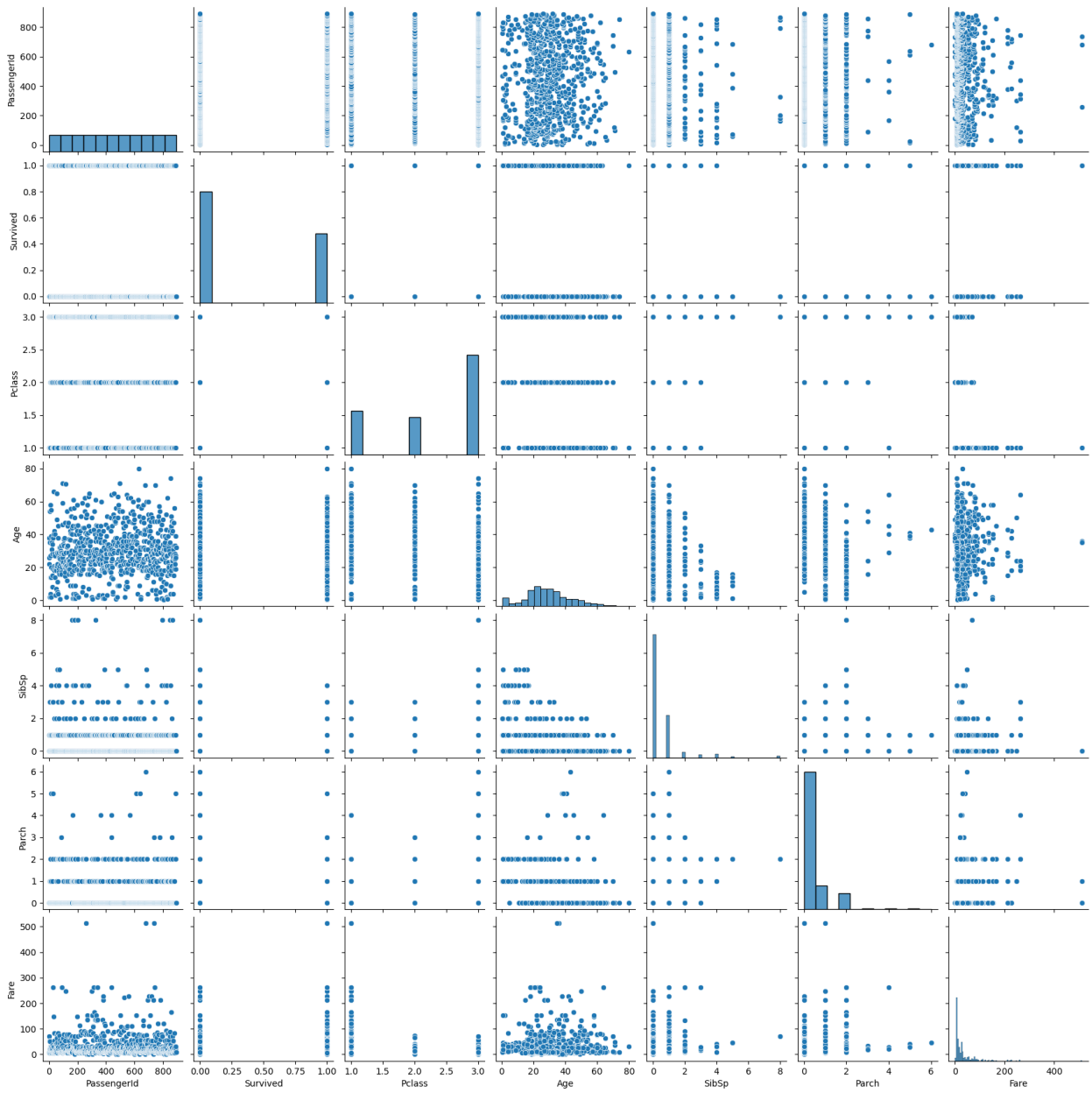
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

titanic= pd.read_csv(r"C:\Users\SUJITHA MANDAVILLI\Downloads\
train.csv")

titanic.describe()
```

```
       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
```

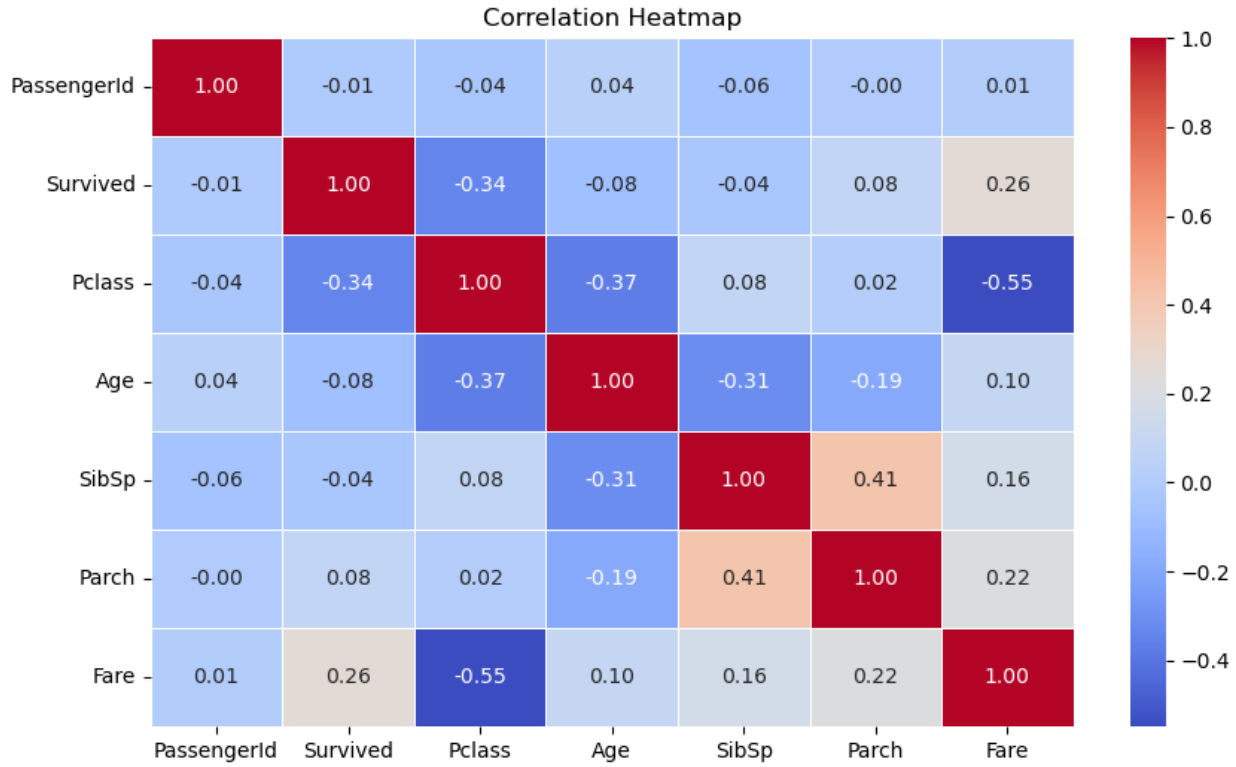
```
titanic.info()
```

```
<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
```

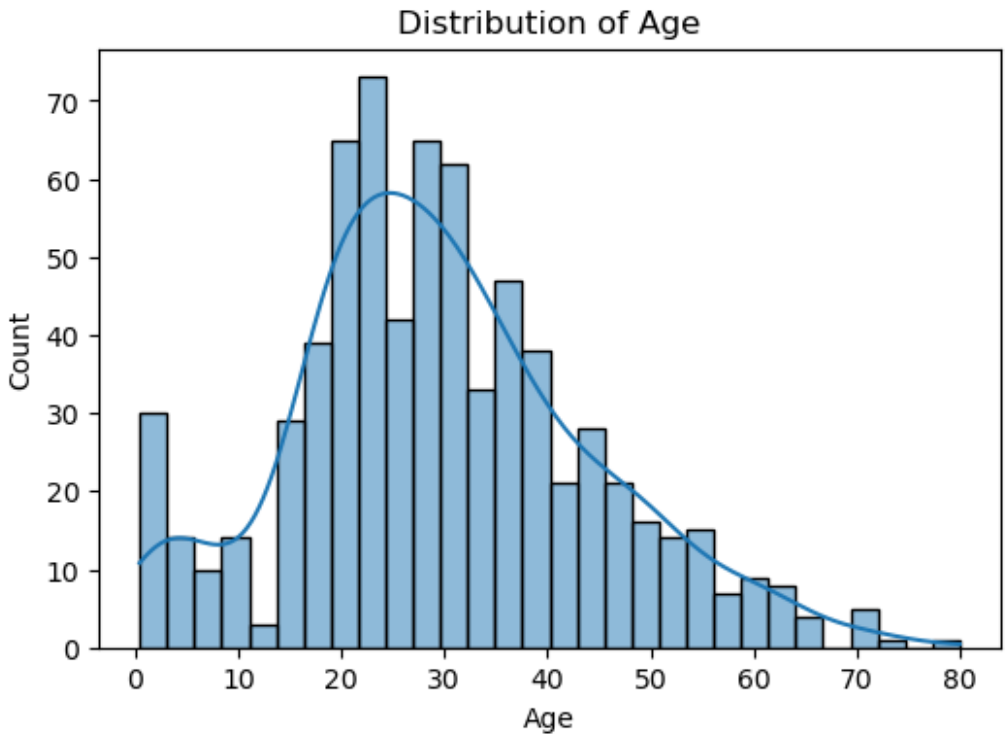
```
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB

titanic.value_counts()

PassengerId  Survived  Pclass  Name
Sex     Age   SibSp  Parch  Ticket    Fare      Cabin  Embarked
2            1         1       Cumings, Mrs. John Bradley (Florence
Briggs Thayer)   female  38.0  1      0      PC 17599  71.2833   C85
C           1
572          1         1       Appleton, Mrs. Edward Dale (Charlotte
Lamson)          female  53.0  2      0      11769     51.4792   C101
S           1
578          1         1       Silvey, Mrs. William Baird (Alice
Munger)             female  39.0  1      0      13507     55.9000
E44    S           1
582          1         1       Thayer, Mrs. John Borland (Marian
Longstreth Morris)  female  39.0  1      1      17421     110.8833
C68    C           1
584          0         1       Ross, Mr. John Hugo
male    36.0  0      0      13049     40.1250   A10    C            1

..
328          1         2       Ball, Mrs. (Ada E Hall)
female  36.0  0      0      28551     13.0000   D      S            1
330          1         1       Hippach, Miss. Jean Gertrude
female  16.0  0      1      111361    57.9792   B18    C            1
332          0         1       Partner, Mr. Austen
male    45.5  0      0      113043    28.5000   C124   S            1
333          0         1       Graham, Mr. George Edward
male    38.0  0      1      PC 17582  153.4625  C91    S            1
890          1         1       Behr, Mr. Karl Howell
male    26.0  0      0      111369    30.0000   C148   C            1
Name: count, Length: 183, dtype: int64

titplot=sns.pairplot(titanic)
```



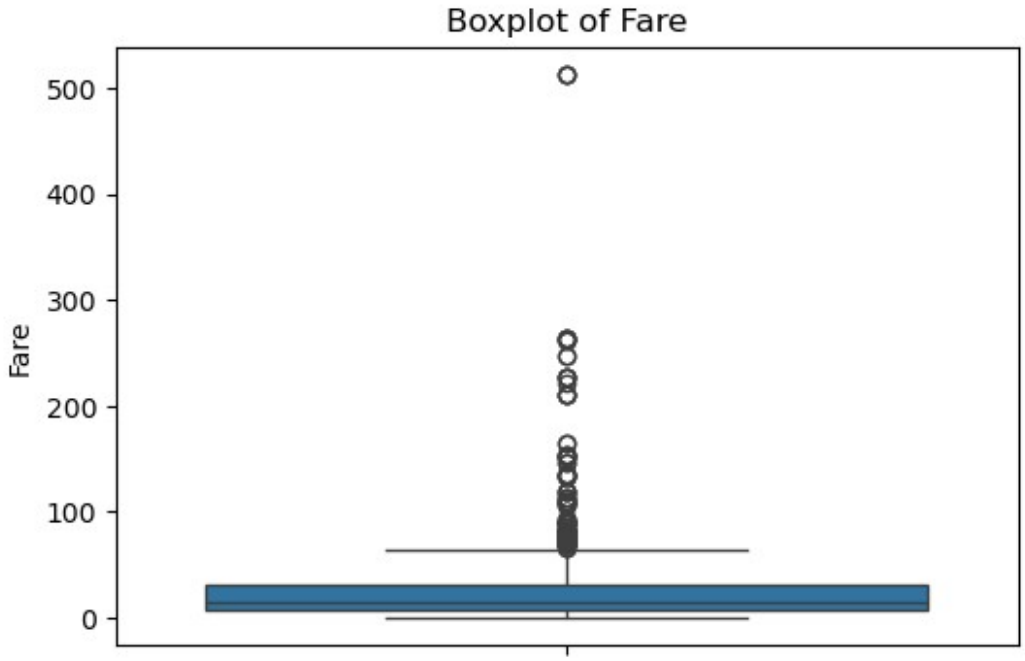

```
numeric_df = titanic.select_dtypes(include=['int64', 'float64'])

plt.figure(figsize=(10, 6))
sns.heatmap(numeric_df.corr(), annot=True, fmt=".2f", cmap="coolwarm",
linewidths=0.5)
plt.title("Correlation Heatmap")
plt.show()
```




```
plt.figure(figsize=(6,4))
sns.histplot(data=titanic, x='Age', bins=30, kde=True)
plt.title('Distribution of Age')
plt.show()
```




```
plt.figure(figsize=(6,4))
sns.boxplot(data=titanic, y='Fare')
plt.title('Boxplot of Fare')
plt.show()
```



```
plt.figure(figsize=(6,4))
sns.scatterplot(data=titanic, x='Age', y='Fare')
plt.title('Age vs Fare')
plt.show()
```