plots we can observe that by increasing speed,alcohol consumption,not_distracted rate and no previous accidents rates will increase the total rate of accidents.

```
In [9]: fig, axes = plt.subplots(2, 3, figsize=(18, 10))
sn.lineplot(ax=axes[0, 0], x="speeding", y="total", data=df)
sn.lineplot(ax=axes[0, 1], x="alcohol", y="total", data=df)
sn.lineplot(ax=axes[0, 2], x="not_distracted", y="total", data=df)
sn.lineplot(ax=axes[1, 0], x="no_previous", y="total", data=df)
sn.lineplot(ax=axes[1, 1], x="ins_premium", y="total", data=df)
sn.lineplot(ax=axes[1, 2], x="ins_losses", y="total", data=df)
```

```
Out[9]: <AxesSubplot:xlabel='ins_losses', ylabel='total'>
```

from the above graphs we can see that line plots of total vs speed,alcohol,not-distracted,no-previous accidents.we can observe that at some points the slope is increasing and at some points it is decreasing.

```
In [16]: fig, axes = plt.subplots(2, 3, figsize=(18, 10))
sn.distplot(df["total"], ax=axes[0, 0])
sn.distplot(df["speeding"], ax=axes[0, 1])
sn.distplot(df["alcohol"], ax=axes[0, 2])
sn.distplot(df["not_distracted"], ax=axes[1, 0])
sn.distplot(df["no_previous"], ax=axes[1, 1])
```

C:\Users\Bhagya Sri\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: 'distplot' is a deprecated function and will be removed in a future version. Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or 'histplot' (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

C:\Users\Bhagya Sri\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: 'distplot' is a deprecated function and will be removed in a future version. Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or 'histplot' (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

C:\Users\Bhagya Sri\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: 'distplot' is a deprecated function and will be removed in a future version. Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or 'histplot' (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

C:\Users\Bhagya Sri\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: 'distplot' is a deprecated function and will be removed in a future version. Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or 'histplot' (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

C:\Users\Bhagya Sri\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: 'distplot' is a deprecated function and will be removed in a future version. Please adapt your code to use either 'displot' (a figure-level function with similar flexibility) or 'histplot' (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

```
Out[16]: <AxesSubplot:xlabel='no_previous', ylabel='Density'>
```

we can see that all the values are normally distributed because all the graphs are in bell shaped curve.

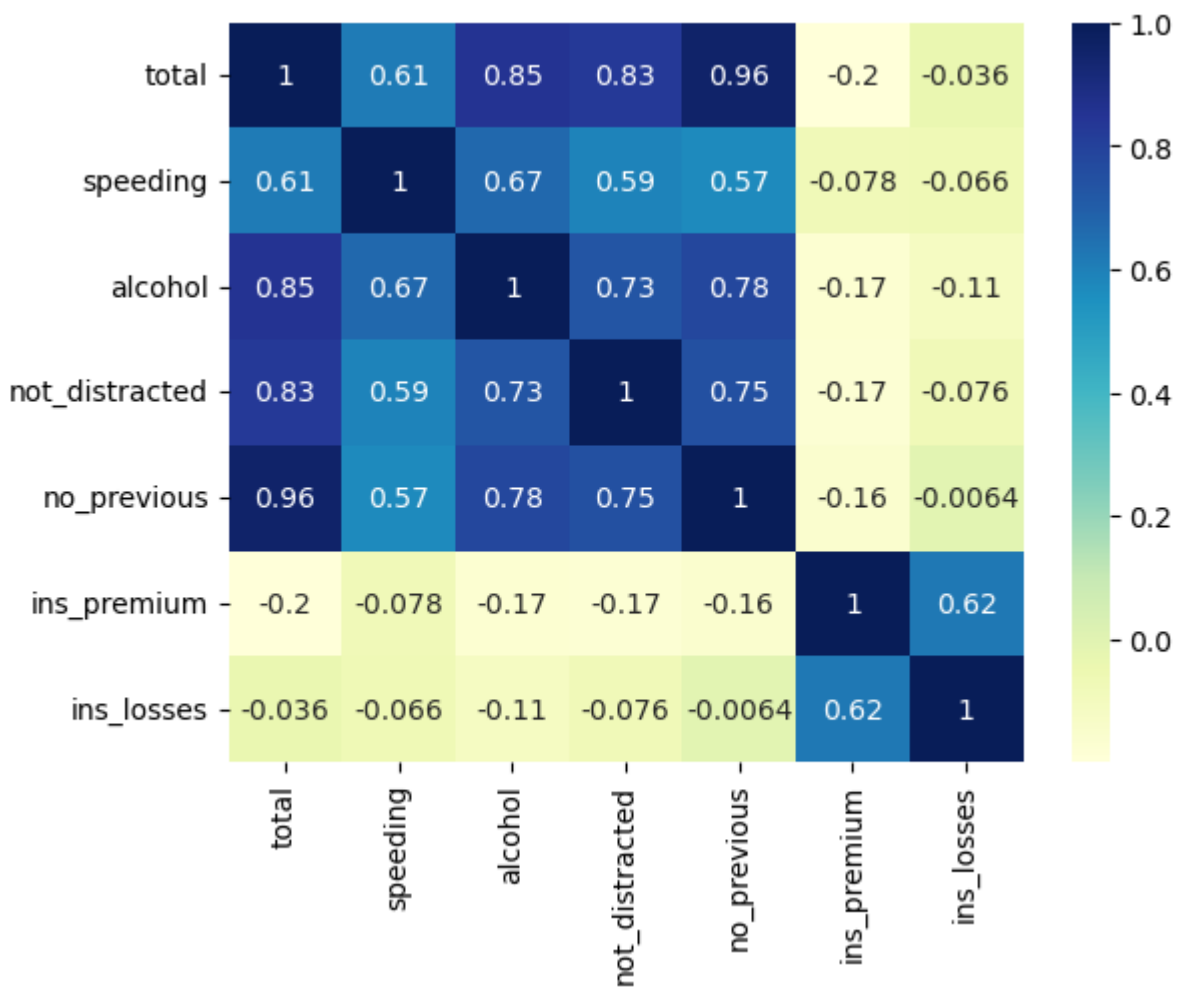
```
In [18]: corr=df.corr()
corr
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses
total	1.000000	0.611548	0.852613	0.827560	0.956179	-0.199702	-0.036011
speeding	0.611548	1.000000	0.669719	0.588010	0.571976	-0.077675	-0.065928
alcohol	0.852613	0.669719	1.000000	0.732816	0.783520	-0.170612	-0.112547
not_distracted	0.827560	0.588010	0.732816	1.000000	0.747307	-0.174856	-0.075970
no_previous	0.956179	0.571976	0.783520	0.747307	1.000000	-0.156895	-0.006359
ins_premium	-0.199702	-0.077675	-0.170612	-0.174856	-0.156895	1.000000	0.623116
ins_losses	-0.036011	-0.065928	-0.112547	-0.075970	-0.006359	0.623116	1.000000

from the above correlation table we can say that total and no previous accidents are strongly related. speeding and alcohol are strongly related. alcohol and total are strongly related. not_distracted and total and no_previous and total are related.

```
In [20]: sn.heatmap(corr, annot=True, cmap="YlGnBu")
```

```
Out[20]: <AxesSubplot:~>
```



```
In [22]: sn.boxplot(x="total", y="speeding", data=df)
```

```
Out[22]: <AxesSubplot:xlabel='total', ylabel='speeding'>
```

