ASSIGNMENT 2

NAME: SUBHANGHI SAHA

REG NO:21BCE7835

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
10	15.6	2.964	3.900	14.820	14.508	913.15	142.80	GA
11	17.5	9.450	7.175	14.350	15.225	861.18	120.92	HI
12	15.3	5.508	4.437	13.005	14.994	641.96	82.75	ID
13	12.8	4.608	4.352	12.032	12.288	803.11	139.15	IL
14	14.5	3.625	4.205	13.775	13.775	710.46	108.92	IN
15	15.7	2.669	3.925	15.229	13.659	649.06	114.47	IA
16	17.8	4.806	4.272	13.706	15.130	780.45	133.80	KS
17	21.4	4.066	4.922	16.692	16.264	872.51	137.13	KY
18	20.5	7.175	6.765	14.965	20.090	1281.55	194.78	LA
19	15.1	5.738	4.530	13.137	12.684	661.88	96.57	ME

In [4]: `car_crash.tail()`

Out[4]:

	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

In [5]: `car_crash.describe().transpose()`

```
In [7]: import numpy as np
        import pandas as pd
        import seaborn as sns
        from matplotlib import pyplot as plt
```

```
In [3]: car_crash= pd.read_csv(r"C:\Users\Subhanghi\Downloads\car_crashes.csv")
        car_crash.head(20)
```

Out[3]:

	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
5	13.6	5.032	3.808	10.744	12.920	835.50	139.91	CO
6	10.8	4.968	3.888	9.396	8.856	1068.73	167.02	CT
7	16.2	6.156	4.860	14.094	16.038	1137.87	151.48	DE
8	5.9	2.006	1.593	5.900	5.900	1273.89	136.05	DC
9	17.9	3.759	5.191	16.468	16.826	1160.13	144.18	FL
10	15.6	2.964	3.900	14.820	14.508	913.15	142.80	GA
11	17.5	9.450	7.175	14.350	15.225	861.18	120.92	HI

| ins_premium | 51.0 | 886.957647 | 178.296285 | 641.960 | 768.4300 | 858.970 | 1007.945 | 1301.520 |
| ins_losses | 51.0 | 134.493137 | 24.835922 | 82.750 | 114.6450 | 136.050 | 151.870 | 194.780 |

```
In [10]: plt.hist(car_crash['alcohol'])
         plt.show()
```
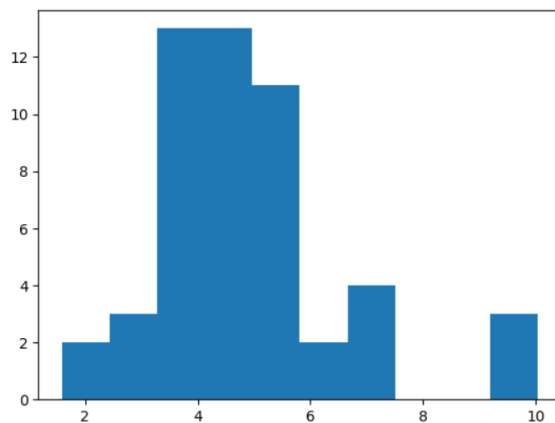

INFERENCE: IT shows the skewness of data. IT also shows outliers or gaps present if any.

```
In [13]: car_crash.boxplot(column='alcohol',by='not_distracted')
```
Out[13]: <Axes: title={'center': 'alcohol'}, xlabel='not_distracted'>

Boxplot grouped by not_distracted
**alcohol**
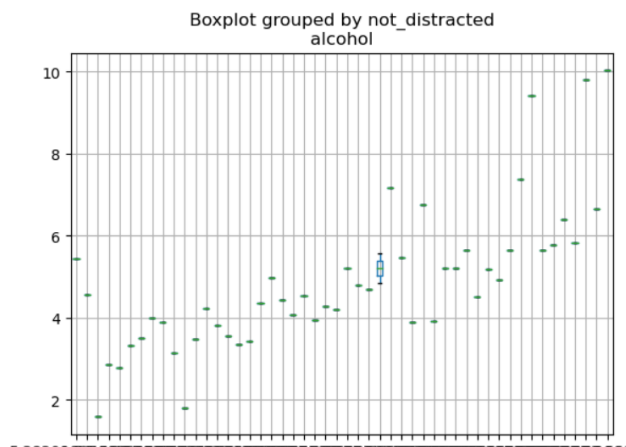
INFERENCE: It shows the average score of the data. It divides the date between 25% median. They show the quartile range.

```
In [20]: speeding=[7.332,7.421,6.51,4.032,4.2,5.032,4.968,6.156,2.006,3.759,2.964,9.45,5.508,4.608,3.625,2.669,4.806,4.066,7.175]
         car_crash.plot.pie(y='speeding',figsize=(5,5),legend=False)
```
Out[20]: <Axes: ylabel='speeding'>
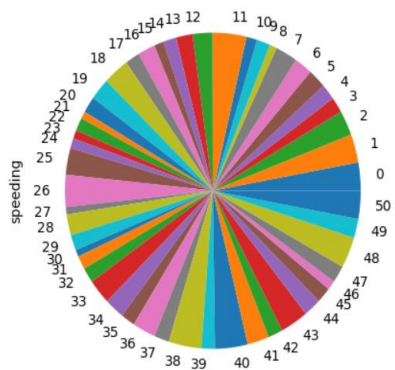
INFERENCE: The graph displays total number of car crash due to various reasons like speeding, alcohol and their proportions.

```
In [23]: car_crash.plot.scatter(x='alcohol',y='speeding')
         plt.grid(True)
```
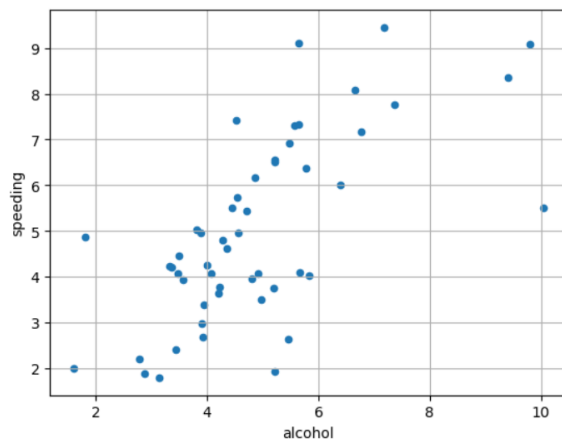

INFERENCE: It explains whether a corelation/relationship exists between vraibles. Here in this case it shows if relationship exits between various car crash type/incidents.