ASSIGNMENT NO 1

Question 7

```
In [8]:   import pandas as pd
          import matplotlib.pyplot as plt
```

```
In [5]:   df=pd.read_csv(r'C:\Users\Vrushabh\Downloads\Q7.csv')
```

```
In [6]:   df
```

Out[6]:

	Unnamed: 0	Points	Score	Weigh
0	Mazda RX4	3.90	2.620	16.46
1	Mazda RX4 Wag	3.90	2.875	17.02
2	Datsun 710	3.85	2.320	18.61
3	Hornet 4 Drive	3.08	3.215	19.44
4	Hornet Sportabout	3.15	3.440	17.02
5	Valiant	2.76	3.460	20.22
6	Duster 360	3.21	3.570	15.84
7	Merc 240D	3.69	3.190	20.00
8	Merc 230	3.92	3.150	22.90
9	Merc 280	3.92	3.440	18.30
10	Merc 280C	3.92	3.440	18.90
11	Merc 450SE	3.07	4.070	17.40
12	Merc 450SL	3.07	3.730	17.60
13	Merc 450SLC	3.07	3.780	18.00

	Unnamed: 0	Points	Score	Weigh
14	Cadillac Fleetwood	2.93	5.250	17.98
15	Lincoln Continental	3.00	5.424	17.82
16	Chrysler Imperial	3.23	5.345	17.42
17	Fiat 128	4.08	2.200	19.47
18	Honda Civic	4.93	1.615	18.52
19	Toyota Corolla	4.22	1.835	19.90
20	Toyota Corona	3.70	2.465	20.01
21	Dodge Challenger	2.76	3.520	16.87
22	AMC Javelin	3.15	3.435	17.30
23	Camaro Z28	3.73	3.840	15.41
24	Pontiac Firebird	3.08	3.845	17.05
25	Fiat X1-9	4.08	1.935	18.90
26	Porsche 914-2	4.43	2.140	16.70
27	Lotus Europa	3.77	1.513	16.90
28	Ford Pantera L	4.22	3.170	14.50
29	Ferrari Dino	3.62	2.770	15.50
30	Maserati Bora	3.54	3.570	14.60
31	Volvo 142E	4.11	2.780	18.60

In [7]: `df.mean()`

Out[7]:
```
Points      3.596563
Score       3.217250
Weigh      17.848750
dtype: float64
```

In [8]: `df.median()`

```
Points     3.695
```

```
Out[8]:   Score      3.325
          Weigh     17.710
          dtype: float64
```

In [12]:
```python
df.var()
```

```
Out[12]:  Points    0.285881
          Score     0.957379
          Weigh     3.193166
          dtype: float64
```

In [14]:
```python
df.std()
```

```
Out[14]:  Points    0.534679
          Score     0.978457
          Weigh     1.786943
          dtype: float64
```

In [15]:
```python
print('Range of Points',df.Points.max()-df.Points.min())
```

```
Range of Points 2.17
```

In [16]:
```python
print('Range of Score',df.Score.max()-df.Score.min())
```

```
Range of Score 3.9109999999999996
```

In [17]:
```python
print('Range of Weigh',df.Weigh.max()-df.Weigh.min())
```

```
Range of Weigh 8.399999999999999
```
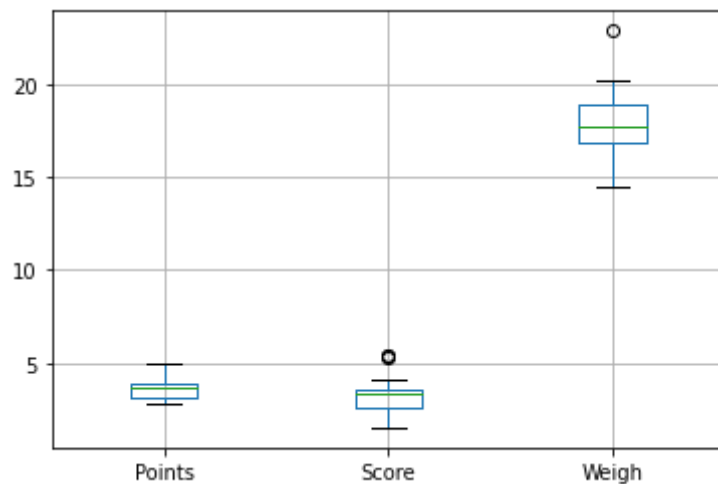
In [18]:
```python
df.mode()
```

Out[18]:

	Unnamed: 0	Points	Score	Weigh
0	AMC Javelin	3.07	3.44	17.02
1	Cadillac Fleetwood	3.92	NaN	18.90
2	Camaro Z28	NaN	NaN	NaN
3	Chrysler Imperial	NaN	NaN	NaN
4	Datsun 710	NaN	NaN	NaN
5	Dodge Challenger	NaN	NaN	NaN

	Unnamed: 0	Points	Score	Weigh
6	Duster 360	NaN	NaN	NaN
7	Ferrari Dino	NaN	NaN	NaN
8	Fiat 128	NaN	NaN	NaN
9	Fiat X1-9	NaN	NaN	NaN
10	Ford Pantera L	NaN	NaN	NaN
11	Honda Civic	NaN	NaN	NaN
12	Hornet 4 Drive	NaN	NaN	NaN
13	Hornet Sportabout	NaN	NaN	NaN
14	Lincoln Continental	NaN	NaN	NaN
15	Lotus Europa	NaN	NaN	NaN
16	Maserati Bora	NaN	NaN	NaN
17	Mazda RX4	NaN	NaN	NaN
18	Mazda RX4 Wag	NaN	NaN	NaN
19	Merc 230	NaN	NaN	NaN
20	Merc 240D	NaN	NaN	NaN
21	Merc 280	NaN	NaN	NaN
22	Merc 280C	NaN	NaN	NaN
23	Merc 450SE	NaN	NaN	NaN
24	Merc 450SL	NaN	NaN	NaN
25	Merc 450SLC	NaN	NaN	NaN
26	Pontiac Firebird	NaN	NaN	NaN
27	Porsche 914-2	NaN	NaN	NaN
28	Toyota Corolla	NaN	NaN	NaN
29	Toyota Corona	NaN	NaN	NaN
30	Valiant	NaN	NaN	NaN

	Unnamed: 0	Points	Score	Weigh
31	Volvo 142E	NaN	NaN	NaN

In [19]: `df.boxplot()`

Out[19]: `<AxesSubplot:>`

In [20]: `df.hist()`

Out[20]:
```
array([[<AxesSubplot:title={'center':'Points'}>,
        <AxesSubplot:title={'center':'Score'}>],
       [<AxesSubplot:title={'center':'Weigh'}>, <AxesSubplot:>]],
      dtype=object)
```
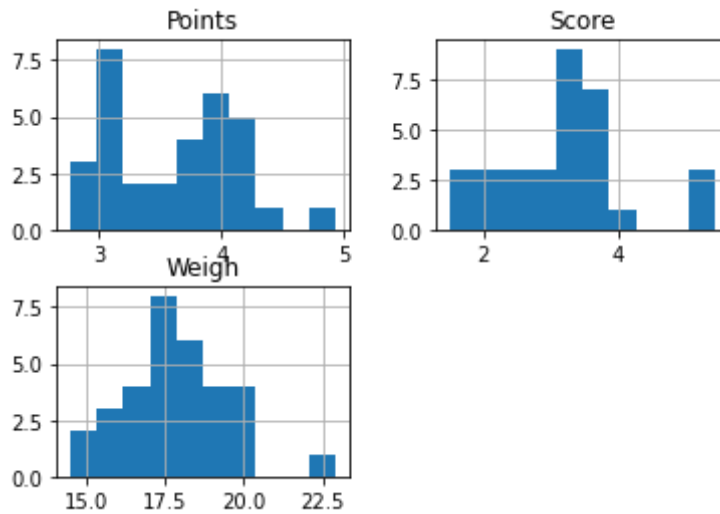
In [23]: `df.mode()`

Out[23]:

	Unnamed: 0	Points	Score	Weigh
0	AMC Javelin	3.07	3.44	17.02
1	Cadillac Fleetwood	3.92	NaN	18.90
2	Camaro Z28	NaN	NaN	NaN
3	Chrysler Imperial	NaN	NaN	NaN
4	Datsun 710	NaN	NaN	NaN
5	Dodge Challenger	NaN	NaN	NaN
6	Duster 360	NaN	NaN	NaN
7	Ferrari Dino	NaN	NaN	NaN
8	Fiat 128	NaN	NaN	NaN
9	Fiat X1-9	NaN	NaN	NaN
10	Ford Pantera L	NaN	NaN	NaN
11	Honda Civic	NaN	NaN	NaN
12	Hornet 4 Drive	NaN	NaN	NaN

	Unnamed: 0	Points	Score	Weigh
13	Hornet Sportabout	NaN	NaN	NaN
14	Lincoln Continental	NaN	NaN	NaN
15	Lotus Europa	NaN	NaN	NaN
16	Maserati Bora	NaN	NaN	NaN
17	Mazda RX4	NaN	NaN	NaN
18	Mazda RX4 Wag	NaN	NaN	NaN
19	Merc 230	NaN	NaN	NaN
20	Merc 240D	NaN	NaN	NaN
21	Merc 280	NaN	NaN	NaN
22	Merc 280C	NaN	NaN	NaN
23	Merc 450SE	NaN	NaN	NaN
24	Merc 450SL	NaN	NaN	NaN
25	Merc 450SLC	NaN	NaN	NaN
26	Pontiac Firebird	NaN	NaN	NaN
27	Porsche 914-2	NaN	NaN	NaN
28	Toyota Corolla	NaN	NaN	NaN
29	Toyota Corona	NaN	NaN	NaN
30	Valiant	NaN	NaN	NaN
31	Volvo 142E	NaN	NaN	NaN

In [ ]:

Question 9

Que9 a

```
In [79]:  df1=pd.read_csv(r'C:\Users\Vrushabh\Downloads\Q9_a.csv')
```

```
In [80]:  df1
```

Out[80]:

	Index	speed	dist
0	1	4	2
1	2	4	10
2	3	7	4
3	4	7	22
4	5	8	16
5	6	9	10
6	7	10	18
7	8	10	26
8	9	10	34
9	10	11	17
10	11	11	28
11	12	12	14
12	13	12	20
13	14	12	24
14	15	12	28
15	16	13	26
16	17	13	34
17	18	13	34
18	19	13	46
19	20	14	26
20	21	14	36
21	22	14	60

	Index	speed	dist
22	23	14	80
23	24	15	20
24	25	15	26
25	26	15	54
26	27	16	32
27	28	16	40
28	29	17	32
29	30	17	40
30	31	17	50
31	32	18	42
32	33	18	56
33	34	18	76
34	35	18	84
35	36	19	36
36	37	19	46
37	38	19	68
38	39	20	32
39	40	20	48
40	41	20	52
41	42	20	56
42	43	20	64
43	44	22	66
44	45	23	54
45	46	24	70
46	47	24	92

Index	speed	dist
47	48	24	93
48	49	24	120
49	50	25	85

In [81]:
```python
df1.skew()
```

Out[81]:
```
Index     0.000000
speed    -0.117510
dist      0.806895
dtype: float64
```

In [82]:
```python
df1.kurtosis()
```

Out[82]:
```
Index    -1.200000
speed    -0.508994
dist      0.405053
dtype: float64
```

In [83]:
```python
df1.plot(x="speed", y="dist")
```

Out[83]:
```
<AxesSubplot:xlabel='speed'>
```
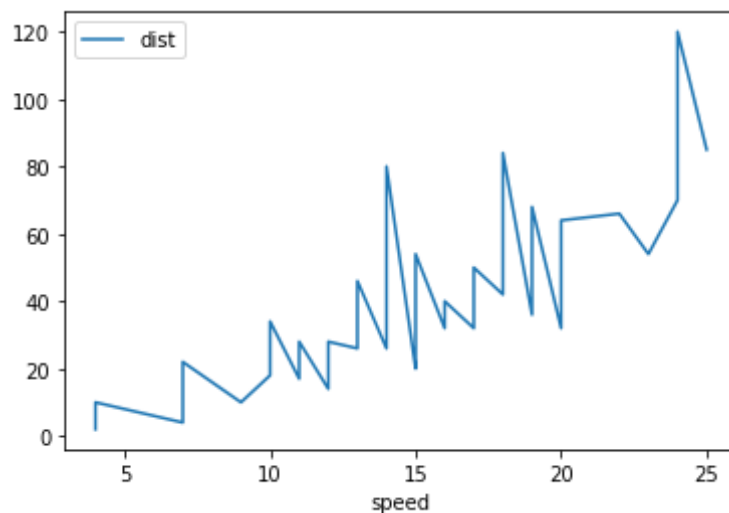
In [84]: `df1.plot.bar(x="speed", y="dist")`

Out[84]: `<AxesSubplot:xlabel='speed'>`
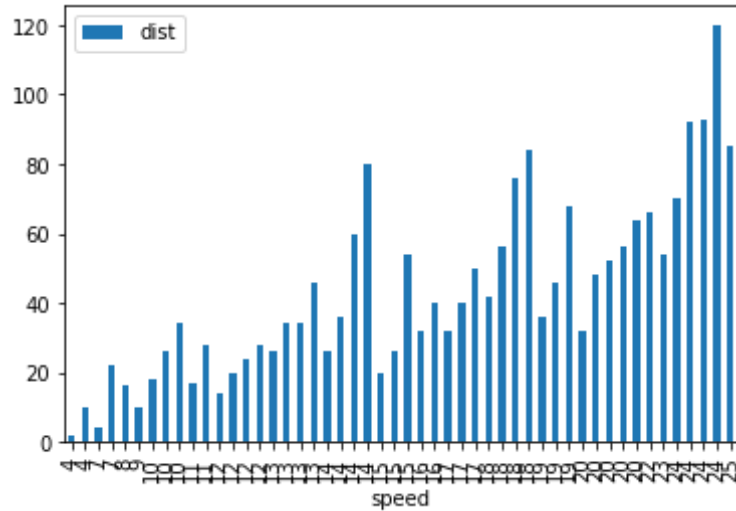
In [85]: `df1.plot.scatter(x="speed", y="dist")`

Out[85]: `<AxesSubplot:xlabel='speed', ylabel='dist'>`
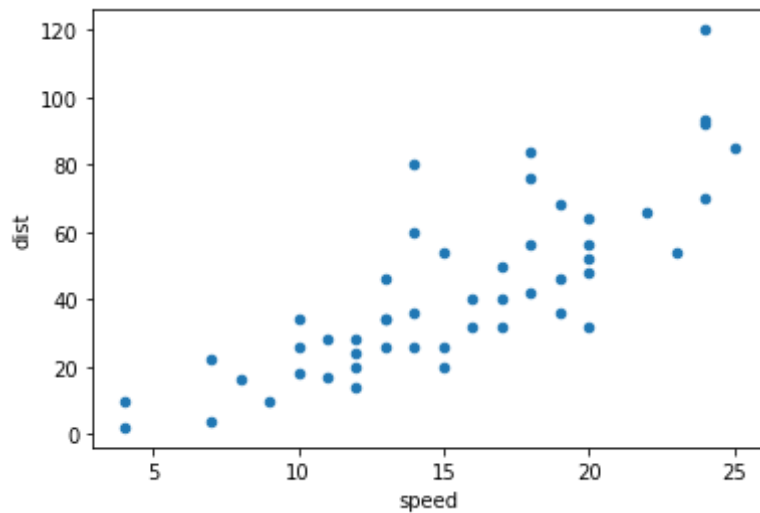
In [86]: `df1.plot.density(x="speed", y="dist")`

Out[86]: `<AxesSubplot:ylabel='Density'>`
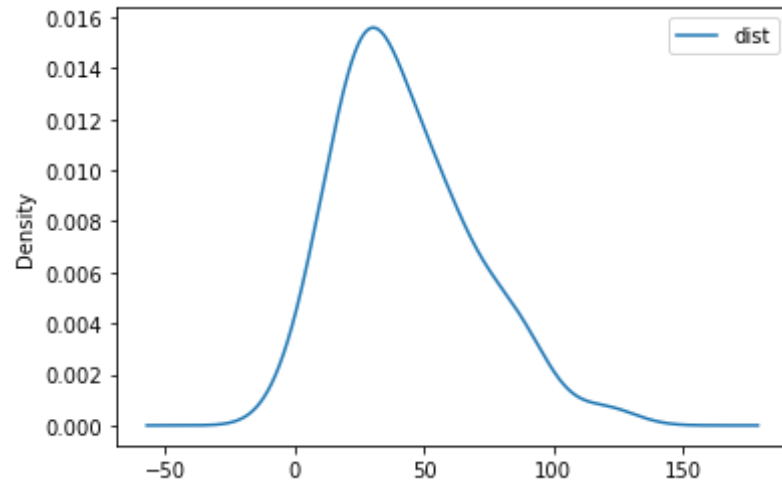
In [87]: `df1.plot.line(x="speed", y="dist")`

Out[87]: `<AxesSubplot:xlabel='speed'>`
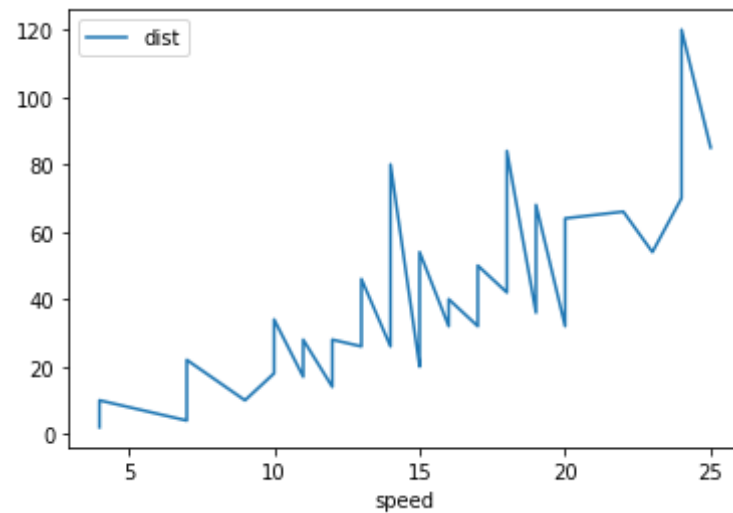
```
In [88]:  df1.plot.hexbin(x="speed", y="dist")
```

Out[88]: <AxesSubplot:xlabel='speed', ylabel='dist'>
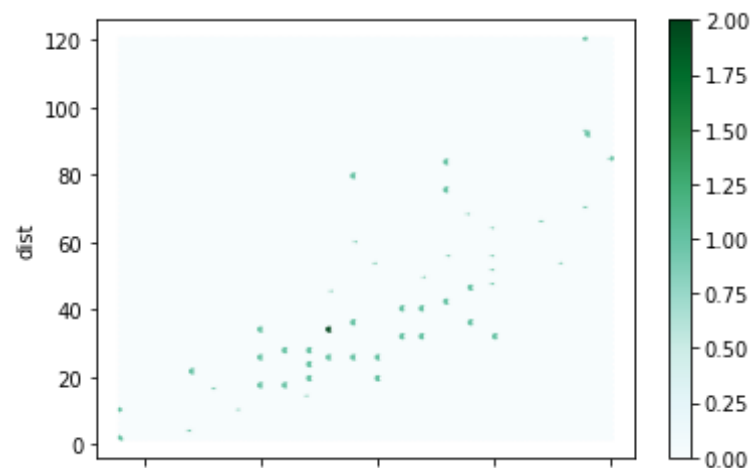
Que9 b

```
In [89]:  df2=pd.read_csv(r'C:\Users\Vrushabh\Downloads\Q9_b.csv')
```

```
In [90]:  df2
```

Out[90]:

	Unnamed: 0	SP	WT
0	1	104.185353	28.762059
1	2	105.461264	30.466833
2	3	105.461264	30.193597
3	4	113.461264	30.632114
4	5	104.461264	29.889149
...	...	...	...
76	77	169.598513	16.132947
77	78	150.576579	37.923113

Unnamed: 0		SP	WT
78	79	151.598513	15.769625
79	80	167.944460	39.423099
80	81	139.840817	34.948615

81 rows × 3 columns

In [91]: `df2.skew()`

Out[91]:
```
Unnamed: 0     0.000000
SP             1.611450
WT            -0.614753
dtype: float64
```

In [92]: `df2.kurtosis()`

Out[92]:
```
Unnamed: 0    -1.200000
SP             2.977329
WT             0.950291
dtype: float64
```
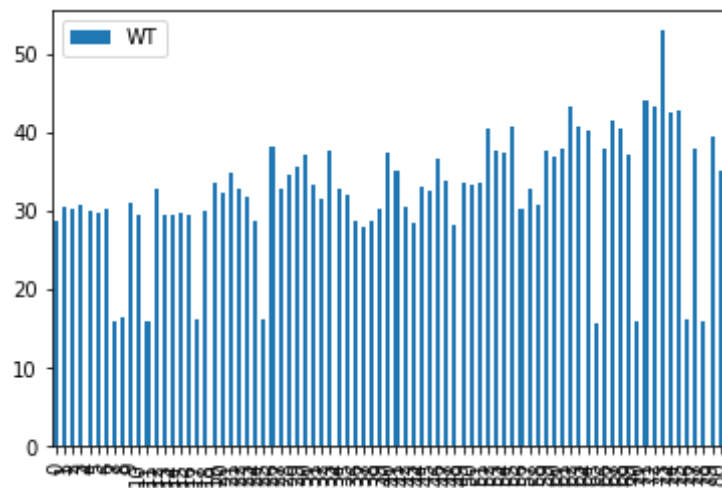
In [93]: `df2.plot(x="SP", y="WT")`

Out[93]: `<AxesSubplot:xlabel='SP'>`
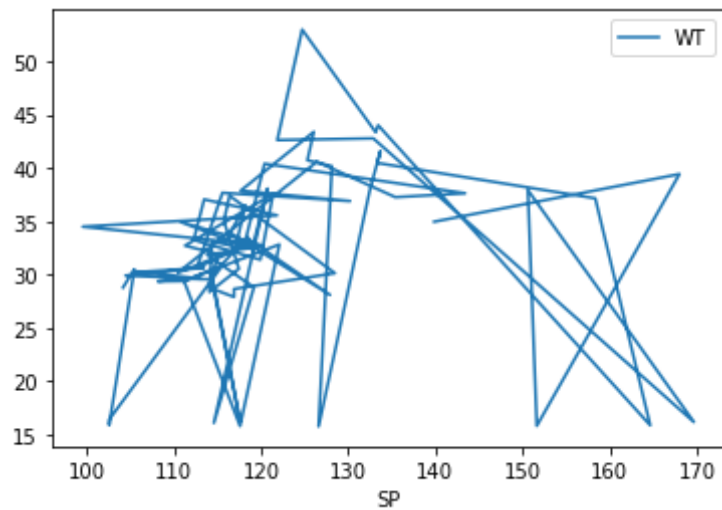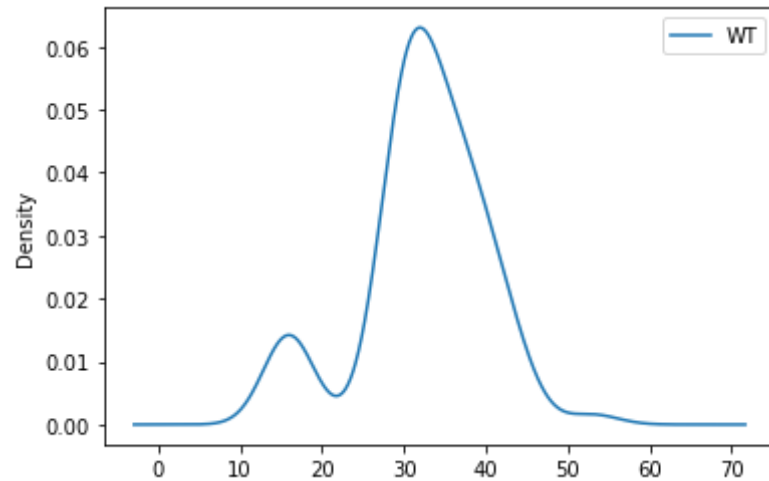
```
In [94]:  df2.plot.bar(y="WT")
```

Out[94]: <AxesSubplot:>
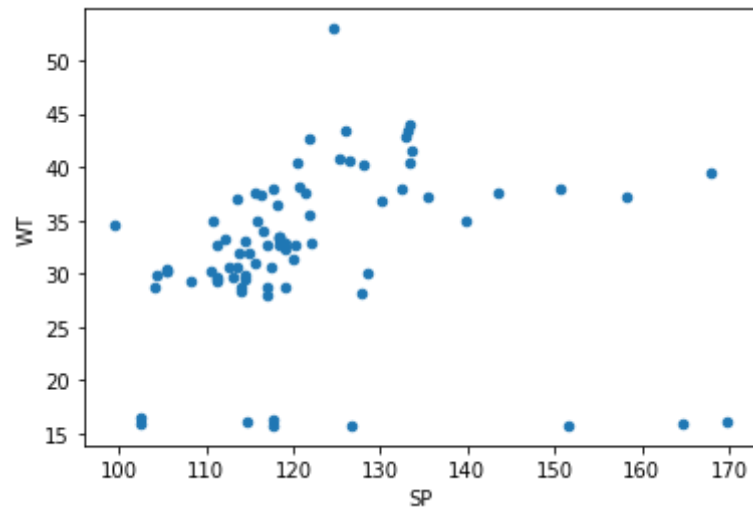


```
In [95]:  df2.plot.density(x="SP", y="WT")
```

Out[95]: <AxesSubplot:ylabel='Density'>
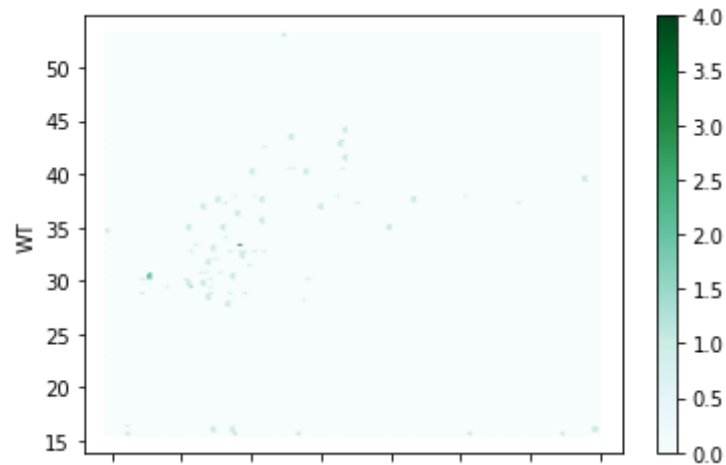
```
In [96]:   df2.plot.scatter(x="SP",y="WT")
```

Out[96]: <AxesSubplot:xlabel='SP', ylabel='WT'>


```
In [97]:   df2.plot.hexbin(x="SP",y="WT")
```

Out[97]: <AxesSubplot:xlabel='SP', ylabel='WT'>

Que20

```
In [10]:   df3=pd.read_csv(r'C:\Users\Vrushabh\Downloads\Cars.csv')
```

```
In [11]:   df3
```

Out[11]:

	HP	MPG	VOL	SP	WT
0	49	53.700681	89	104.185353	28.762059
1	55	50.013401	92	105.461264	30.466833
2	55	50.013401	92	105.461264	30.193597
3	70	45.696322	92	113.461264	30.632114
4	53	50.504232	92	104.461264	29.889149
...	...	...	...	...	...
76	322	36.900000	50	169.598513	16.132947
77	238	19.197888	115	150.576579	37.923113
78	263	34.000000	50	151.598513	15.769625
79	295	19.833733	119	167.944460	39.423099

	HP	MPG	VOL	SP	WT
80	236	12.101263	107	139.840817	34.948615

81 rows × 5 columns

In [7]:
```python
#   P(MPG>38)
```

In [12]:
```python
mpg=(df3['MPG']>38).sum()
total=len(df3)
print("Probability of (MPG>38)= {}".format(mpg/total))
```

Probability of (MPG>38)= 0.4074074074074074

In [13]:
```python
#   P(MPG>40)
```

In [14]:
```python
mpg=(df3['MPG']<40).sum()
total=len(df3)
print("Probability of (MPG<40)= {}".format(mpg/total))
```

Probability of (MPG<40)= 0.7530864197530864

In [15]:
```python
#   P(20<MPG>50)
```

In [17]:
```python
mpg=(((df3['MPG'])>20) & ((df3['MPG'])<50)).sum()
total=len(df3)
print("Probability of (20<MPG>50)= {}".format(mpg/total))
```

Probability of (20<MPG>50)= 0.8518518518518519

Que 21

In [18]:
```python
df3
```

Out[18]:

	HP	MPG	VOL	SP	WT
0	49	53.700681	89	104.185353	28.762059
1	55	50.013401	92	105.461264	30.466833

	HP	MPG	VOL	SP	WT
2	55	50.013401	92	105.461264	30.193597
3	70	45.696322	92	113.461264	30.632114
4	53	50.504232	92	104.461264	29.889149
...	...	...	...	...	...
76	322	36.900000	50	169.598513	16.132947
77	238	19.197888	115	150.576579	37.923113
78	263	34.000000	50	151.598513	15.769625
79	295	19.833733	119	167.944460	39.423099
80	236	12.101263	107	139.840817	34.948615

81 rows × 5 columns

In [20]: `df3.plot.density(y="MPG")`

Out[20]: `<AxesSubplot:ylabel='Density'>`
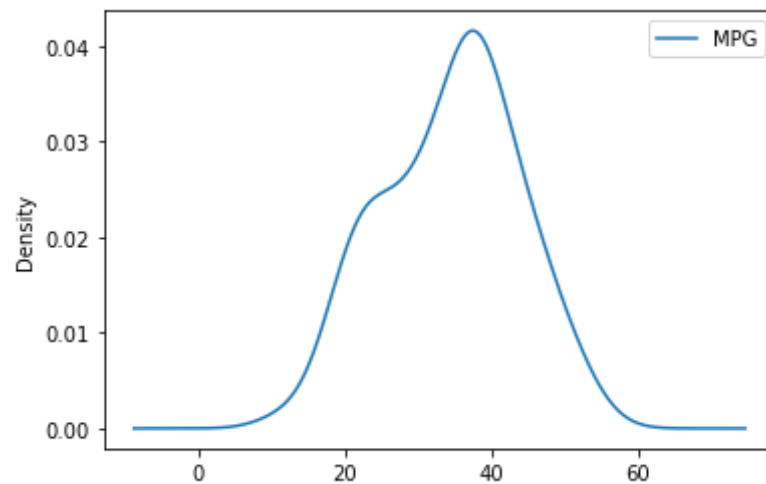
In [21]: `df4=pd.read_csv(r'C:\Users\Vrushabh\Downloads\wc-at.csv')`
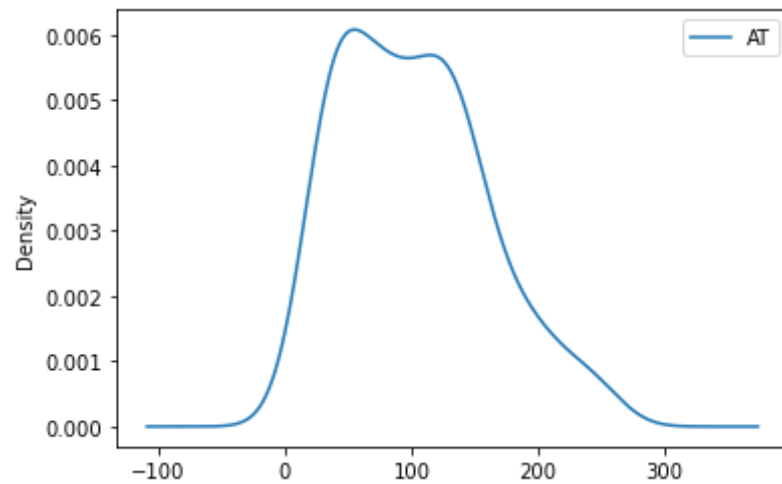
In [22]: `df4`

Out[22]:

	Waist	AT
0	74.75	25.72
1	72.60	25.89
2	81.80	42.60
3	83.95	42.80
4	74.65	29.84
...	...	...
104	100.10	124.00
105	93.30	62.20
106	101.80	133.00
107	107.90	208.00
108	108.50	208.00

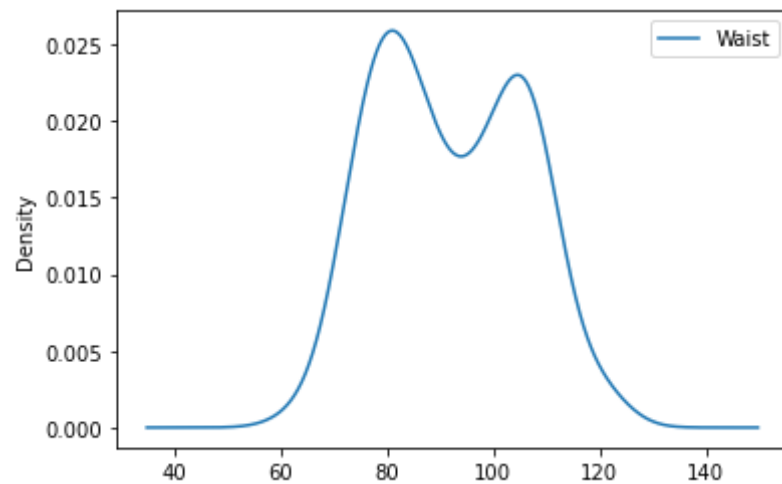109 rows × 2 columns

In [24]: `df4.plot.density(y="AT")`

Out[24]: `<AxesSubplot:ylabel='Density'>`

```
In [26]:  df4.plot.density(y="Waist")
```

Out[26]: <AxesSubplot:ylabel='Density'>


```
In [ ]:
```

```
In [4]:  #------------------------------------------------------------------------------------
```