assignement-3

September 20, 2023

NAME: P VISHAL GUPTA

1 1.Import the necessary libraries

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

2 2.Import the dataset

[2]:
```
dataset=pd.read_csv("Titanic-Dataset.csv")
```

[3]:
```
dataset
```

[3]:
```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           …       …     …
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex   Age  SibSp  \
0                             Braund, Mr. Owen Harris    male  22.0      1
1    Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.0      1
2                              Heikkinen, Miss. Laina  female  26.0      0
3        Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                            Allen, Mr. William Henry    male  35.0      0
..                                                 …     …   …    …
886                            Montvila, Rev. Juozas    male  27.0      0
```

```
887                      Graham, Miss. Margaret Edith  female  19.0      0
888          Johnston, Miss. Catherine Helen "Carrie"  female   NaN      1
889                             Behr, Mr. Karl Howell    male  26.0      0
890                               Dooley, Mr. Patrick    male  32.0      0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
2        0  STON/O2. 3101282   7.9250   NaN        S
3        0            113803  53.1000  C123        S
4        0            373450   8.0500   NaN        S
..     …               …      …   …      …
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
888      2        W./C. 6607  23.4500   NaN        S
889      0            111369  30.0000  C148        C
890      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

[5]:
```
dataset.head(3)
```

[5]:
```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3

                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
```

[6]:
```
dataset.tail()
```

[6]:
```
     PassengerId  Survived  Pclass                                      Name  \
886          887         0       2                     Montvila, Rev. Juozas
887          888         1       1              Graham, Miss. Margaret Edith
888          889         0       3  Johnston, Miss. Catherine Helen "Carrie"
889          890         1       1                     Behr, Mr. Karl Howell
890          891         0       3                       Dooley, Mr. Patrick

        Sex   Age  SibSp  Parch      Ticket   Fare Cabin Embarked
```

```
886    male  27.0      0      0      211536  13.00   NaN        S
887  female  19.0      0      0      112053  30.00   B42        S
888  female   NaN      1      2  W./C. 6607  23.45   NaN        S
889    male  26.0      0      0      111369  30.00  C148        C
890    male  32.0      0      0      370376   7.75   NaN        Q
```

[7]:
```
dataset.shape
```

[7]:
```
(891, 12)
```

[8]:
```
dataset.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  891 non-null    int64  
 1   Survived     891 non-null    int64  
 2   Pclass       891 non-null    int64  
 3   Name         891 non-null    object 
 4   Sex          891 non-null    object 
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64  
 7   Parch        891 non-null    int64  
 8   Ticket       891 non-null    object 
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object 
 11  Embarked     889 non-null    object 
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

[10]:
```
dataset.describe()
```

[10]:
```
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000   
mean    446.000000    0.383838    2.308642   29.699118    0.523008   
std     257.353842    0.486592    0.836071   14.526497    1.102743   
min       1.000000    0.000000    1.000000    0.420000    0.000000   
25%     223.500000    0.000000    2.000000   20.125000    0.000000   
50%     446.000000    0.000000    3.000000   28.000000    0.000000   
75%     668.500000    1.000000    3.000000   38.000000    1.000000   
max     891.000000    1.000000    3.000000   80.000000    8.000000   

            Parch        Fare  
count  891.000000  891.000000  
mean     0.381594   32.204208  
```

```
std      0.806057   49.693429
min      0.000000    0.000000
25%      0.000000    7.910400
50%      0.000000   14.454200
75%      0.000000   31.000000
max      6.000000  512.329200
```

3 3.Handling null values

[11]:
```
dataset.isnull().any()
```

[11]:
```
PassengerId    False
Survived       False
Pclass         False
Name           False
Sex            False
Age             True
SibSp          False
Parch          False
Ticket         False
Fare           False
Cabin           True
Embarked        True
dtype: bool
```

[12]:
```
dataset["Age"].fillna(dataset["Age"].mean(),inplace=True)
dataset
```

[12]:
```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           …       …     …
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex         Age  \
0                             Braund, Mr. Owen Harris    male  22.000000
1    Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.000000
2                                Heikkinen, Miss. Laina  female  26.000000
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.000000
```

```
4                              Allen, Mr. William Henry    male  35.000000
..                                                  …     …       …
886                               Montvila, Rev. Juozas    male  27.000000
887                        Graham, Miss. Margaret Edith  female  19.000000
888            Johnston, Miss. Catherine Helen "Carrie"  female  29.699118
889                               Behr, Mr. Karl Howell    male  26.000000
890                                 Dooley, Mr. Patrick    male  32.000000

     SibSp  Parch            Ticket     Fare Cabin Embarked
0        1      0         A/5 21171   7.2500   NaN        S
1        1      0          PC 17599  71.2833   C85        C
2        0      0  STON/O2. 3101282   7.9250   NaN        S
3        1      0            113803  53.1000  C123        S
4        0      0            373450   8.0500   NaN        S
..     …    …                 …       …   …      …
886      0      0            211536  13.0000   NaN        S
887      0      0            112053  30.0000   B42        S
888      1      2        W./C. 6607  23.4500   NaN        S
889      0      0            111369  30.0000  C148        C
890      0      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

[13]:
```
dataset["Embarked"].fillna(dataset["Embarked"].mode()[0],inplace=True)
dataset
```

[13]:
```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           …       …     …
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex        Age  \
0                             Braund, Mr. Owen Harris    male  22.000000
1    Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.000000
2                                Heikkinen, Miss. Laina  female  26.000000
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.000000
4                              Allen, Mr. William Henry    male  35.000000
..                                                   …     …       …
886                               Montvila, Rev. Juozas    male  27.000000
```

```
887                      Graham, Miss. Margaret Edith  female  19.000000
888          Johnston, Miss. Catherine Helen "Carrie"  female  29.699118
889                             Behr, Mr. Karl Howell    male  26.000000
890                               Dooley, Mr. Patrick    male  32.000000

     SibSp  Parch            Ticket     Fare Cabin Embarked
0        1      0         A/5 21171   7.2500   NaN        S
1        1      0          PC 17599  71.2833   C85        C
2        0      0  STON/O2. 3101282   7.9250   NaN        S
3        1      0            113803  53.1000  C123        S
4        0      0            373450   8.0500   NaN        S
..     …    …                …       …   …       …
886      0      0            211536  13.0000   NaN        S
887      0      0            112053  30.0000   B42        S
888      1      2        W./C. 6607  23.4500   NaN        S
889      0      0            111369  30.0000  C148        C
890      0      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

[14]:
```
dataset["Cabin"].fillna(dataset["Cabin"].mode()[1],inplace=True)
dataset
```

[14]:
```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           …       …     …
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                 Name     Sex         Age  \
0                             Braund, Mr. Owen Harris    male  22.000000
1    Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.000000
2                                Heikkinen, Miss. Laina  female  26.000000
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.000000
4                             Allen, Mr. William Henry    male  35.000000
..                                                 …     …         …
886                              Montvila, Rev. Juozas    male  27.000000
887                       Graham, Miss. Margaret Edith  female  19.000000
888          Johnston, Miss. Catherine Helen "Carrie"  female  29.699118
889                              Behr, Mr. Karl Howell    male  26.000000
```

```
890                                Dooley, Mr. Patrick    male  32.000000

     SibSp  Parch            Ticket     Fare        Cabin Embarked
0        1      0         A/5 21171   7.2500  C23 C25 C27        S
1        1      0          PC 17599  71.2833          C85        C
2        0      0  STON/O2. 3101282   7.9250  C23 C25 C27        S
3        1      0            113803  53.1000         C123        S
4        0      0            373450   8.0500  C23 C25 C27        S
..     …    …                …       …          …      …
886      0      0            211536  13.0000  C23 C25 C27        S
887      0      0            112053  30.0000          B42        S
888      1      2        W./C. 6607  23.4500  C23 C25 C27        S
889      0      0            111369  30.0000         C148        C
890      0      0            370376   7.7500  C23 C25 C27        Q

[891 rows x 12 columns]
```

[15]:
```
dataset.isnull().sum()
```

[15]:
```
PassengerId    0
Survived       0
Pclass         0
Name           0
Sex            0
Age            0
SibSp          0
Parch          0
Ticket         0
Fare           0
Cabin          0
Embarked       0
dtype: int64
```

4 4.Data Visualization.

[16]:
```
corr=dataset.corr()
corr
```

```
C:\Users\Vishal Gupta\AppData\Local\Temp\ipykernel_10548\1091080309.py:1:
FutureWarning: The default value of numeric_only in DataFrame.corr is
deprecated. In a future version, it will default to False. Select only valid
columns or specify the value of numeric_only to silence this warning.
  corr=dataset.corr()
```

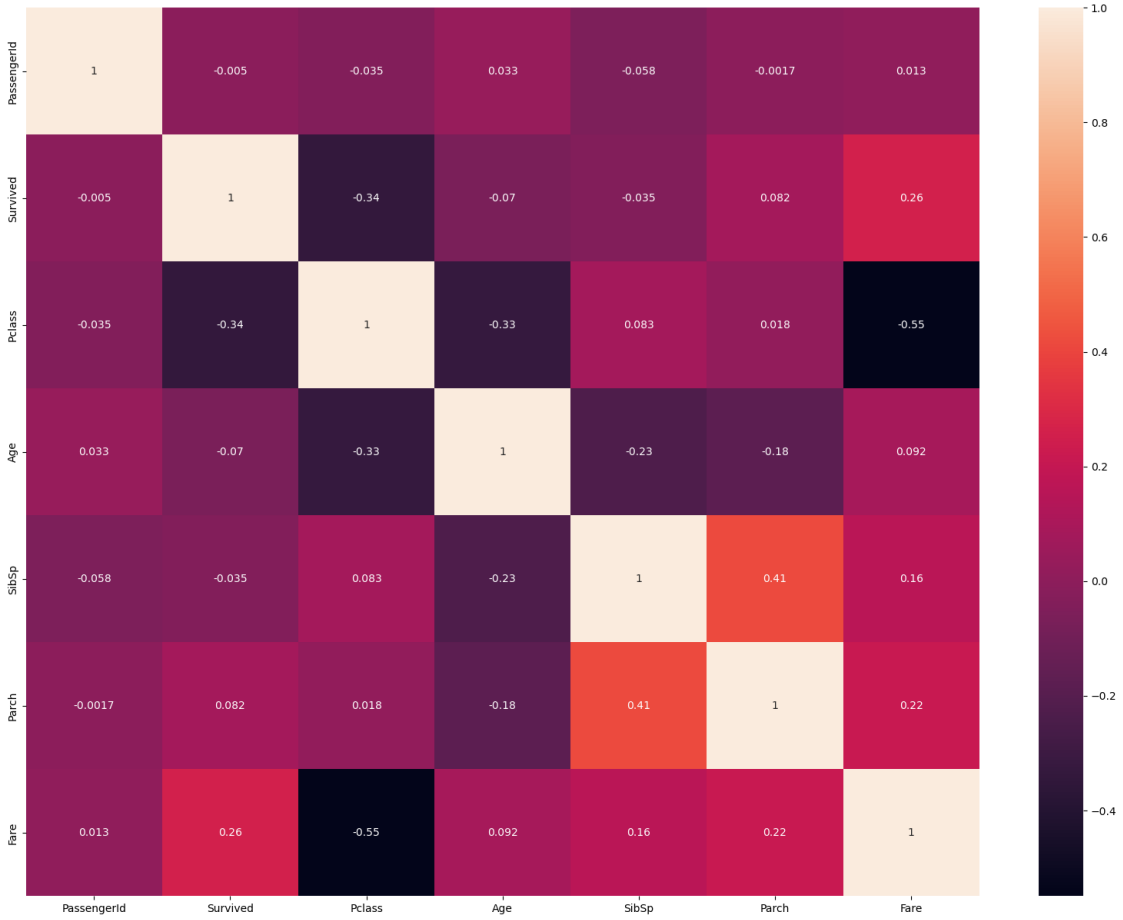
[16]:
```
             PassengerId  Survived    Pclass       Age     SibSp     Parch  \
PassengerId     1.000000 -0.005007 -0.035144  0.033207 -0.057527 -0.001652
```

```
Survived      -0.005007  1.000000 -0.338481 -0.069809 -0.035322  0.081629
Pclass        -0.035144 -0.338481  1.000000 -0.331339  0.083081  0.018443
Age            0.033207 -0.069809 -0.331339  1.000000 -0.232625 -0.179191
SibSp         -0.057527 -0.035322  0.083081 -0.232625  1.000000  0.414838
Parch         -0.001652  0.081629  0.018443 -0.179191  0.414838  1.000000
Fare           0.012658  0.257307 -0.549500  0.091566  0.159651  0.216225

                 Fare
PassengerId  0.012658
Survived     0.257307
Pclass      -0.549500
Age          0.091566
SibSp        0.159651
Parch        0.216225
Fare         1.000000
```

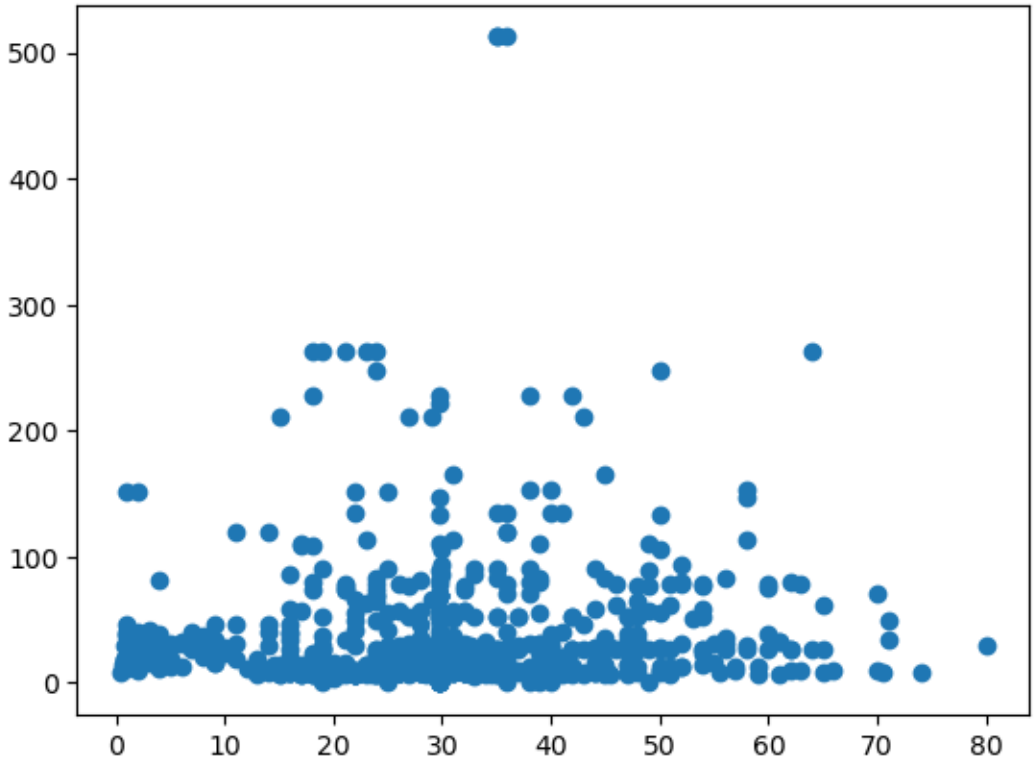
[17]:
```
plt.subplots(figsize=(20,15))
sns.heatmap(corr,annot=True)
```

[17]: <Axes: >



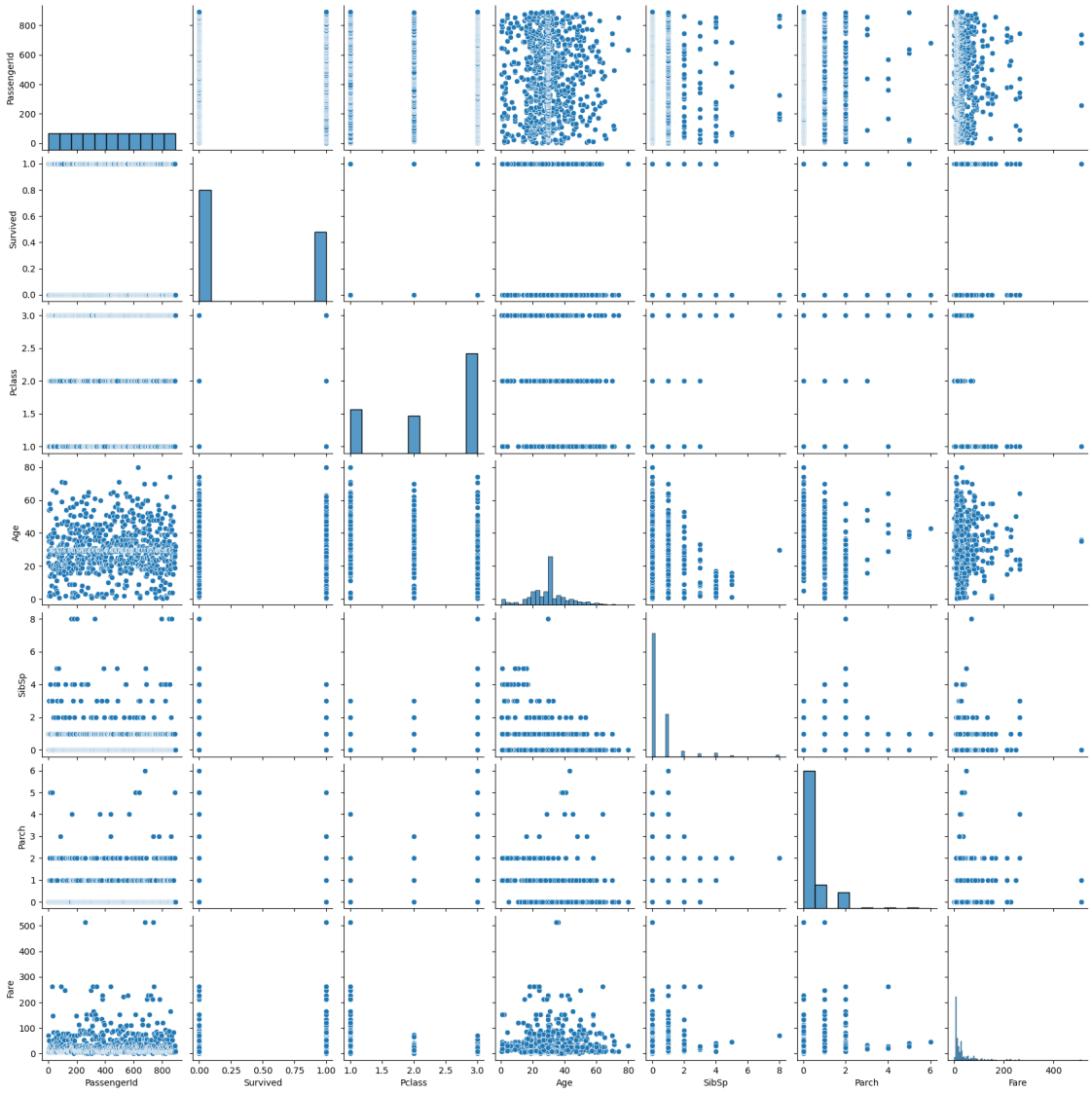
[18]:
```
plt.scatter(dataset["Age"],dataset["Fare"])
```

[18]: <matplotlib.collections.PathCollection at 0x1aab98af210>



[19]:
```
sns.pairplot(dataset)
```

[19]: <seaborn.axisgrid.PairGrid at 0x1aabb8856d0>



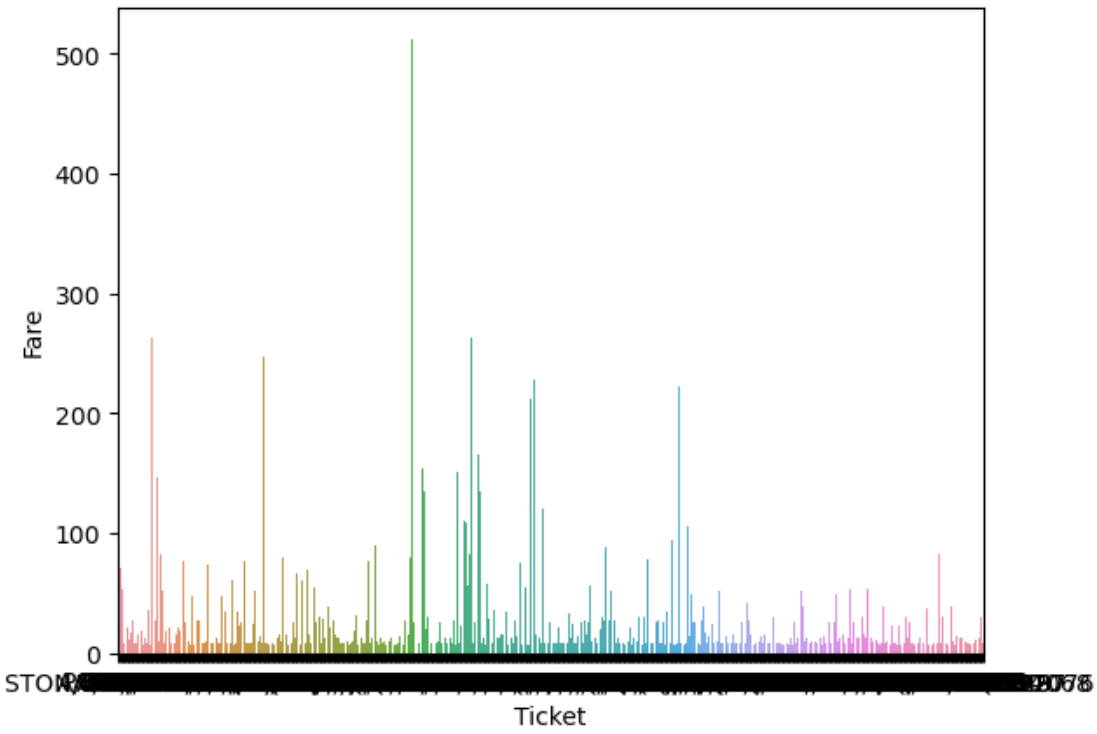
[20]:
```
sns.barplot(x=dataset["Ticket"],y=dataset["Fare"],ci=0)
```

```
C:\Users\Vishal Gupta\AppData\Local\Temp\ipykernel_10548\4120788058.py:1:
FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=('ci', 0)` for the same effect.

  sns.barplot(x=dataset["Ticket"],y=dataset["Fare"],ci=0)
```

[20]: <Axes: xlabel='Ticket', ylabel='Fare'>



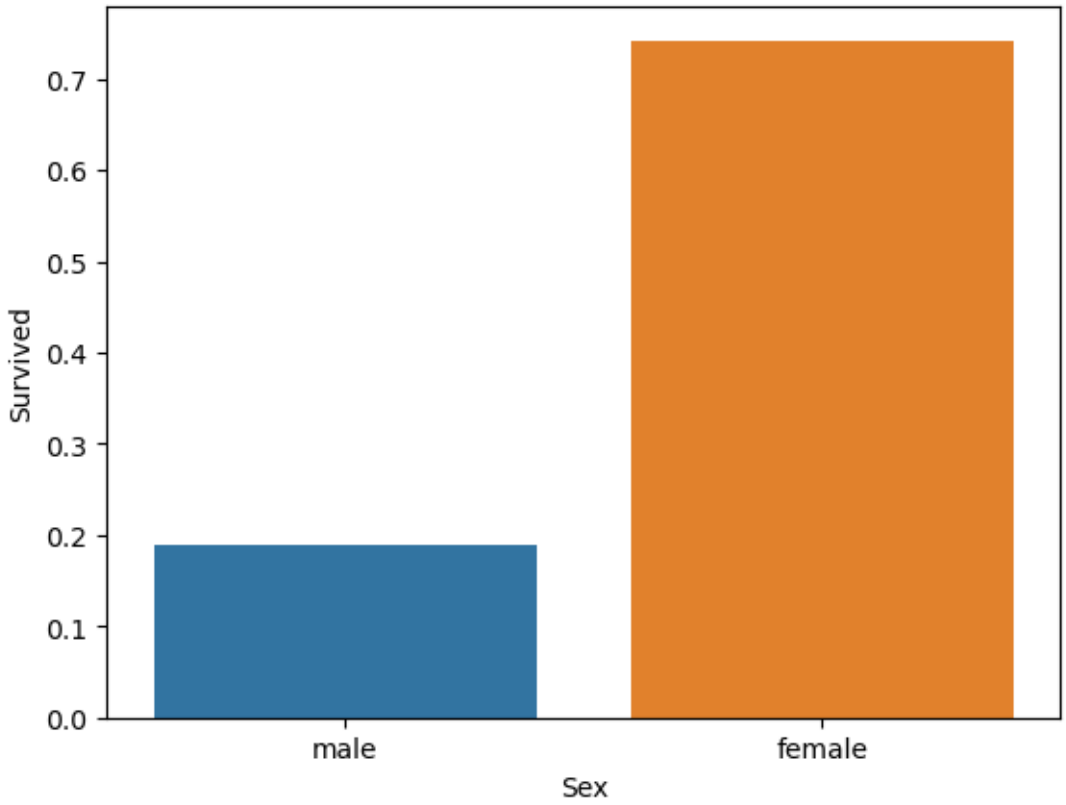
[21]:
```
sns.barplot(x=dataset["Sex"],y=dataset["Survived"],ci=0)
```

```
C:\Users\Vishal Gupta\AppData\Local\Temp\ipykernel_10548\939471306.py:1:
FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=('ci', 0)` for the same effect.

  sns.barplot(x=dataset["Sex"],y=dataset["Survived"],ci=0)
```

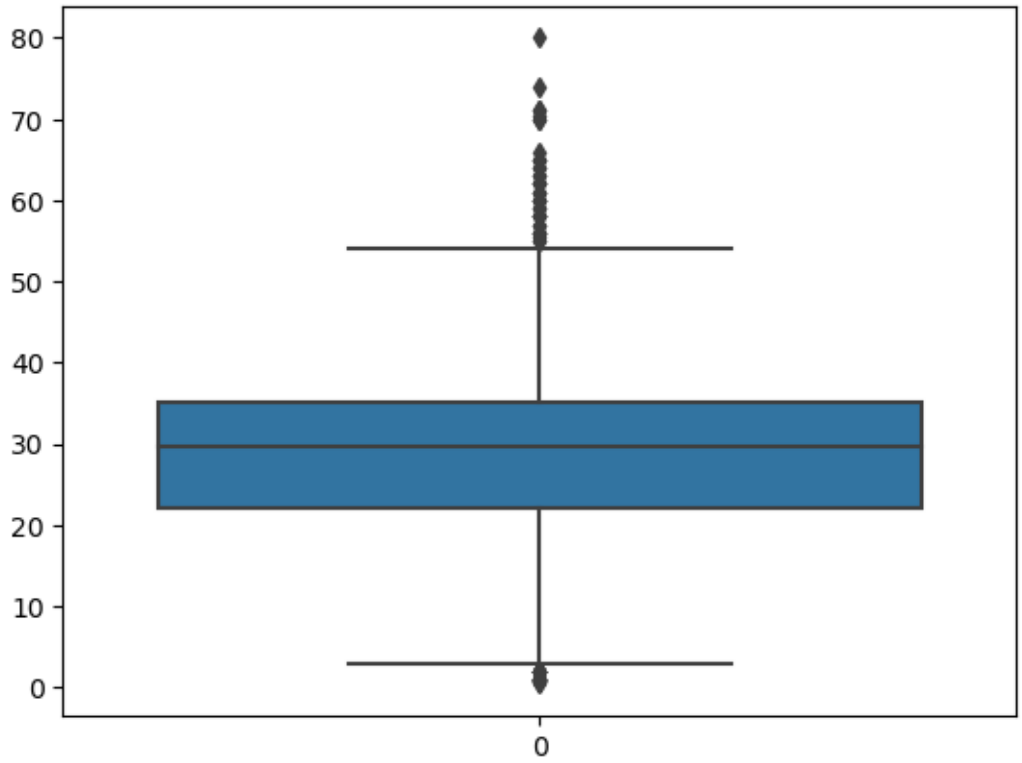
[21]: <Axes: xlabel='Sex', ylabel='Survived'>



5 5. Outlier Detection

[22]:
```
sns.boxplot(dataset.Age)
```

[22]:
```
<Axes: >
```



[23]:
```
q1=dataset.Age.quantile(0.25)
q1
```

[23]: 22.0

[24]:
```
q2=dataset.Age.quantile(0.50)
q2
```

[24]: 29.69911764705882

[25]:
```
q3=dataset.Age.quantile(0.75)
q3
```

[25]: 35.0

[26]:
```
IQR=q3-q1
```

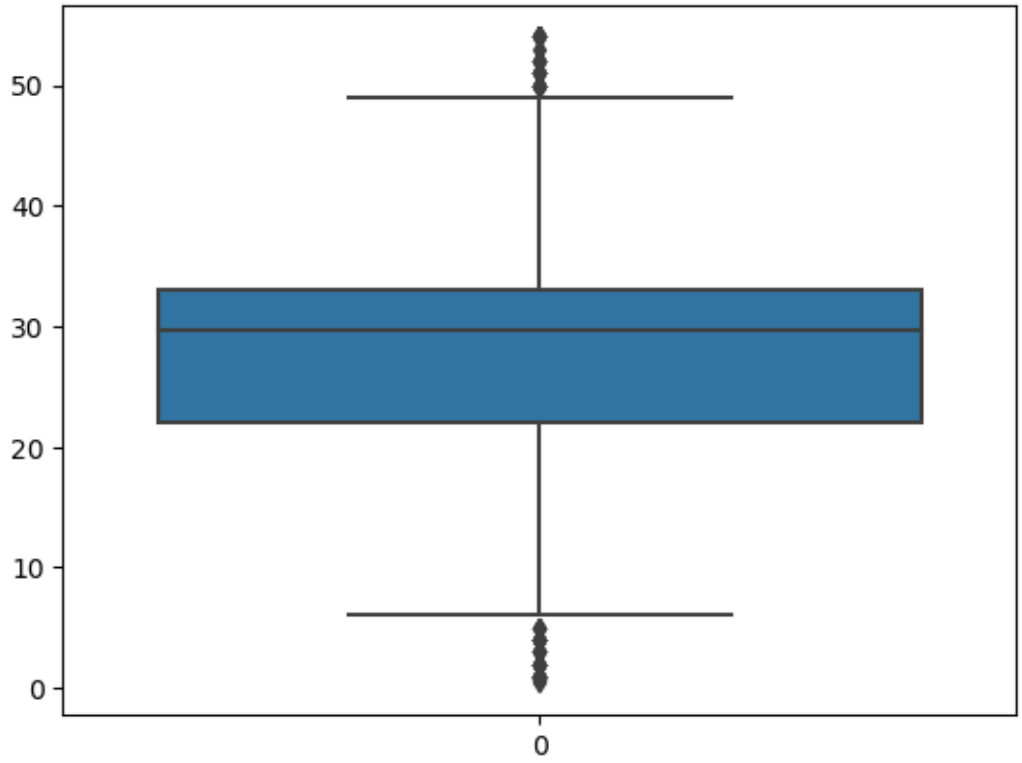
[27]:
```
upper_limit=q3+1.5*IQR
upper_limit
```

[27]: 54.5

[28]:
```
dataset['Age']=np.where(dataset['Age']>upper_limit,30,dataset['Age'])
```

[29]:
```
sns.boxplot(dataset.Age)
```

[29]: <Axes: >



6 6. Split Dependent and Independent variables

[30]:
```
x=dataset.iloc[:,3:13]
y=dataset.iloc[:,1:2]
```

[31]:
```
y.head()
```

[31]:

	Survived
0	0
1	1
2	1
3	1
4	0

[32]:
```
x.head()
```

[32]:
```
                                                Name     Sex   Age  SibSp  \
0                            Braund, Mr. Owen Harris    male  22.0      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th…  female  38.0      1
2                             Heikkinen, Miss. Laina  female  26.0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1
4                           Allen, Mr. William Henry    male  35.0      0

   Parch            Ticket     Fare        Cabin Embarked
0      0         A/5 21171   7.2500  C23 C25 C27        S
1      0          PC 17599  71.2833          C85        C
2      0  STON/O2. 3101282   7.9250  C23 C25 C27        S
3      0            113803  53.1000         C123        S
4      0            373450   8.0500  C23 C25 C27        S
```

[34]:
```
x.shape
```

[34]:
```
(891, 9)
```

[36]:
```
type(x)
```

[36]:
```
pandas.core.frame.DataFrame
```

7 7. Perform Encoding

[38]:
```
from sklearn.preprocessing import LabelEncoder
le=LabelEncoder()
x["Sex"]=le.fit_transform(x["Sex"])
x["Sex"]
```

[38]:
```
0      1
1      0
2      0
3      0
4      1
      ..
886    1
887    0
888    0
889    1
890    1
Name: Sex, Length: 891, dtype: int32
```

[39]:
```
x.head()
```

[39]:
```
                                                Name  Sex   Age  SibSp  Parch  \
0                            Braund, Mr. Owen Harris    1  22.0      1      0
```

```
1  Cumings, Mrs. John Bradley (Florence Briggs Th…    0  38.0      1      0
2                             Heikkinen, Miss. Laina    0  26.0      0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)    0  35.0      1      0
4                           Allen, Mr. William Henry    1  35.0      0      0

             Ticket     Fare        Cabin Embarked
0         A/5 21171   7.2500  C23 C25 C27        S
1          PC 17599  71.2833          C85        C
2  STON/O2. 3101282   7.9250  C23 C25 C27        S
3            113803  53.1000         C123        S
4            373450   8.0500  C23 C25 C27        S
```

[40]:
```
x.Embarked.value_counts()
```

[40]:
```
S    646
C    168
Q     77
Name: Embarked, dtype: int64
```

[41]:
```
Embarked=pd.get_dummies(x["Embarked"],drop_first=True)
Embarked
```

[41]:
```
     Q  S
0    0  1
1    0  0
2    0  1
3    0  1
4    0  1
..  .. ..
886  0  1
887  0  1
888  0  1
889  0  0
890  1  0

[891 rows x 2 columns]
```

[42]:
```
x=pd.concat([x,Embarked],axis=1)
```

[43]:
```
x.head()
```

[43]:
```
                                                Name  Sex   Age  SibSp  Parch  \
0                            Braund, Mr. Owen Harris    1  22.0      1      0
1  Cumings, Mrs. John Bradley (Florence Briggs Th…    0  38.0      1      0
2                             Heikkinen, Miss. Laina    0  26.0      0      0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)    0  35.0      1      0
4                           Allen, Mr. William Henry    1  35.0      0      0
```

```
             Ticket     Fare        Cabin Embarked  Q  S
0         A/5 21171   7.2500  C23 C25 C27        S  0  1
1          PC 17599  71.2833          C85        C  0  0
2  STON/O2. 3101282   7.9250  C23 C25 C27        S  0  1
3            113803  53.1000         C123        S  0  1
4            373450   8.0500  C23 C25 C27        S  0  1
```

[44]:
```
x.drop(["Embarked"],axis=1,inplace=True)
x.head(6)
```

[44]:
```
                                                 Name  Sex        Age  SibSp  \
0                             Braund, Mr. Owen Harris    1  22.000000      1
1  Cumings, Mrs. John Bradley (Florence Briggs Th…    0  38.000000      1
2                              Heikkinen, Miss. Laina    0  26.000000      0
3        Futrelle, Mrs. Jacques Heath (Lily May Peel)    0  35.000000      1
4                            Allen, Mr. William Henry    1  35.000000      0
5                                    Moran, Mr. James    1  29.699118      0

   Parch            Ticket     Fare        Cabin  Q  S
0      0         A/5 21171   7.2500  C23 C25 C27  0  1
1      0          PC 17599  71.2833          C85  0  0
2      0  STON/O2. 3101282   7.9250  C23 C25 C27  0  1
3      0            113803  53.1000         C123  0  1
4      0            373450   8.0500  C23 C25 C27  0  1
5      0            330877   8.4583  C23 C25 C27  1  0
```

8 8. Feature Scaling

[50]:
```
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
y_train=sc.fit_transform(y_train)
y_test=sc.fit_transform(y_test)
y_train
```

[50]:
```
array([[-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
```

```
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
```

```
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
```

```
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
```

```
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
```

```
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
```

```
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453]])
```

[51]:
```
from sklearn.preprocessing import StandardScaler
sc=StandardScaler()
x_train=sc.fit_transform(y_train)
x_test=sc.fit_transform(y_test)
x_train
```

[51]:
```
array([[-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
```

```
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
```

```
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
```

```
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
```

```
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
```

```
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[ 1.21902975],
[-0.82032453],
[-0.82032453],
[-0.82032453],
[ 1.21902975],
[ 1.21902975],
```

```
       [-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [-0.82032453],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [ 1.21902975],
       [-0.82032453],
       [ 1.21902975],
       [-0.82032453]])
```

9 9. Splitting Data into Train and Tes

[48]:
```
from sklearn.model_selection import train_test_split
```

[49]:
```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.5,random_state=0)
x_train.shape,x_test.shape,y_train.shape,y_test.shape
```

[49]: ((445, 10), (446, 10), (445, 1), (446, 1))