

```
[In [1]: ]#Importing the Libraries
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import StandardScaler
from sklearn.model_selection import train_test_split

[In [52]: ]pip install nbconvert

Requirement already satisfied: nbconvert in c:\users\nvthin\anaconda3\lib\site-packages (6.5.4)
Requirement already satisfied: lxml in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (4.9.2)
Requirement already satisfied: beautifulsoup4 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (4.12.2)
Requirement already satisfied: bleach in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (4.1.0)
Requirement already satisfied: defusedxml in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (0.7.1)
Requirement already satisfied: entrypoint>=0.2.2 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (0.4)
Requirement already satisfied: Jinja2>=3.0 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (3.1.2)
Requirement already satisfied: jupyter-core<4.7 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (5.3.0)
Requirement already satisfied: jupyterlab-pygments in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (0.1.2)
Requirement already satisfied: MarkupSafe<2.0 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (2.1.1)
Requirement already satisfied: mistune<2,>=0.8.1 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (0.8.4)
Requirement already satisfied: nbclient<0.5.0 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (0.5.13)
Requirement already satisfied: nbformat<5.1 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (5.7.0)
Requirement already satisfied: packaging in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (23.0)
Requirement already satisfied: pandocfilters<=4.1.1 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (1.5.0)
Requirement already satisfied: pygments<2.4.1 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (2.15.1)
Requirement already satisfied: tinycss2 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (1.2.1)
Requirement already satisfied: traitlets<=0 in c:\users\nvthin\anaconda3\lib\site-packages (from nbconvert) (5.7.1)
Requirement already satisfied: platformdirs<=7.5 in c:\users\nvthin\anaconda3\lib\site-packages (from jupyter-core<4.7->nbconvert) (2.5.2)
Requirement already satisfied: pywin32<300 in c:\users\nvthin\anaconda3\lib\site-packages (from jupyter-core<4.7->nbconvert) (305.1)
Requirement already satisfied: bleach in c:\users\nvthin\anaconda3\lib\site-packages (from nbclient<0.5.0->nbconvert) (4.1.0)
Requirement already satisfied: nest-asyncio in c:\users\nvthin\anaconda3\lib\site-packages (from nbclient<0.5.0->nbconvert) (1.5.6)
Requirement already satisfied: fastjsonschema in c:\users\nvthin\anaconda3\lib\site-packages (from nbformat<5.1->nbconvert) (2.16.2)
Requirement already satisfied: jsonschema<2.6 in c:\users\nvthin\anaconda3\lib\site-packages (from nbformat<5.1->nbconvert) (4.17.3)
Requirement already satisfied: soupsieve<1.2 in c:\users\nvthin\anaconda3\lib\site-packages (from beautifulsoup4->nbconvert) (2.4)
Requirement already satisfied: nbclient<0.5.0 in c:\users\nvthin\anaconda3\lib\site-packages (from bleach->nbconvert) (0.5.13)
Requirement already satisfied: webencodings in c:\users\nvthin\anaconda3\lib\site-packages (from bleach->nbconvert) (0.5.1)
Requirement already satisfied: attrs<=17.4.0 in c:\users\nvthin\anaconda3\lib\site-packages (from jsonschema<2.6->nbformat<5.1->nbconvert) (22.1.0)
Requirement already satisfied: pyrsistent<0.17.0,!=0.17.1,!=0.17.2,>=0.14.0 in c:\users\nvthin\anaconda3\lib\site-packages (from jsonschema<2.6->nbformat<5.1->nbconvert) (0.18.0)
Requirement already satisfied: python-dateutil<=2.8.2 in c:\users\nvthin\anaconda3\lib\site-packages (from jupyter-client<6.1.5->nbclient<0.5.0->nbconvert) (2.8.2)
Requirement already satisfied: pyzmq<23.0 in c:\users\nvthin\anaconda3\lib\site-packages (from jupyter-client<6.1.5->nbclient<0.5.0->nbconvert) (23.2.0)
Requirement already satisfied: tornado<6.2 in c:\users\nvthin\anaconda3\lib\site-packages (from jupyter-client<6.1.5->nbclient<0.5.0->nbconvert) (6.3.3)
Note: you may need to restart the kernel to use updated packages.

[In [3]: ]df=pd.read_csv("Titanic-Dataset.csv")

[In [4]: ]df

Out[4]:
   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  Furetle, 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
...      ...      ...      ...      ...  ...  ...      ...      ...      ...      ...      ...      ...
886           887         0      2   Montvila, Rev. Juozas   male  27.0      0      0      211536  13.0000   NaN      S
887           888         1      1   Graham, Miss. Margaret Edith  female  19.0      0      0      112053  30.0000   B42      S
888           889         0      3   Johnson, Miss. Catherine Helen "Carnie"  female  NaN      1      2      W/C 6607  23.4500   NaN      S
889           890         1      1   Behr, Mr. Karl Howell   male  26.0      0      0      111369  30.0000   C148      C
890           891         0      3   Dooley, Mr. Patrick   male  32.0      0      0      370376   7.7500   NaN      Q
891 rows x 12 columns

[In [5]: ]df.head()

Out[5]:
   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  Furetle, 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 [6]: ]df.tail()

Out[6]:
   PassengerId  Survived  Pclass    Name  Sex  Age  SibSp  Parch    Ticket   Fare  Cabin  Embarked
886           887         0      2   Montvila, Rev. Juozas   male  27.0      0      0      211536  13.0000   NaN      S
887           888         1      1   Graham, Miss. Margaret Edith  female  19.0      0      0      112053  30.0000   B42      S
888           889         0      3   Johnson, Miss. Catherine Helen "Carnie"  female  NaN      1      2      W/C 6607  23.4500   NaN      S
889           890         1      1   Behr, Mr. Karl Howell   male  26.0      0      0      111369  30.0000   C148      C
890           891         0      3   Dooley, Mr. Patrick   male  32.0      0      0      370376   7.7500   NaN      Q

[In [8]: ]df.shape

Out[8]:
(891, 12)

[In [9]: ]df.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    object
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      294 non-null    object
11 Embarked   889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.74 KB

[In [10]: ]df.describe()

Out[10]:
   PassengerId  Survived  Pclass    Age  SibSp  Parch    Fare
count  891.000000  891.000000  891.000000  714.000000  891.000000  891.000000
mean   446.000000  0.383838  2.308642  29.699118  0.523008  0.381594  32.204208
std    257.353842  0.486592  0.836071  14.526487  1.102743  0.806507  49.693429
min      1.000000  0.000000  1.000000  0.420000  0.000000  0.000000  0.000000
25%    223.500000  0.000000  2.000000  20.125000  0.000000  0.000000  7.910400
50%    446.000000  0.000000  3.000000  28.000000  0.000000  0.000000  14.454200
75%    668.500000  1.000000  38.000000  38.000000  1.000000  0.000000  31.000000
max    891.000000  1.000000  3.000000  80.000000  8.000000  6.000000  512.329000

[In [11]: ]corr=df.corr()
corr
corr

C:\Users\NVTHIN\AppData\Local\Temp\ipykernel_34436\3182148910.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.
  df.corr()

Out[11]:
   PassengerId  Survived  Pclass    Age  SibSp  Parch    Fare
PassengerId  1.000000  -0.050007  -0.035144  0.036847  -0.067527  -0.001652  0.012658
Survived      -0.050007  1.000000  -0.338481  -0.077221  -0.093232  0.081629  0.257307
Pclass        -0.035144  -0.338481  1.000000  -0.369226  -0.030261  0.018443  -0.549500
Age           0.036847  -0.077221  -0.369226  1.000000  -0.308247  -0.198119  0.090607
SibSp         -0.057527  -0.035232  0.083081  -0.308247  1.000000  0.414838  0.096601
Parch         -0.001652  0.081629  0.018443  -0.198119  0.414838  1.000000  0.216225
Fare          0.012658  0.257307  -0.549500  0.090607  0.159651  0.216225  1.000000

[In [13]: ]plt.subplots(figsize=(15,10))
sns.heatmap(corr,annot=True)

Out[13]:
<Axes: >

PassengerId  Survived  Pclass    Age  SibSp  Parch    Fare
1            1         0      3   Braund, Mr. Owen Harris   male  22.0      1      0      A/5 21171   7.2500   NaN      S
2            2         1      1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0      1      0      PC 17599  71.2833   C85      C
3            3         1      3   Heikkinen, Miss. Laina   female  26.0      0      0      STON/O2 3101282   7.9250   NaN      S
4            4         1      1  Furetle, Mrs. Jacques Heath (Lily May Peel)  female  35.0      1      0      113803  53.1000   C123      S
5            5         0      3      Allen, Mr. William Henry   male  35.0      0      0      373450   8.0500   NaN      S
...      ...      ...      ...      ...  ...  ...      ...      ...      ...      ...      ...      ...
886           887         0      2   Montvila, Rev. Juozas   male  27.0      0      0      211536  13.0000   NaN      S
887           888         1      1   Graham, Miss. Margaret Edith  female  19.0      0      0      112053  30.0000   B42      S
888           889         0      3   Johnson, Miss. Catherine Helen "Carnie"  female  NaN      1      2      W/C 6607  23.4500   NaN      S
889           890         1      1   Behr, Mr. Karl Howell   male  26.0      0      0      111369  30.0000   C148      C
890           891         0      3   Dooley, Mr. Patrick   male  32.0      0      0      370376   7.7500   NaN      Q

[In [15]: ]df.isnull().sum()

Out[15]:
PassengerId    0
Survived        0
Pclass          0
Sex             0
Age            177
SibSp           0
Parch           0
Ticket         0
Fare            0
Cabin         687
Embarked        2
dtype: int64

[In [16]: ]df.describe()

Out[16]:
   PassengerId  Survived  Pclass    Age  SibSp  Parch    Fare
count  891.000000  891.000000  891.000000  714.000000  891.000000  891.000000
mean   446.000000  0.383838  2.308642  29.699118  0.523008  0.381594  32.204208
std    257.353842  0.486592  0.836071  14.526487  1.102743  0.806507  49.693429
min      1.000000  0.000000  1.000000  0.420000  0.000000  0.000000  0.000000
25%    223.500000  0.000000  2.000000  20.125000  0.000000  0.000000  7.910400
50%    446.000000  0.000000  3.000000  28.000000  0.000000  0.000000  14.454200
75%    668.500000  1.000000  38.000000  38.000000  1.000000  0.000000  31.000000
max    891.000000  1.000000  3.000000  80.000000  8.000000  6.000000  512.329000

[In [17]: ]df.isnull().any()

Out[17]:
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

[In [18]: ]df.isnull().sum()

Out[18]:
PassengerId    0
Survived        0
Pclass          0
Sex             0
Age            177
SibSp           0
Parch           0
Ticket         0
Fare            0
Cabin         687
Embarked        2
dtype: int64

[In [19]: ]df.describe()

Out[19]:
   PassengerId  Survived  Pclass    Age  SibSp  Parch    Fare
count  891.000000  891.000000  891.000000  714.000000  891.000000  891.000000
mean   446.000000  0.383838  2.308642  29.699118  0.523008  0.381594  32.204208
std    257.353842  0.486592  0.836071  14.526487  1.102743  0.806507  49.693429
min      1.000000  0.000000  1.000000  0.420000  0.000000  0.000000  0.000000
25%    223.500000  0.000000  2.000000  20.125000  0.000000  0.000000  7.910400
50%    446.000000  0.000000  3.000000  28.000000  0.000000  0.000000  14.454200
75%    668.500000  1.000000  38.000000  38.000000  1.000000  0.000000  31.000000
max    891.000000  1.000000  3.000000  80.000000  8.000000  6.000000  512.329000

[In [20]: ]sns.countplot(data=df, x='Survived')
plt.title('Survival Count')
plt.xlabel('Survived')
plt.ylabel('Count')
plt.show()

Survival Count

[In [21]: ]sns.histplot(data=df, x='Age', bins=20, kde=True)
plt.title('Age Distribution')
plt.xlabel('Age')
plt.ylabel('Count')
plt.show()

Age Distribution

[In [22]: ]sns.histplot(data=df, x='Fare', bins=20, kde=True)
plt.title('Fare Distribution')
plt.xlabel('Fare')
plt.ylabel('Count')
plt.show()

Fare Distribution

[In [23]: ]sns.pairplot(data=df[['Fare', 'SibSp', 'Parch']])
plt.title('Pair Plot')
plt.show()

Pair Plot

[In [24]: ]corr_matrix = df.corr()
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')
plt.show()

Correlation Heatmap

[In [25]: ]sns.boxplot(df_cleaned)

Out[25]:
<Axes: >

df_cleaned

[In [26]: ]df.head()

Out[26]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [27]: ]sns.boxplot(df_cleaned)

Out[27]:
<Axes: >

df_cleaned

[In [28]: ]df.head()

Out[28]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [29]: ]sns.boxplot(df_cleaned)

Out[29]:
<Axes: >

df_cleaned

[In [30]: ]df.head()

Out[30]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [31]: ]sns.boxplot(df_cleaned)

Out[31]:
<Axes: >

df_cleaned

[In [32]: ]df.head()

Out[32]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [33]: ]sns.boxplot(df_cleaned)

Out[33]:
<Axes: >

df_cleaned

[In [34]: ]df.head()

Out[34]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [35]: ]sns.boxplot(df_cleaned)

Out[35]:
<Axes: >

df_cleaned

[In [36]: ]df.head()

Out[36]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [37]: ]sns.boxplot(df_cleaned)

Out[37]:
<Axes: >

df_cleaned

[In [38]: ]df.head()

Out[38]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [39]: ]sns.boxplot(df_cleaned)

Out[39]:
<Axes: >

df_cleaned

[In [40]: ]df.head()

Out[40]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [41]: ]sns.boxplot(df_cleaned)

Out[41]:
<Axes: >

df_cleaned

[In [42]: ]df.head()

Out[42]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [43]: ]sns.boxplot(df_cleaned)

Out[43]:
<Axes: >

df_cleaned

[In [44]: ]df.head()

Out[44]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500      S
1            2         1   female  38.000000      1      0      71.2500      S
2            3         1   female  26.000000      0      0      7.9250      S
3            4         1   female  35.000000      1      0      53.1000      S
4            5         3   male  35.000000      0      0      8.0500      S
5            6         3   male  29.699118      0      0      8.4583      Q
...      ...      ...      ...      ...      ...      ...      ...
886           887         2   male  27.000000      0      0      13.0000      S
887           888         1   female  19.000000      0      0      30.0000      S
888           889         1   male  29.699118      1      2      23.4500      C
889           890         1   male  26.000000      0      0      30.0000      C
890           891         3   male  32.000000      0      0      7.7500      Q
775 rows x 9 columns

[In [45]: ]sns.boxplot(df_cleaned)

Out[45]:
<Axes: >

df_cleaned

[In [46]: ]df.head()

Out[46]:
   PassengerId  Pclass  Sex  Age  SibSp  Parch  Fare  Embarked
0            1         3   male  22.000000      1      0      7.2500     
```