

```
In [1]: ▶ import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

```
In [2]: ▶ data = pd.read_excel(r'C:\Users\hp\OneDrive\Documents\Desktop\car_sell.xls')
```

```
In [3]: ▶ data
```

Out[3]:

	Millage	Age(year)	sell price
0	69000	6	18000
1	35000	3	34000
2	57000	5	26100
3	22500	2	40000
4	46000	4	31500
5	59000	5	26750
6	52000	5	32000
7	72000	6	19300
8	91000	8	12000
9	67000	6	22000
10	83000	7	18700
11	79000	7	19500
12	59000	5	26000
13	58780	4	27500
14	82450	7	19400
15	25400	3	35000
16	28000	2	35500
17	69000	5	19700
18	87600	8	12800
19	52000	5	28200

```
In [4]: data.head()
```

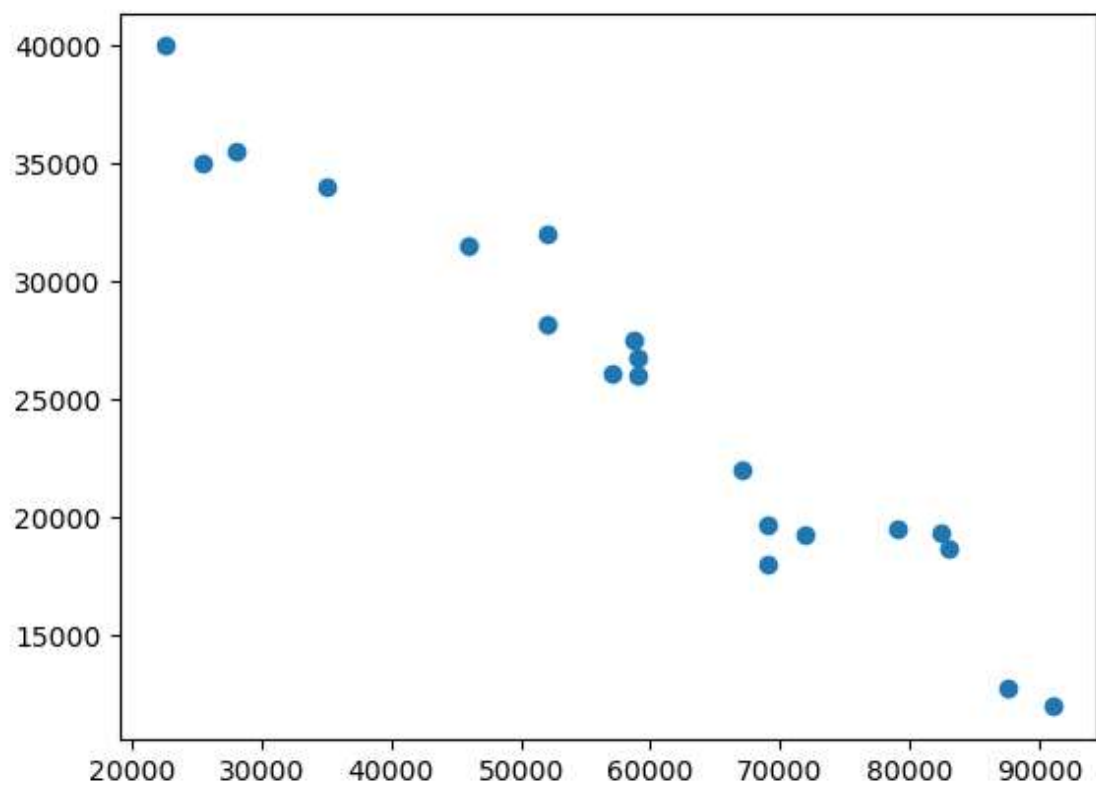
Out[4]:

	Millage	Age(year)	sell price
0	69000	6	18000
1	35000	3	34000
2	57000	5	26100
3	22500	2	40000
4	46000	4	31500

```
In [5]: # Car milage vs sell price
```

```
In [6]: plt.scatter(data["Millage"],data["sell price"])
```

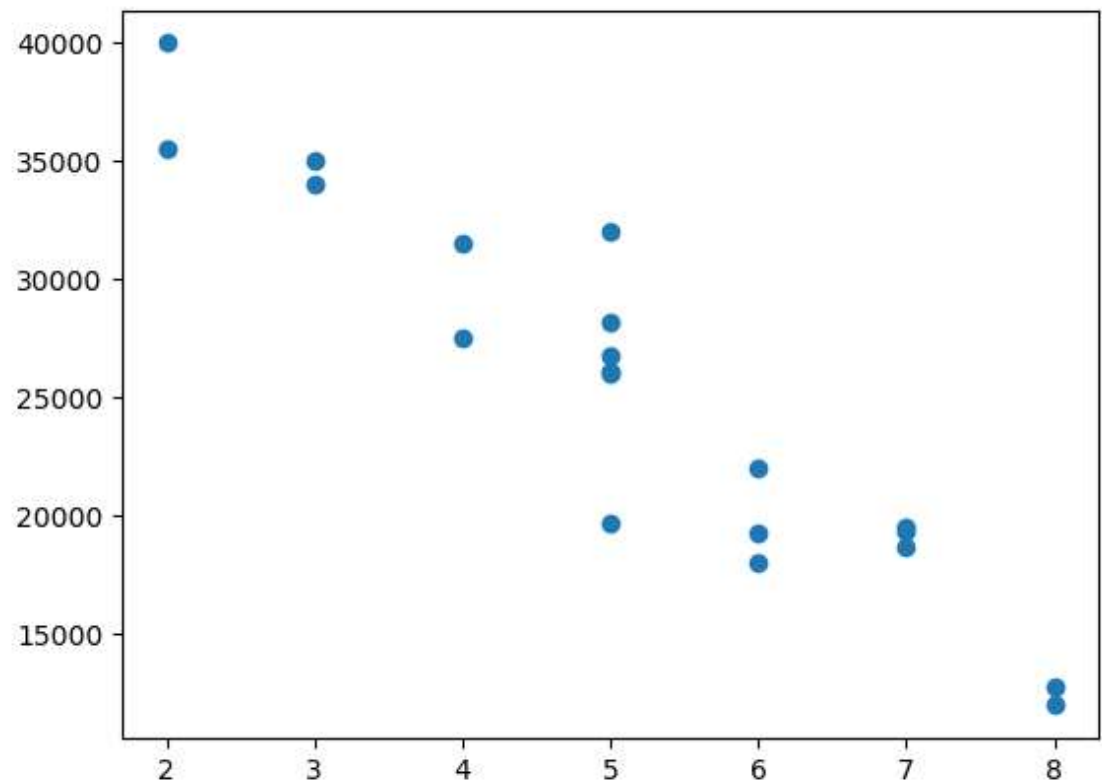
Out[6]: <matplotlib.collections.PathCollection at 0x2360bb28340>



```
In [7]: # car aAge vs sell price
```

```
In [8]: ▶ plt.scatter(data["Age(year)"],data["sell price"])
```

```
Out[8]: <matplotlib.collections.PathCollection at 0x2360c3f9f60>
```



```
In [9]: ▶ x = data[["Millage","Age(year)"]]  
y = data["sell price"]
```

In [10]:  x

Out[10]:

	Millage	Age(year)
0	69000	6
1	35000	3
2	57000	5
3	22500	2
4	46000	4
5	59000	5
6	52000	5
7	72000	6
8	91000	8
9	67000	6
10	83000	7
11	79000	7
12	59000	5
13	58780	4
14	82450	7
15	25400	3
16	28000	2
17	69000	5
18	87600	8
19	52000	5

In [11]: `y`

```
Out[11]: 0    18000
         1    34000
         2    26100
         3    40000
         4    31500
         5    26750
         6    32000
         7    19300
         8    12000
         9    22000
        10    18700
        11    19500
        12    26000
        13    27500
        14    19400
        15    35000
        16    35500
        17    19700
        18    12800
        19    28200
        Name: sell price, dtype: int64
```

In [12]: `from sklearn.model_selection import train_test_split`

In [13]: `x_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2,rand`

In [14]: `len(x_train)`

```
Out[14]: 16
```

In [15]: `len(x_test)`

```
Out[15]: 4
```

In [16]: `from sklearn.linear_model import LinearRegression`  
`clf = LinearRegression()`

In [17]: `clf.fit(x_train, y_train)`

```
Out[17]: ▼ LinearRegression
         LinearRegression()
```

In [18]: `clf.predict(x_test)`


```
Out[18]: array([20474.0745775 , 16352.07892168, 25174.14834912, 27197.42175439])
```

In [19]:  y\_test

Out[19]:

7	19300
10	18700
5	26750
6	32000

Name: sell price, dtype: int64

In [20]:  clf.score(x\_test, y\_test) *# Accuracy is 73 percent*

Out[20]: 0.7332339593090137

In [ ]: 

In [ ]: 