In [1]:
```
import pandas as pd
df = pd.read_csv("C:/Users/aziz/Desktop/Downloads/6_train_test_split/carprices.cs
df.head()
```

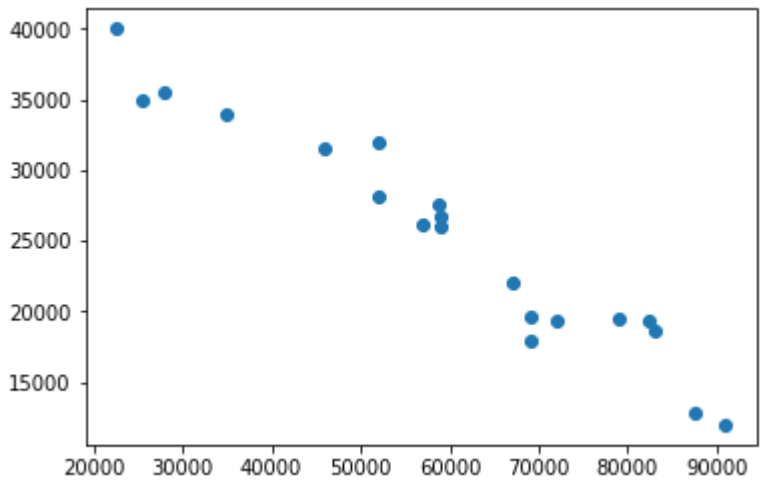
Out[1]:

	Mileage	Age(yrs)	Sell Price($)
0	69000	6	18000
1	35000	3	34000
2	57000	5	26100
3	22500	2	40000
4	46000	4	31500

In [3]:
```
import matplotlib.pyplot as plt
%matplotlib inline
```

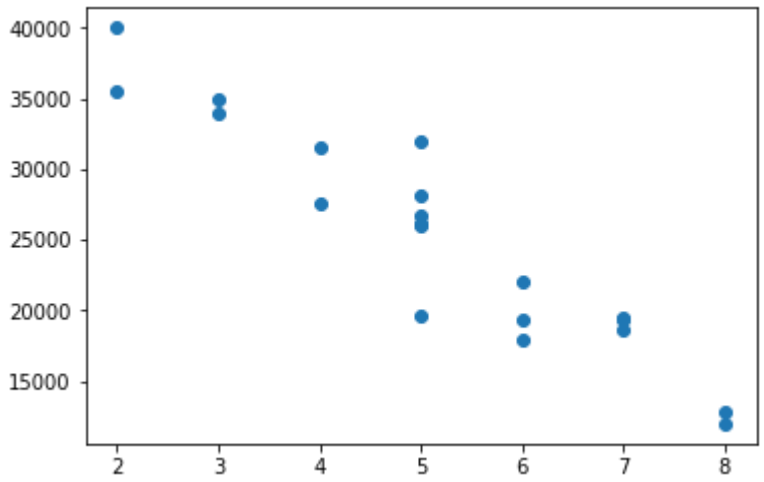
In [4]:
```
plt.scatter(df['Mileage'],df['Sell Price($)'])
```

Out[4]:
```
<matplotlib.collections.PathCollection at 0x1fad60f2eb8>
```



In [5]:
```
plt.scatter(df['Age(yrs)'],df['Sell Price($)'])
```

Out[5]:
```
<matplotlib.collections.PathCollection at 0x1fad6193e10>
```



In [6]:
```
X = df[['Mileage','Age(yrs)']]
```

In [7]:
```
y = df['Sell Price($)']
```

In [8]:
```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.3)
```

In [9]:
```
X_train
```

Out[9]:

	Mileage	Age(yrs)
10	83000	7
0	69000	6
17	69000	5
19	52000	5
12	59000	5
16	28000	2
11	79000	7
18	87600	8
3	22500	2
14	82450	7
15	25400	3
5	59000	5
6	52000	5
7	72000	6

In [11]:
```
X_test
```

Out[11]:

	Mileage	Age(yrs)
4	46000	4
8	91000	8
1	35000	3
13	58780	4
9	67000	6
2	57000	5

In [12]:
```
y_train
```

Out[12]:
```
10    18700
0     18000
17    19700
19    28200
12    26000
16    35500
11    19500
18    12800
3     40000
14    19400
15    35000
5     26750
6     32000
7     19300
Name: Sell Price($), dtype: int64
```

In [13]:
```
y_test
```

Out[13]:
```
4     31500
8     12000
1     34000
13    27500
9     22000
2     26100
Name: Sell Price($), dtype: int64
```

In [14]:
```
from sklearn.linear_model import LinearRegression
clf = LinearRegression()
clf.fit(X_train, y_train)
```

Out[14]:
```
LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None,
         normalize=False)
```

In [15]:
```
X_test
```

Out[15]:

	Mileage	Age(yrs)
4	46000	4
8	91000	8
1	35000	3
13	58780	4
9	67000	6
2	57000	5

In [16]:
```
clf.predict(X_test)
```

Out[16]:
```
array([30019.27563313, 14131.29279112, 33890.42390824, 24863.95473603,
       22680.36882446, 26148.12735802])
```

In [17]:
```
y_test
```

Out[17]:
```
4     31500
8     12000
1     34000
13    27500
9     22000
2     26100
Name: Sell Price($), dtype: int64
```

In [18]:
```
clf.score(X_test, y_test)
```

Out[18]:
```
0.9538895184252855
```

In [20]:
```
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size=0.3,random_sta
X_test
```

Out[20]:

	Mileage	Age(yrs)
7	72000	6
10	83000	7
5	59000	5
6	52000	5
3	22500	2
18	87600	8

In [ ]:
```

```