

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sn
from sklearn import linear_model
```

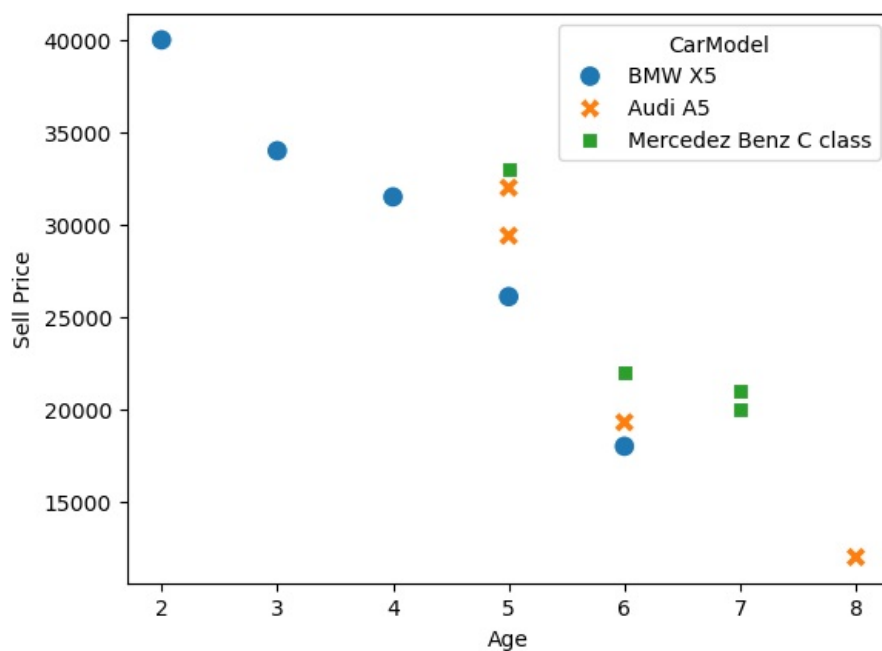
```
In [2]: df = pd.read_csv("carprices.csv")
df
```

```
Out[2]:
```

	CarModel	Mileage	Sell Price	Age
0	BMW X5	69000	18000	6
1	BMW X5	35000	34000	3
2	BMW X5	57000	26100	5
3	BMW X5	22500	40000	2
4	BMW X5	46000	31500	4
5	Audi A5	59000	29400	5
6	Audi A5	52000	32000	5
7	Audi A5	72000	19300	6
8	Audi A5	91000	12000	8
9	Mercedes Benz C class	67000	22000	6
10	Mercedes Benz C class	83000	20000	7
11	Mercedes Benz C class	79000	21000	7
12	Mercedes Benz C class	59000	33000	5

```
In [7]: plt.xlabel("Age")
plt.ylabel("Sell Price")
sn.scatterplot(data=df, x='Age', y='Sell Price', hue='CarModel', style='CarModel', s=100)
```

```
Out[7]: <Axes: xlabel='Age', ylabel='Sell Price'>
```



```
In [9]: dummies = pd.get_dummies(df['CarModel']).astype(int)
dummies
```

Out[9]:

	Audi A5	BMW X5	Mercedes Benz C class
0	0	1	0
1	0	1	0
2	0	1	0
3	0	1	0
4	0	1	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	0	0	1
10	0	0	1
11	0	0	1
12	0	0	1

```
In [13]: merged=pd.concat([df,dummies],axis='columns')
merged
```

Out[13]:

	CarModel	Mileage	Sell Price	Age	Audi A5	BMW X5	Mercedes Benz C class
0	BMW X5	69000	18000	6	0	1	0
1	BMW X5	35000	34000	3	0	1	0
2	BMW X5	57000	26100	5	0	1	0
3	BMW X5	22500	40000	2	0	1	0
4	BMW X5	46000	31500	4	0	1	0
5	Audi A5	59000	29400	5	1	0	0
6	Audi A5	52000	32000	5	1	0	0
7	Audi A5	72000	19300	6	1	0	0
8	Audi A5	91000	12000	8	1	0	0
9	Mercedes Benz C class	67000	22000	6	0	0	1
10	Mercedes Benz C class	83000	20000	7	0	0	1
11	Mercedes Benz C class	79000	21000	7	0	0	1
12	Mercedes Benz C class	59000	33000	5	0	0	1

```
In [21]: final=merged.drop(['CarModel','Mercedes Benz C class','Sell Price'],axis='columns')
final
```

Out[21]:

	Mileage	Age	Audi A5	BMW X5
0	69000	6	0	1
1	35000	3	0	1
2	57000	5	0	1
3	22500	2	0	1
4	46000	4	0	1
5	59000	5	1	0
6	52000	5	1	0
7	72000	6	1	0
8	91000	8	1	0
9	67000	6	0	0
10	83000	7	0	0
11	79000	7	0	0
12	59000	5	0	0

```
In [29]: y = df['Sell Price']
y
```

```
Out[29]: 0      18000
         1      34000
         2      26100
         3      40000
         4      31500
         5      29400
         6      32000
         7      19300
         8      12000
         9      22000
        10      20000
        11      21000
        12      33000
        Name: Sell Price, dtype: int64
```

```
In [31]: model = linear_model.LinearRegression()
        model.fit(final,y)
```

```
Out[31]: LinearRegression ⓘ ?
        LinearRegression()
```

```
In [33]: model.score(final,y)
```

```
Out[33]: 0.9417050937281083
```

```
In [35]: model.predict([[34000,6,1,0]])
```

```
C:\Users\91901\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature names
, but LinearRegression was fitted with feature names
warnings.warn(
```

```
Out[35]: array([35944.2122564])
```

```
In [37]: model.predict([[69000,6,0,1]])
```

```
C:\Users\91901\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature names
, but LinearRegression was fitted with feature names
warnings.warn(
```

```
Out[37]: array([18705.2723644])
```

```
In [39]: model.predict([[45000,4,0,0]])
```

```
C:\Users\91901\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature names
, but LinearRegression was fitted with feature names
warnings.warn(
```

```
Out[39]: array([36991.31721061])
```

```
In [41]: model.predict([[86000,7,0,1]])
```

```
C:\Users\91901\anaconda3\Lib\site-packages\sklearn\base.py:493: UserWarning: X does not have valid feature names
, but LinearRegression was fitted with feature names
warnings.warn(
```

```
Out[41]: array([11080.74313219])
```

```
In [43]: model.score(final,y)
```

```
Out[43]: 0.9417050937281083
```

```
In [ ]:
```