```python
In [1]:  # Import libraries
         import pandas as pd
         import numpy as np
         import matplotlib.pyplot as plt
         import seaborn as sns
```

C:\Users\UDAY\AppData\Roaming\Python\Python39\site-packages\matplotlib\projecti
ons__init__.py:63: UserWarning: Unable to import Axes3D. This may be due to mu
ltiple versions of Matplotlib being installed (e.g. as a system package and as
a pip package). As a result, the 3D projection is not available.
  warnings.warn("Unable to import Axes3D. This may be due to multiple versions
of "

```python
In [2]:  # Load dataset
         df = pd.read_csv("tested.csv")
```

```python
In [3]:  df.head()
```

Out[3]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	892	0	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN
1	893	1	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN
2	894	0	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN
3	895	0	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN
4	896	1	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN

```python
# 1. Basic Information

print("===== INFO =====")
print(df.info())
```

```
===== INFO =====
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  418 non-null    int64
 1   Survived     418 non-null    int64
 2   Pclass       418 non-null    int64
 3   Name         418 non-null    object
 4   Sex          418 non-null    object
 5   Age          332 non-null    float64
 6   SibSp        418 non-null    int64
 7   Parch        418 non-null    int64
 8   Ticket       418 non-null    object
 9   Fare         417 non-null    float64
 10  Cabin        91 non-null     object
 11  Embarked     418 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 39.3+ KB
None
```

```python
print("\n===== DESCRIBE =====")

print(df.describe())
```
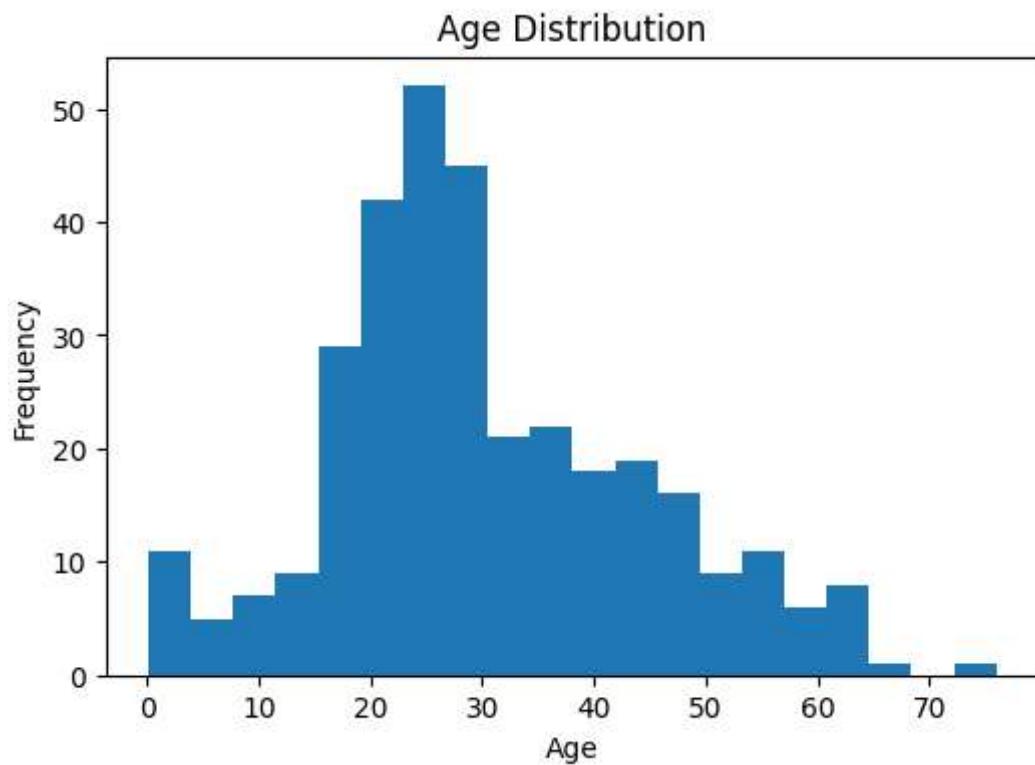
```
===== DESCRIBE =====
       PassengerId    Survived      Pclass         Age       SibSp  \
count   418.000000  418.000000  418.000000  332.000000  418.000000
mean   1100.500000    0.363636    2.265550   30.272590    0.447368
std     120.810458    0.481622    0.841838   14.181209    0.896760
min     892.000000    0.000000    1.000000    0.170000    0.000000
25%     996.250000    0.000000    1.000000   21.000000    0.000000
50%    1100.500000    0.000000    3.000000   27.000000    0.000000
75%    1204.750000    1.000000    3.000000   39.000000    1.000000
max    1309.000000    1.000000    3.000000   76.000000    8.000000

            Parch        Fare
count  418.000000  417.000000
mean     0.392344   35.627188
std      0.981429   55.907576
min      0.000000    0.000000
25%      0.000000    7.895800
50%      0.000000   14.454200
75%      0.000000   31.500000
max      9.000000  512.329200
```

```
In [6]: print("\n===== VALUE COUNTS =====")

        print("Sex:\n", df['Sex'].value_counts())
        print("\nPclass:\n", df['Pclass'].value_counts())
        print("\nEmbarked:\n", df['Embarked'].value_counts())
        print("\nSurvived:\n", df['Survived'].value_counts())
```
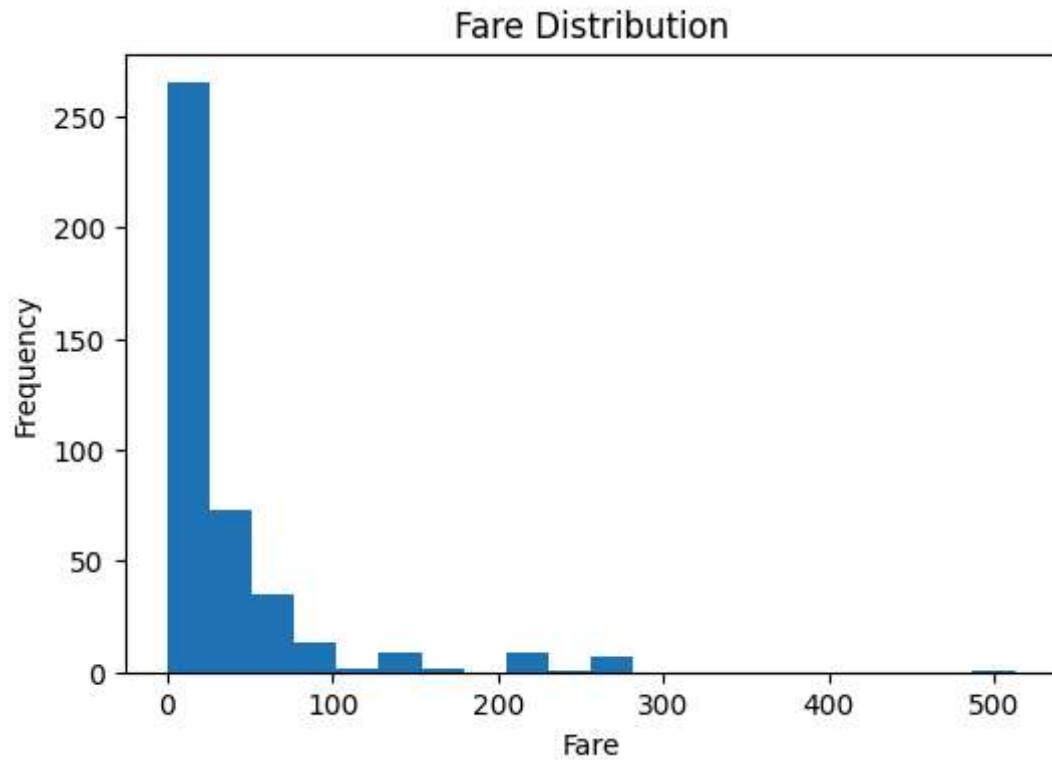
```
===== VALUE COUNTS =====
Sex:
 male      266
female    152
Name: Sex, dtype: int64

Pclass:
 3    218
1    107
2     93
Name: Pclass, dtype: int64

Embarked:
 S    270
C    102
Q     46
Name: Embarked, dtype: int64

Survived:
 0    266
1    152
Name: Survived, dtype: int64
```

In [7]:
```python
# Histograms

plt.figure(figsize=(6,4))
plt.hist(df['Age'].dropna(), bins=20)
plt.title("Age Distribution")
plt.xlabel("Age")
plt.ylabel("Frequency")
plt.show()
```
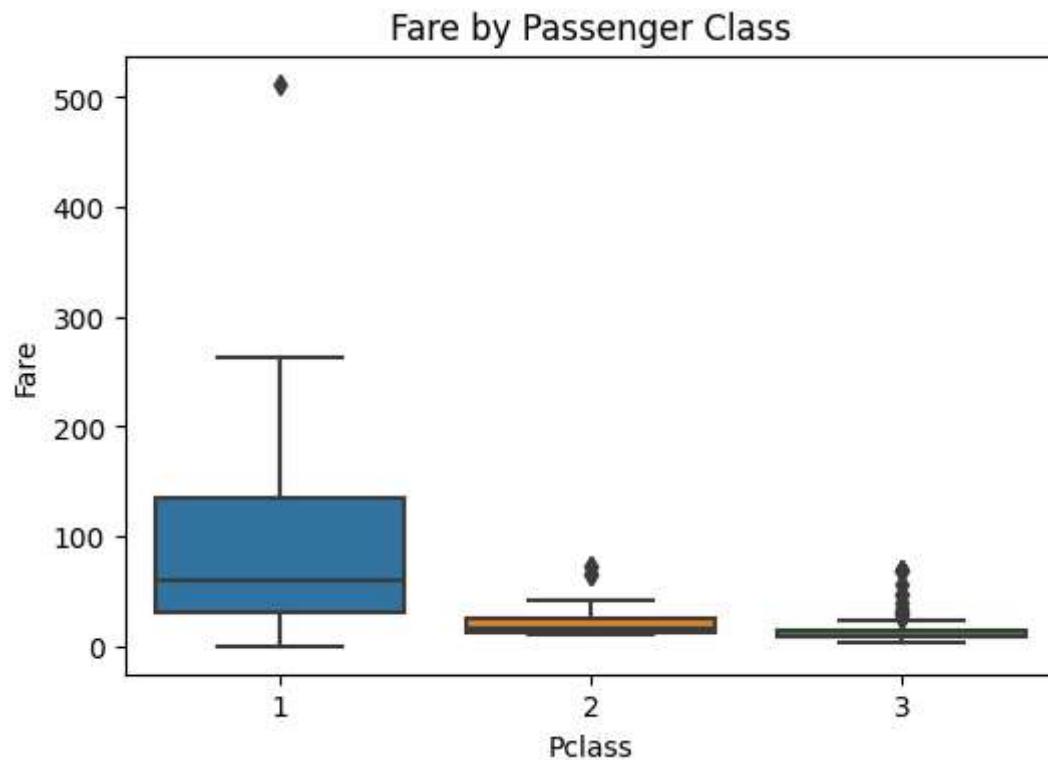
Age Distribution

```
In [8]: plt.figure(figsize=(6,4))
         plt.hist(df['Fare'].dropna(), bins=20)
         plt.title("Fare Distribution")
         plt.xlabel("Fare")
         plt.ylabel("Frequency")
         plt.show()
```



Fare Distribution
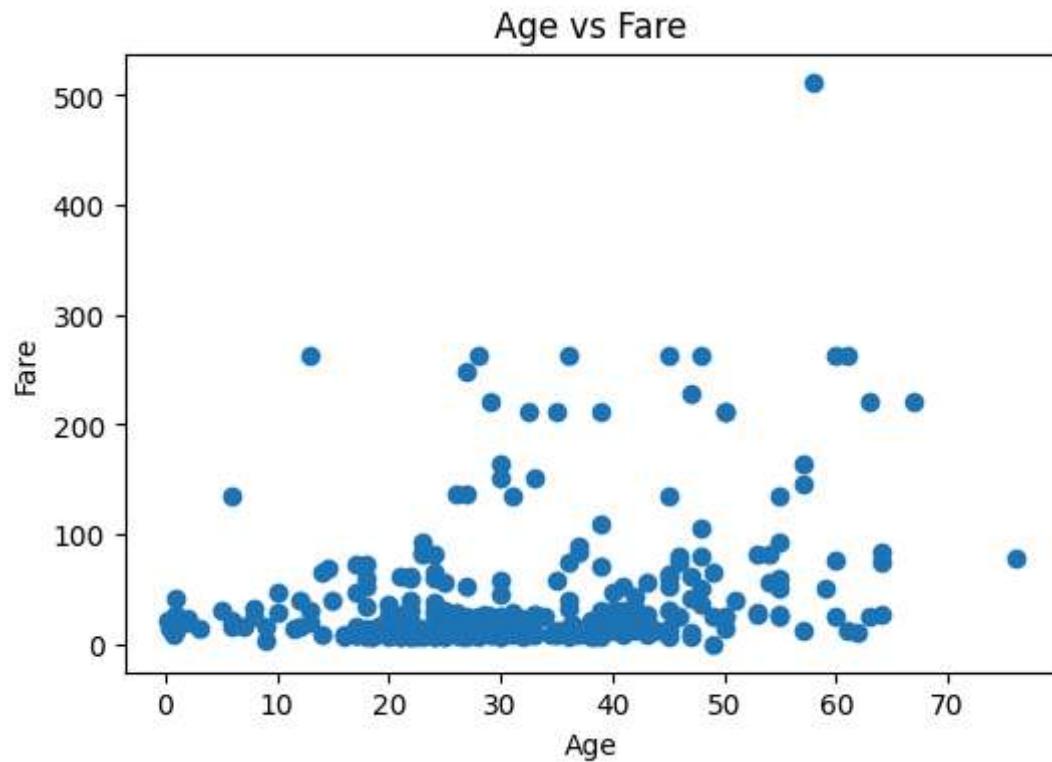
In [9]:
```python
# Boxplot

plt.figure(figsize=(6,4))
sns.boxplot(x="Pclass", y="Fare", data=df)
plt.title("Fare by Passenger Class")
plt.show()
```

In [10]:
```python
# Scatter Plot

plt.figure(figsize=(6,4))
plt.scatter(df['Age'], df['Fare'])
plt.title("Age vs Fare")
plt.xlabel("Age")
plt.ylabel("Fare")
plt.show()
```


Age vs Fare
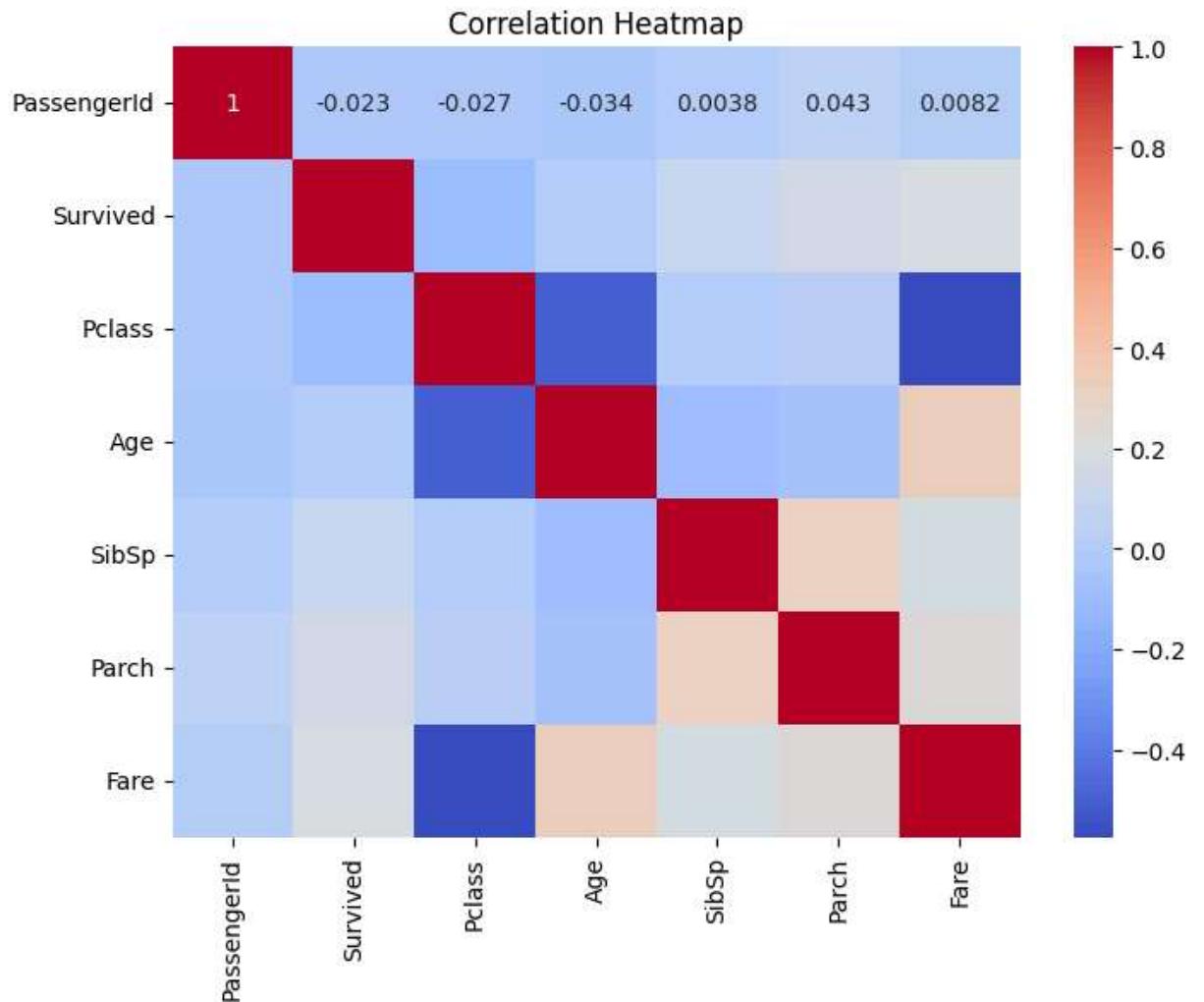
```python
# 6. Correlation Heatmap

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

# Compute correlation
corr = numeric_df.corr()

# Plot heatmap
plt.figure(figsize=(8,6))
sns.heatmap(corr, annot=True, cmap="coolwarm")
plt.title("Correlation Heatmap")
plt.show()
```
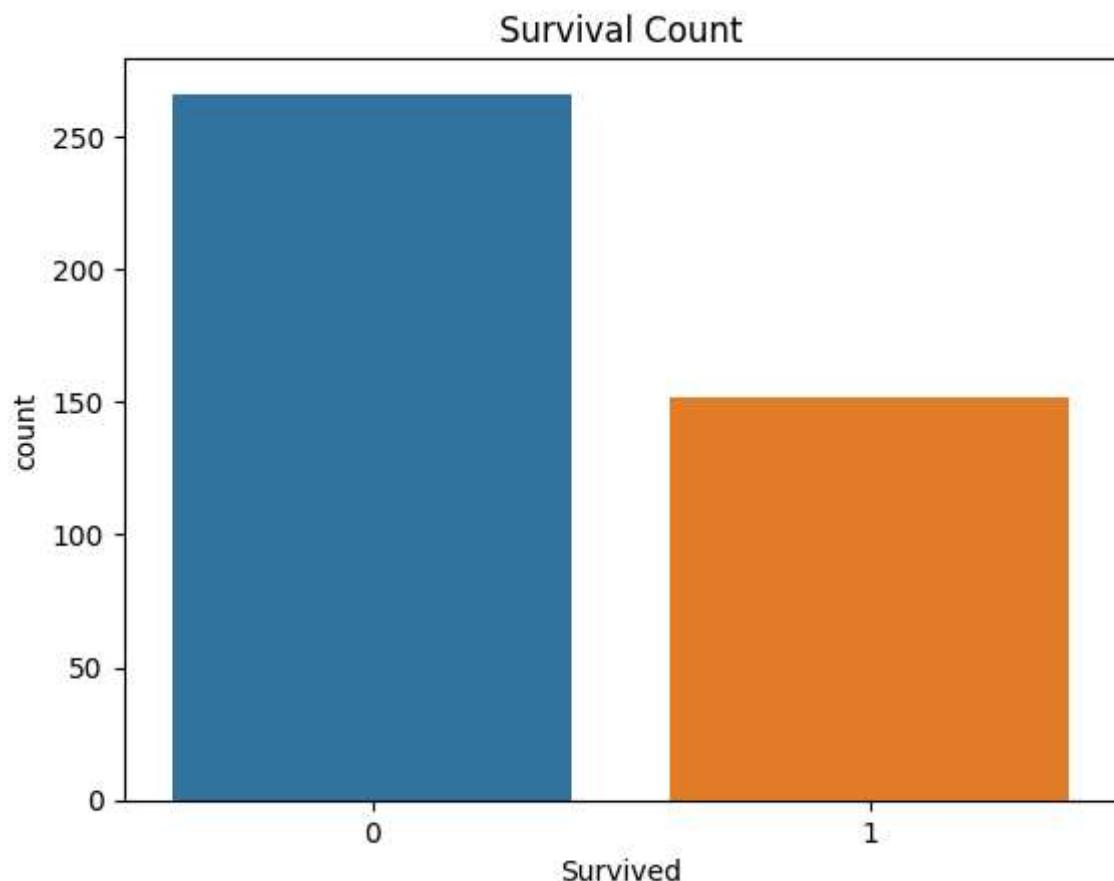
In [16]:
```python
# 7. Survival Count Plot

sns.countplot(x="Survived", data=df)
plt.title("Survival Count")
plt.show()
```

```
In [17]: print("""
         FINAL SUMMARY OF FINDINGS:

         1. Most passengers were aged between 20-40.
         2. Fare distribution is highly skewed.
         3. First-class passengers paid highest fares.
         4. Fare and passenger class have strong relationship.
         5. Weak relationship between Age and Fare.
         6. Majority passengers did not survive.
         7. Most passengers embarked from port 'S'.
         """)
```

```
FINAL SUMMARY OF FINDINGS:

1. Most passengers were aged between 20-40.
2. Fare distribution is highly skewed.
3. First-class passengers paid highest fares.
4. Fare and passenger class have strong relationship.
5. Weak relationship between Age and Fare.
6. Majority passengers did not survive.
7. Most passengers embarked from port 'S'.
```

In [ ]: