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


sns.get_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']
```

```
dataset=sns.load_dataset("car_crashes")
dataset
```

dataset

16	17.8	4.806	4.272	13.706	15.130	780.45	133.80	
17	21.4	4.066	4.922	16.692	16.264	872.51	137.13	
18	20.5	7.175	6.765	14.965	20.090	1281.55	194.78	
19	15.1	5.738	4.530	13.137	12.684	661.88	96.57	N
20	12.5	4.250	4.000	8.875	12.375	1048.78	192.70	N
21	8.2	1.886	2.870	7.134	6.560	1011.14	135.63	N
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	N
24	17.6	2.640	5.456	1.760	17.600	896.07	155.77	N
25	16.1	6.923	5.474	14.812	13.524	790.32	144.45	N
26	21.4	8.346	9.416	17.976	18.190	816.21	85.15	N
27	14.9	1.937	5.215	13.857	13.410	732.28	114.82	
28	14.7	5.439	4.704	13.965	14.553	1029.87	138.71	
29	11.6	4.060	3.480	10.092	9.628	746.54	120.21	N
30	11.2	1.792	3.136	9.632	8.736	1301.52	159.85	
31	18.4	3.496	4.968	12.328	18.032	869.85	120.75	N
32	12.3	3.936	3.567	10.824	9.840	1234.31	150.01	
33	16.8	6.552	5.208	15.792	13.608	708.24	127.82	
34	23.9	5.497	10.038	23.661	20.554	688.75	109.72	N
35	14.1	3.948	4.794	13.959	11.562	697.73	133.52	

36	19.9	6.368	5.771	18.308	18.706	881.51	178.86	
37	12.8	4.224	3.328	8.576	11.520	804.71	104.61	
38	18.2	9.100	5.642	17.472	16.016	905.99	153.86	
39	11.1	3.774	4.218	10.212	8.769	1148.99	148.58	
40	23.9	9.082	9.799	22.944	19.359	858.97	116.29	
41	19.4	6.014	6.402	19.012	16.684	669.31	96.87	
42	19.5	4.095	5.655	15.990	15.795	767.91	155.57	
43	19.4	7.760	7.372	17.654	16.878	1004.75	156.83	
44	11.3	4.859	1.808	9.944	10.848	809.38	109.48	
45	13.6	4.080	4.080	13.056	12.920	716.20	109.61	
46	12.7	2.413	3.429	11.049	11.176	768.95	153.72	
47	10.6	4.452	3.498	8.692	9.116	890.03	111.62	
48	23.8	8.092	6.664	23.086	20.706	992.61	152.56	
49	13.8	4.968	4.554	5.382	11.592	670.31	106.62	

```
dataset.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

```
dataset.tail()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
46	12.7	2.413	3.429	11.049	11.176	768.95	153.72	VA
47	10.6	4.452	3.498	8.692	9.116	890.03	111.62	WA
48	23.8	8.092	6.664	23.086	20.706	992.61	152.56	WV
49	13.8	4.968	4.554	5.382	11.592	670.31	106.62	WI
50	17.4	7.308	5.568	14.094	15.660	791.14	122.04	WY

```
dataset.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
```

```
dataset.shape
```

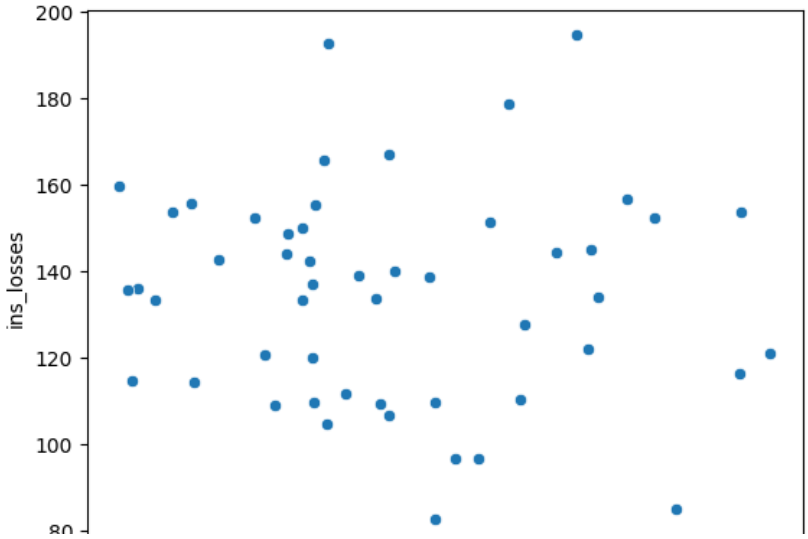
```
(51, 8)
```

```
dataset.describe()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses
	1.000000	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000
	5.790196	4.998196	4.886784	13.573176	14.004882	886.957647	134.493137
	4.122002	2.017747	1.729133	4.508977	3.764672	178.296285	24.835922
	5.900000	1.792000	1.593000	1.760000	5.900000	641.960000	82.750000
	2.750000	3.766500	3.894000	10.478000	11.348000	768.430000	114.645000
	5.600000	4.608000	4.554000	13.857000	13.775000	858.970000	136.050000
	8.500000	6.439000	5.604000	16.140000	16.755000	1007.945000	151.870000
	3.900000	9.450000	10.038000	23.661000	21.280000	1301.520000	194.780000

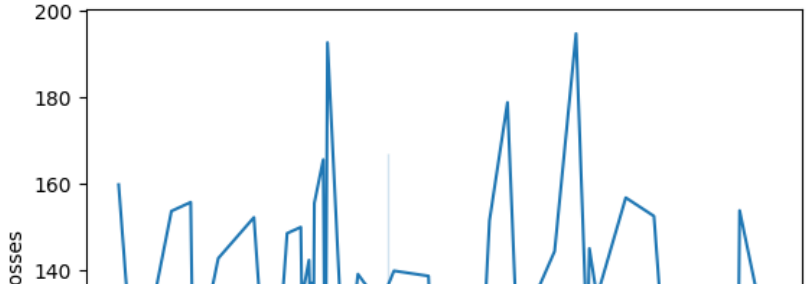
```
sns.scatterplot(x="speeding",y="ins_losses",data=dataset)
```

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




```
sns.lineplot(x="speeding",y="ins_losses",data=dataset)
```

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




```
sns.lineplot(x="speeding",y="ins_losses",data=dataset,ci=None)
```

```
<ipython-input-15-ddb143aea53b>:1: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.lineplot(x="speeding",y="ins_losses",data=dataset,ci=None)
<Axes: xlabel='speeding', ylabel='ins_losses'>
```


```
sns.distplot(dataset["speeding"])
```

```
<ipython-input-17-30d53bfbfda2>:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(dataset["speeding"])
```

```
sns.relplot(x="speeding",y="ins_losses",data=dataset,hue="alcohol")
```

```
<seaborn.axisgrid.FacetGrid at 0x7ae3d8360220>
```


```
dataset["speeding"].value_counts()
```

```
5.439    1
4.060    1
1.792    1
3.496    1
3.936    1
6.552    1
5.497    1
3.948    1
6.368    1
4.224    1
3.774    1
8.346    1
9.082    1
6.014    1
4.095    1
7.760    1
4.859    1
4.080    1
2.413    1
4.452    1
8.092    1
```

```
2.040    1
6.510    1
4.032    1
4.200    1
5.032    1
6.156    1
2.006    1
3.759    1
2.964    1
9.450    1
5.508    1
4.608    1
3.625    1
2.669    1
4.806    1
4.066    1
7.175    1
5.738    1
4.250    1
1.886    1
3.384    1
2.208    1
7.308    1
Name: speeding, dtype: int64
```

```
dataset["ins_losses"].value_counts()
```

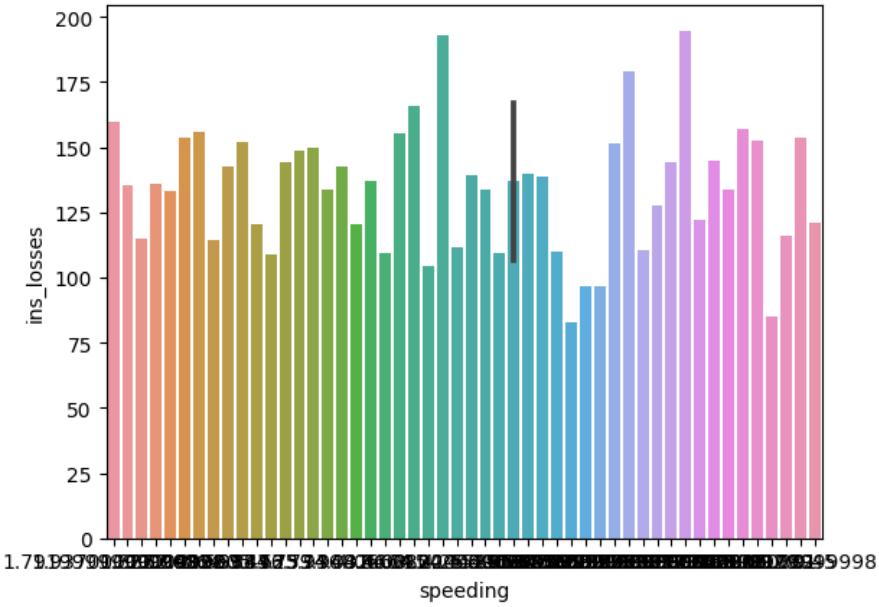
```
159.85    1
120.75    1
```

```
178.80    1
104.61    1
148.58    1
85.15     1
116.29    1
96.87     1
155.57    1
156.83    1
109.48    1
109.61    1
153.72    1
111.62    1
152.56    1
106.62    1
114.82    1
144.45    1
133.93    1
82.75     1
110.35    1
142.39    1
165.63    1
139.91    1
167.02    1
151.48    1
136.05    1
144.18    1
142.80    1
120.92    1
139.15    1
```

```
194.78    1
96.57     1
192.70    1
135.63    1
152.26    1
133.35    1
122.04    1
```

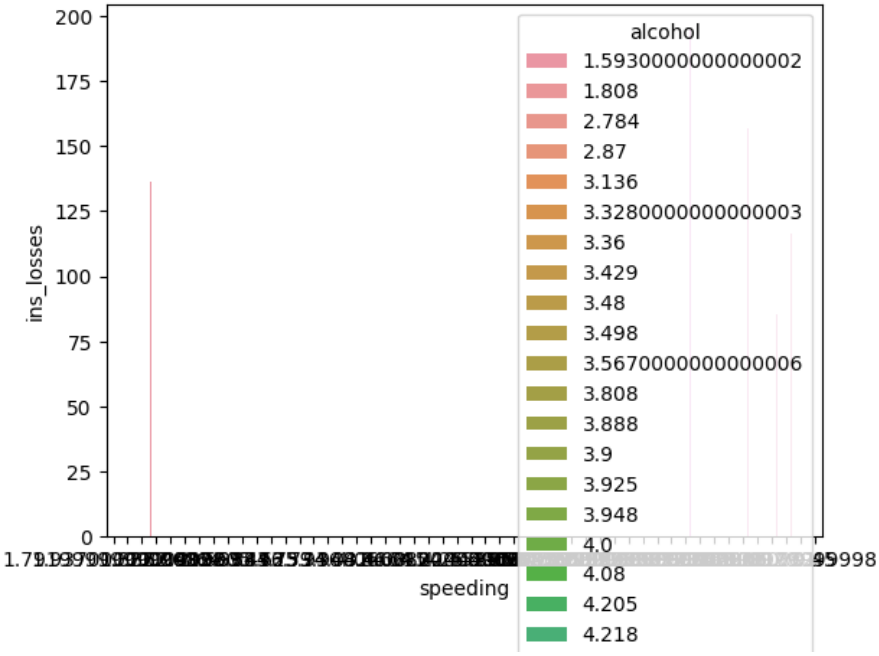
```
sns.barplot(x="speeding",y="ins_losses",data=dataset,)
```

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



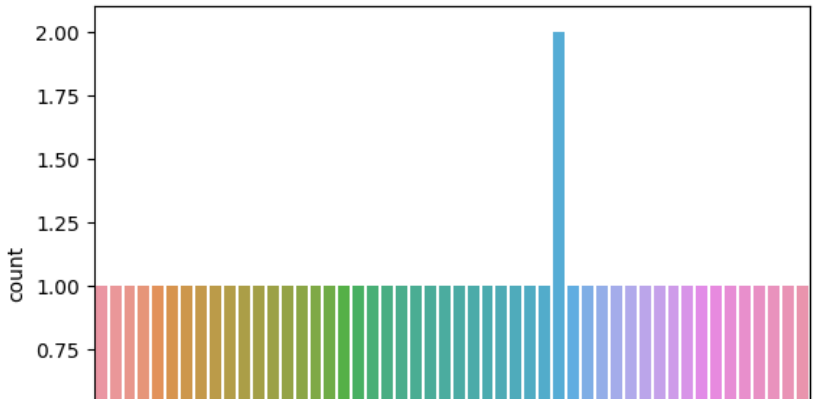
```
sns.barplot(x="speeding",y="ins_losses",data=dataset,hue="alcohol")
```

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



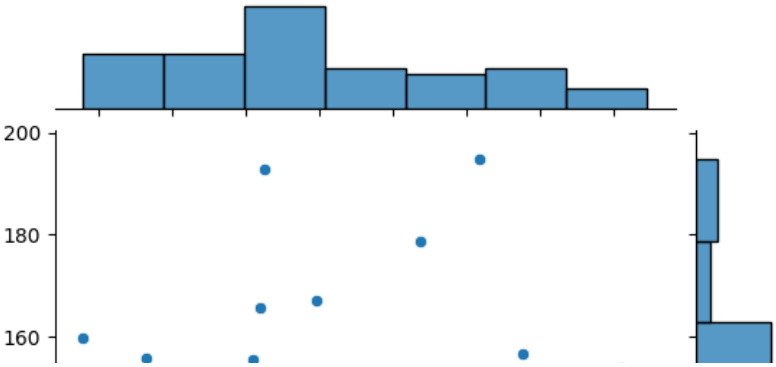

```
sns.countplot(x="alcohol",data=dataset)
```

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




```
sns.jointplot(x="speeding",y="ins_losses",data=dataset)
```

```
<seaborn.axisgrid.JointGrid at 0x7ae3cf54be80>
```




```
sns.boxplot(x="speeding",y="ins_losses",data=dataset)
```

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

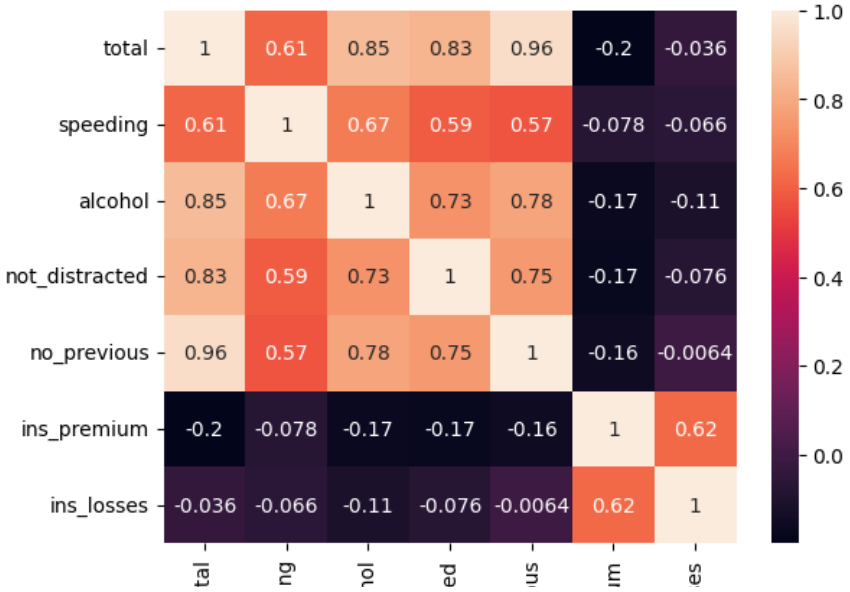

```
corr=dataset.corr()
corr
```

```
<ipython-input-32-f22ca9e9dc13>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is dep
  corr=dataset.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

```
sns.heatmap(corr,annot=True,)
```

```
<Axes: >
```




```
sns.heatmap(corr,annot=True,cmap="YlGnBu")
```