In [ ]:

```
#Use the car dataset containing information of Mileage, Age, and Sell-Price. Split the dataset into training and test dataset in 80:20 ratio. Train the Linear Regression model on the training dataset and predict the Sell-Price for test dataset. (Multivariate Linear Regression)
```

In [2]:

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn import metrics
df=pd.read_csv('C:/Users/Deep/Desktop/car-dataset.csv')
```

In [3]:

```
df.head()
```

Out[3]:

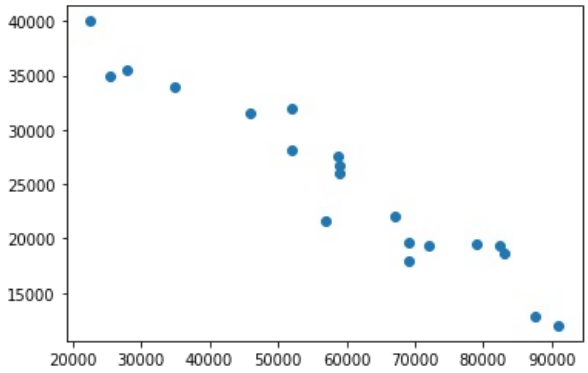
	Mileage	Age	Sell-price
0	69000	6	18000
1	35000	3	34000
2	57000	5	21600
3	22500	2	40000
4	46000	4	31500

In [4]:

```
plt.scatter(df['Mileage'],df['Sell-price'])
```

Out[4]:

```
<matplotlib.collections.PathCollection at 0x21e6c86ef48>
```

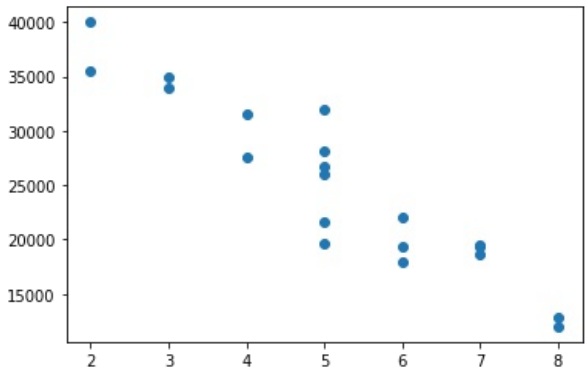


In [5]:

```
plt.scatter(df['Age'],df['Sell-price'])
```

Out[5]:

```
<matplotlib.collections.PathCollection at 0x21e6d904648>
```



In [6]:

```
x=df[['Mileage','Age']]
y=df['Sell-price']
```

In [7]:

```
x
```

Out[7]:

	Mileage	Age
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 [8]:

```
y
```

Out[8]:

```
0     18000
1     34000
2     21600
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 [9]:

```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2)
len(x_train)
```

Out[9]:

```
16
```

In [10]:

```
x_train
```

Out[10]:

	Mileage	Age
16	28000	2
4	46000	4
18	87600	8
1	35000	3
10	83000	7
2	57000	5
8	91000	8
14	82450	7
17	69000	5
5	59000	5
3	22500	2
9	67000	6
15	25400	3
19	52000	5
11	79000	7
7	72000	6

In [11]:

```
from sklearn.linear_model import LinearRegression
#from sklearn.metrics import score
clf=LinearRegression()
clf.fit(x_train,y_train)
```

Out[11]:

```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

In [12]:

```
y_pred=clf.predict(x_test)
y_pred
```

Out[12]:

```
array([25068.0639464 , 21422.03073711, 27317.01327595, 25571.99366383])
```

In [13]:

```
y_test
```

Out[13]:

```
12    26000
0     18000
6     32000
13    27500
Name: Sell-price, dtype: int64
```

In [14]:

```
print(x_test)
len(x_test)
```

```
    Mileage  Age
12    59000    5
0     69000    6
6     52000    5
13    58780    4
```

Out[14]:

```
4
```

In [15]:

```
r_sq=clf.score(x,y)
print(r_sq)
```

```
0.9183108649948636
```