

In [1]: `import pandas as pd
import seaborn as sns`

In [2]: `df=pd.read_csv("C:\\Users\\Pranav\\Downloads\\train.csv")
df.head()`

Out[2]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

In [3]: `df2=df[['Survived','Pclass','Age','Parch']]`

In [4]: `df3=df2.fillna(df2.mean())`

In [5]: `df3`

Out[5]:

	Survived	Pclass	Age	Parch
0	0	3	22.000000	0
1	1	1	38.000000	0
2	1	3	26.000000	0
3	1	1	35.000000	0
4	0	3	35.000000	0
...	...	...	...	...
886	0	2	27.000000	0
887	1	1	19.000000	0
888	0	3	29.699118	2
889	1	1	26.000000	0
890	0	3	32.000000	0

891 rows × 4 columns

In [6]: `X=df3.drop("Survived",axis=1)
y=df3["Survived"]
print("shape of x= ",x.shape)
print("shape of y= ",y.shape)`

shape of x= (891, 3)  
shape of y= (891,)

In [7]: `from sklearn.model_selection import train_test_split`

In [8]: `X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.2,random_state=51)`

In [9]: `print("shape of X_train= ",X_train.shape)
print("shape of X_test= ",X_test.shape)
print("shape of y_train= ",y_train.shape)
print("shape of y_test= ",y_test.shape)`

shape of X\_train= (712, 3)  
shape of X\_test= (179, 3)  
shape of y\_train= (712,)  
shape of y\_test= (179,)

In [10]: `X_train`

Out[10]:

	Pclass	Age	Parch
770	3	24.000000	0
152	3	55.500000	0
731	3	11.000000	0
775	3	18.000000	0
324	3	29.699118	2
...	...	...	...
528	3	39.000000	0
709	3	29.699118	1
736	3	48.000000	3
485	3	29.699118	1
57	3	28.500000	0

712 rows × 3 columns

In [11]: `cars=pd.read_csv("C:\\Users\\Pranav\\Downloads\\carprices.csv")`

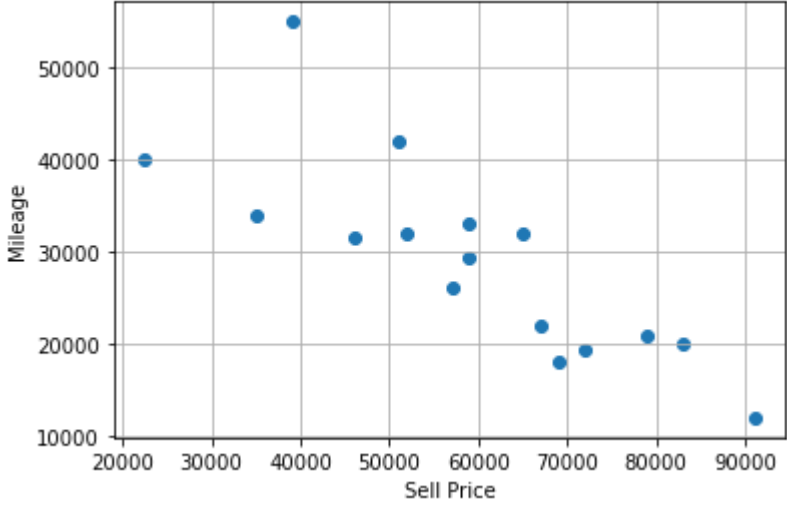
In [12]: `cars.head()`

Out[12]:

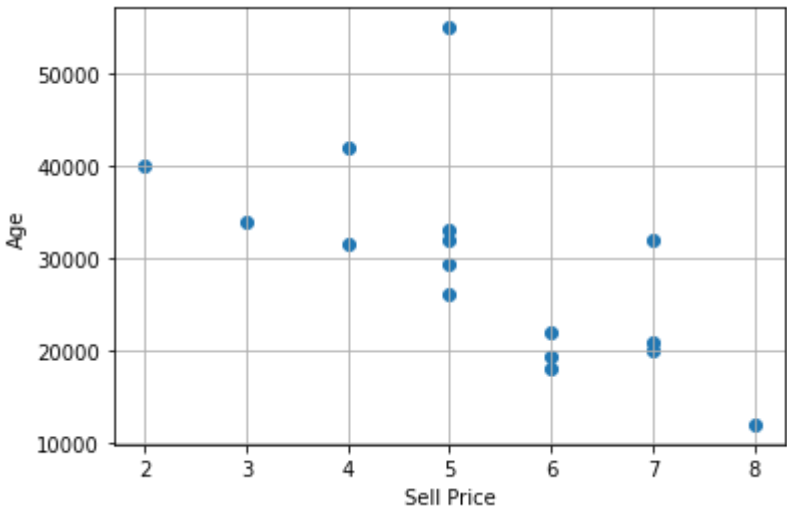
	Car Model	Mileage	Sell Price	Age
0	BMW X5	69000	18000	6
1	BMW X5	35000	34000	3
2	BMW X5	57000	26100	5
3	BMW X5	22500	40000	2
4	BMW X5	46000	31500	4

In [13]: `from matplotlib import pyplot as plt`

In [18]: `plt.scatter(cars['Mileage'],cars['Sell Price'])
plt.xlabel('Sell Price')
plt.ylabel('Mileage')
plt.grid(True)
plt.show()`



In [21]: `plt.scatter(cars['Age'],cars['Sell Price'])
plt.xlabel('Sell Price')
plt.ylabel('Age')
plt.grid(True)
plt.show()`



In [26]: `X=cars[['Mileage','Age']]
y=cars['Sell Price']`

In [27]: `X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.25,random_state=51)`

In [28]: `print("shape of X_train= ",X_train.shape)
print("shape of X_test= ",X_test.shape)
print("shape of y_train= ",y_train.shape)
print("shape of y_test= ",y_test.shape)`

shape of X\_train= (12, 2)  
shape of X\_test= (4, 2)  
shape of y\_train= (12,)  
shape of y\_test= (4,)

In [29]: `X_train`

Out[29]:

	Mileage	Age
2	57000	5
1	35000	3
4	46000	4
6	52000	5
10	83000	7
12	59000	5
13	51000	4
15	39000	5
0	69000	6
5	59000	5
9	67000	6
14	65000	7

In [ ]: