Start coding or generate with AI.

⌄ VISUALIZATION ON TITANIC DATASET

The Titanic dataset provides comprehensive information about the passengers who traveled on the RMS Titanic, covering both demographic and travel-related aspects. One primary feature is the **Survived** column,
⌄ which specifies each passenger's fate, with 1 indicating survival and 0 indicating otherwise. The **Pclass** column categorizes passengers by socio-economic status: First, Second, or Third Class. The dataset also records each passenger's **Sex** and **Age**, although some ages are missing.

Family relationships on board are represented by **SibSp** (the count of siblings or spouses) and **Parch** (the count of parents or children). The **Fare** column shows the ticket price paid by each passenger, reflecting economic differences. The **Embarked** column identifies where each passenger boarded the ship: Cherbourg, Queenstown, or Southampton. Additional details include **Name**, **Ticket** (ticket number), and **Cabin**, which indicates cabin assignment (though many entries are missing here). Overall, this dataset offers a valuable resource for examining how variables such as socio-economic class, family composition, age, and gender may have influenced passengers' chances of survival on the Titanic.

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

file_path = "/content/Titanic-Dataset.csv"
```

```
titanic_df = pd.read_csv(file_path)
```

∨ DETAILS INCLUDED IN TITANIC DATASET:

```
print(titanic_df.head())
titanic_df = pd.read_csv(file_path)

print(titanic_df.head())
```

```
      PassengerId  Survived  Pclass  \
0               1         0       3
1               2         1       1
2               3         1       3
3               4         1       1
4               5         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

   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
   PassengerId  Survived  Pclass  \
0               1         0       3
1               2         1       1
2               3         1       3
3               4         1       1
4               5         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

   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
```

⌄ DATA PROCESSING

The missing values for Age and Embarkation (the passengers' boarding location) were filled using the median and mode, respectively, while the Cabin column, which indicates room numbers, was dropped due to a high volume of missing data.
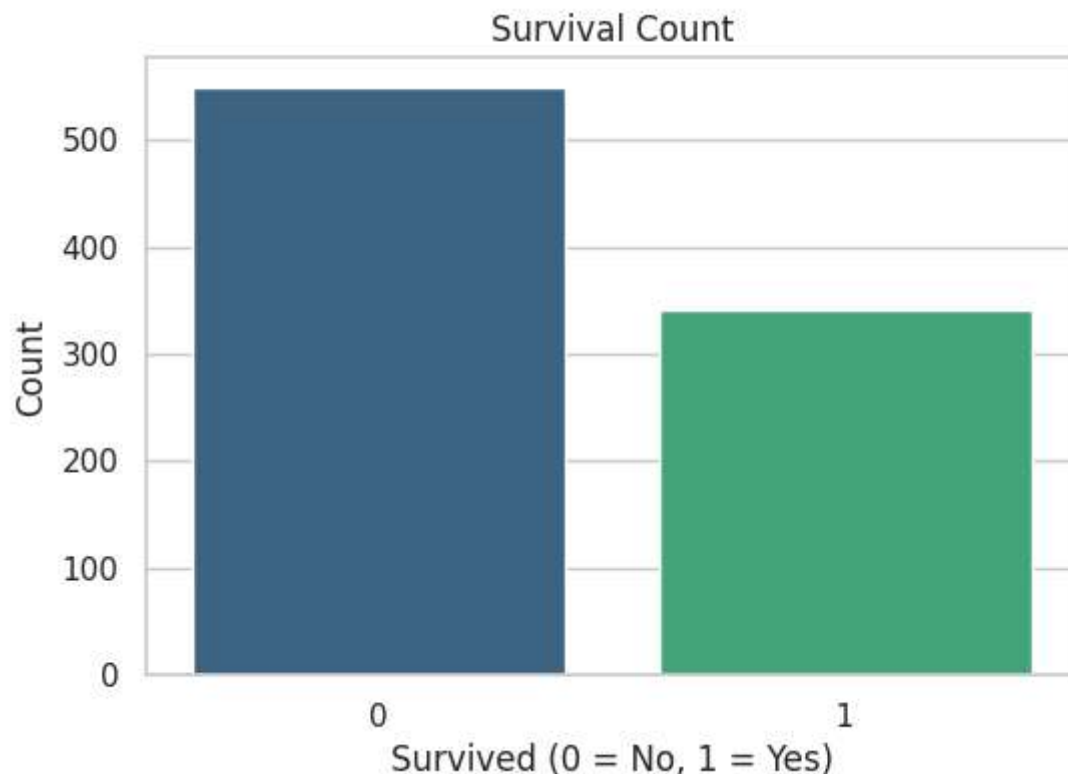
Double-click (or enter) to edit

⌄ VISUALIZATION USING BAR PLOT

```
plt.figure(figsize=(6, 4))
sns.countplot(data=titanic_df, x='Survived', palette='viridis')
plt.title('Survival Count')
plt.xlabel('Survived (0 = No, 1 = Yes)')
plt.ylabel('Count')
plt.show()
```

```
<ipython-input-14-693388f66871>:3: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0.

  sns.countplot(data=titanic_df, x='Survived', palette='viridis')
```


The bar plot displays the number of passengers who survived (1) compared to those who did not (0). This visualization offers a clear overview of the Titanic passenger survival rate. It highlights a notable disparity, with over 500 passengers not surviving, while the number of survivors falls between 300 and 400.

˅ VISUALIZATION USING BARPLOT

```python
plt.figure(figsize=(8, 5))
sns.countplot(data=titanic_df, x='Pclass', hue='Survived', palette='viridis')
plt.title('Survival by Passenger Class')
plt.xlabel('Passenger Class')
plt.ylabel('Count')
plt.legend(title='Survived')
plt.show()
```
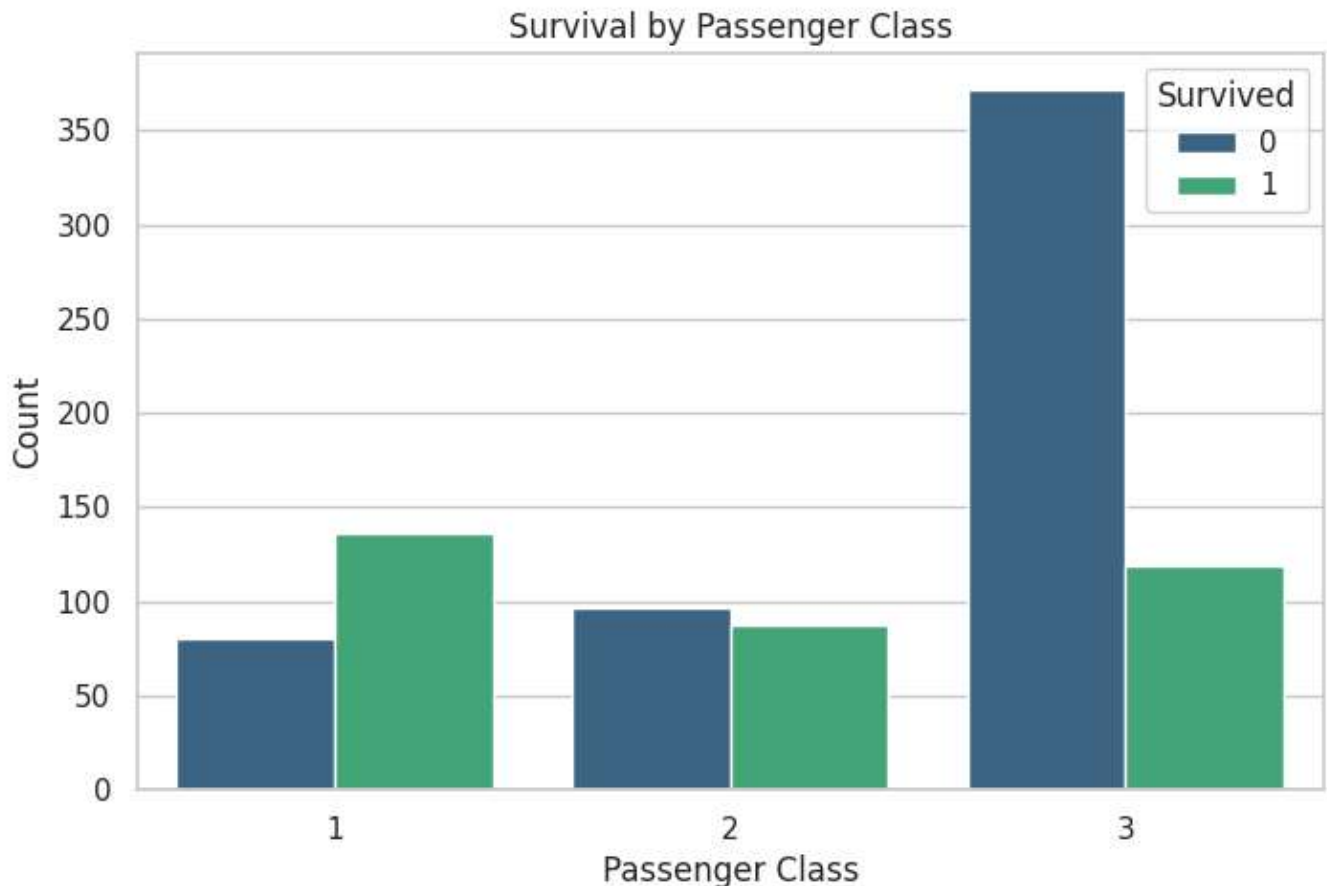

The barplot clearly reveals that passengers in the first class had the highest survival rate, with a significant proportion surviving. Second-class passengers experienced a moderate survival rate, while third-class passengers had the lowest, with the majority not surviving. This graph underscores a clear trend: higher passenger classes are strongly associated with better chances of survival.

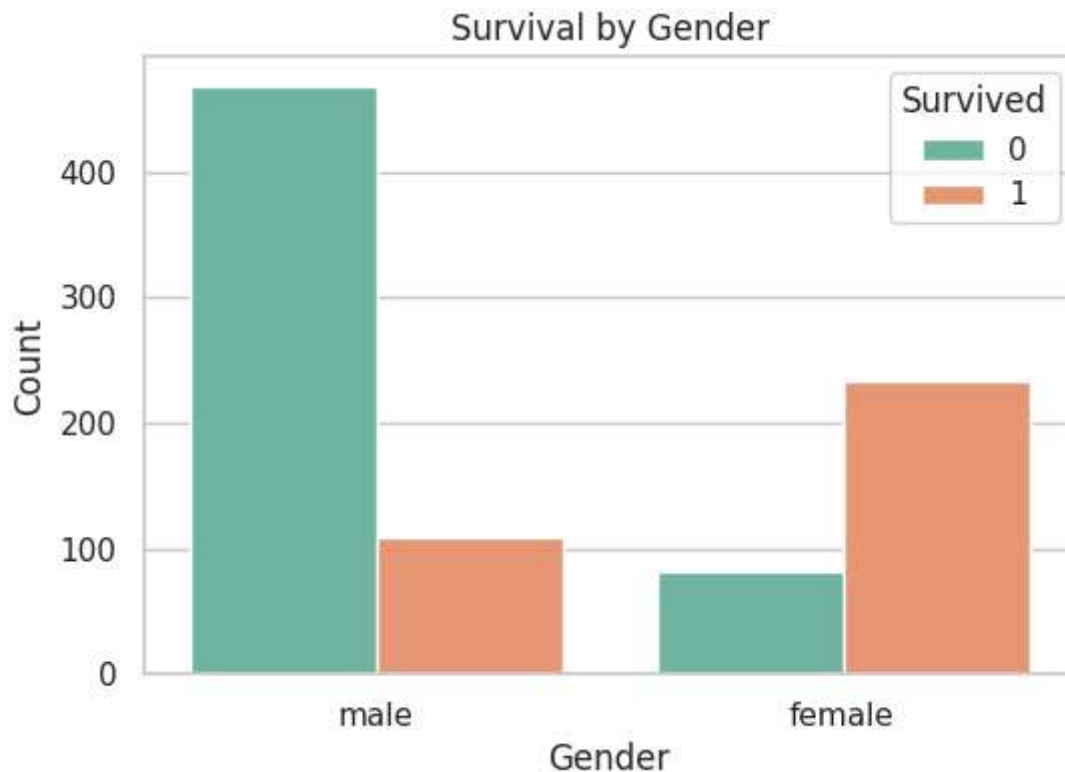## ✓ VISUALIZATION USING SCATTER PLOT

```
plt.figure(figsize=(10, 6))
sns.scatterplot(data=titanic_df, x='Age', y='Fare', hue='Survived', palette='coolwarm')
plt.title('Fare vs Age (Colored by Survival)')
plt.xlabel('Age')
plt.ylabel('Fare')
plt.show()
```
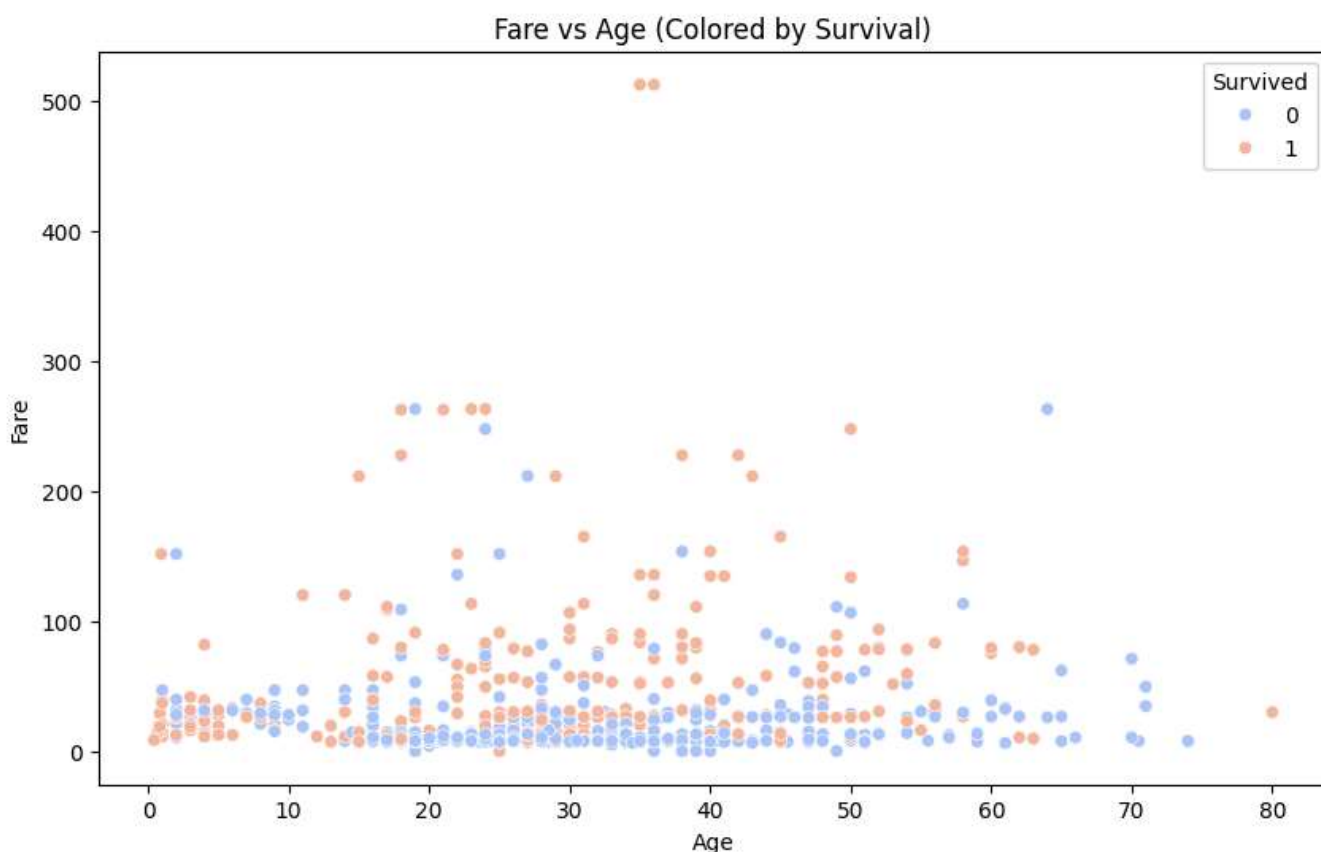

This scatter plot compares fare and age, with red dots representing survivors and blue dots indicating those who did not survive. While there is no obvious correlation between age and fare, younger passengers, particularly children, had a higher likelihood of survival. Those who paid higher fares were more likely to survive, with first-class passengers showing the highest chances. Older adults (over 60) had a lower survival rate. In conclusion, both age and fare appear to influence survival, but other factors may have also played a significant role.

⌄ VISUALIZATION USING BARPLOT

```
plt.figure(figsize=(6, 4))
sns.countplot(data=titanic_df, x='Sex', hue='Survived', palette='Set2')
plt.title('Survival by Gender')
plt.xlabel('Gender')
plt.ylabel('Count')Visualization 5: Survival by Gender (Bar Plot)
plt.legend(title='Survived')
plt.show()
```


The bar plot shows how many men and women survived the Titanic disaster. It splits the counts into two categories for each gender: survived and not survived. The bars are colored to show survival, with different colors for people who made it and those who didn't. From the plot, it's clear that more women survived than men, suggesting that women had a better chance of surviving the disaster

✓ VISUALIZATION USING BARPLOT OF MEAN SURVIVAL

```
plt.figure(figsize=(8, 5))
```

```
sns.barplot(data=titanic_df, x='Pclass', y='Survived', ci=None, palette='viridis')
plt.title('Survival Rate by Passenger Class')
plt.xlabel('Passenger Class')
plt.ylabel('Survival Rate')
plt.show()
```

⇥ thon-input-9-914c543df4d1>:3: FutureWarning:

`ci` parameter is deprecated. Use `errorbar=None` for the same effect.

s.barplot(data=titanic_df, x='Pclass', y='Survived', ci=None, palette='viridis')
thon-input-9-914c543df4d1>:3: FutureWarning:

ing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. /

s.barplot(data=titanic_df, x='Pclass', y='Survived', ci=None, palette='viridis')
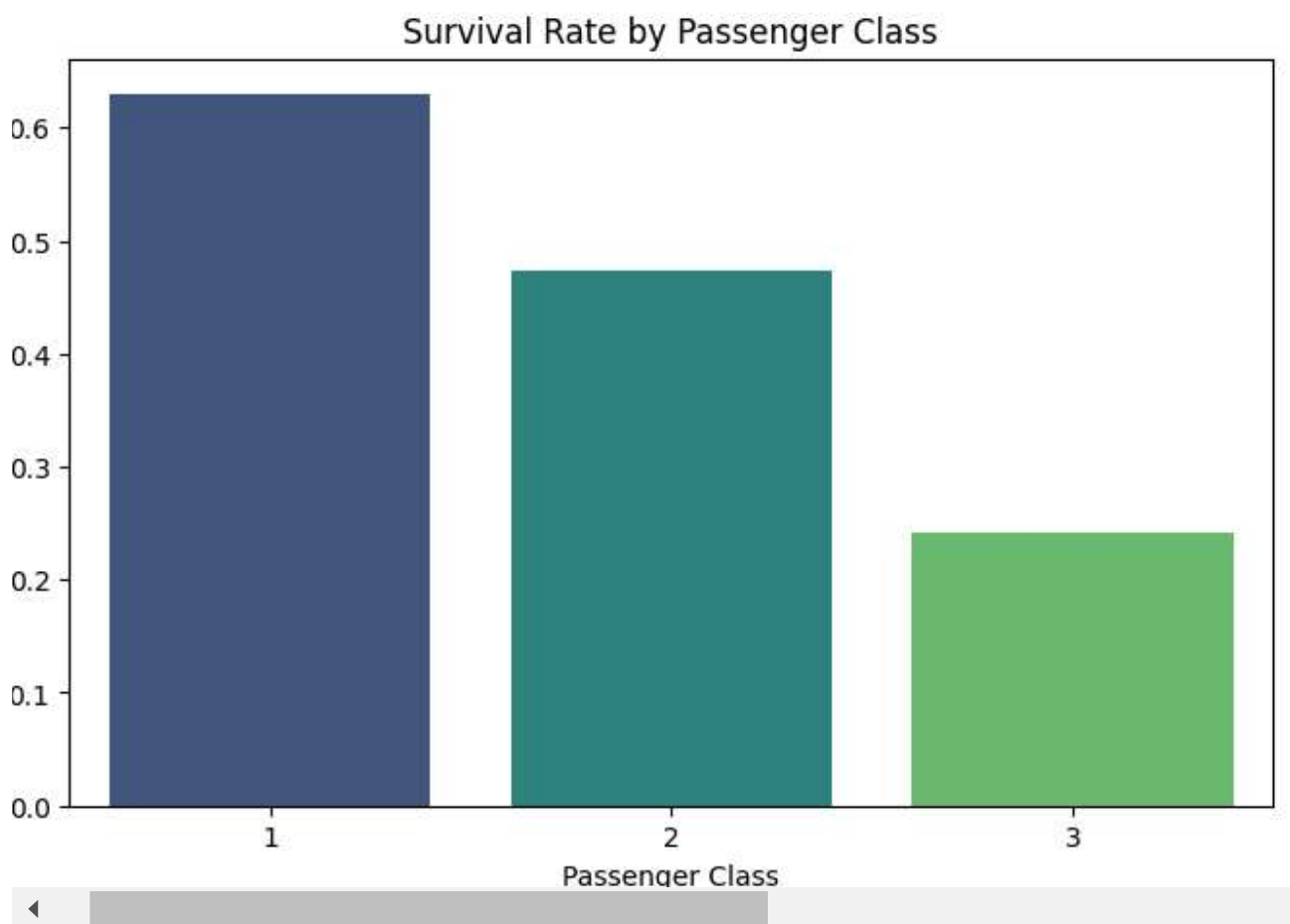
The bar plot shows the survival rate for each passenger class on the Titanic. It displays the average survival rate for passengers in first, second, and third class. The plot uses the 'viridis' color palette to differentiate the classes, and it shows that first-class passengers had the highest survival rate, followed by second-class passengers, with third-class passengers having the lowest

survival rate. This suggests that passengers in higher classes were more likely to survive the disaster.

✓ VISUALIZATION USING BOXPLOT

```python
plt.figure(figsize=(8, 5))
sns.boxplot(data=titanic_df, x='Survived', y='Age', palette='Set2')
plt.title('Age Distribution by Survival')
plt.xlabel('Survived (0 = No, 1 = Yes)')
plt.ylabel('Age')
plt.show()
```

```
<ipython-input-19-27b0a4b0234d>:3: FutureWarning:

  Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0.

    sns.boxplot(data=titanic_df, x='Survived', y='Age', palette='Set2')
```
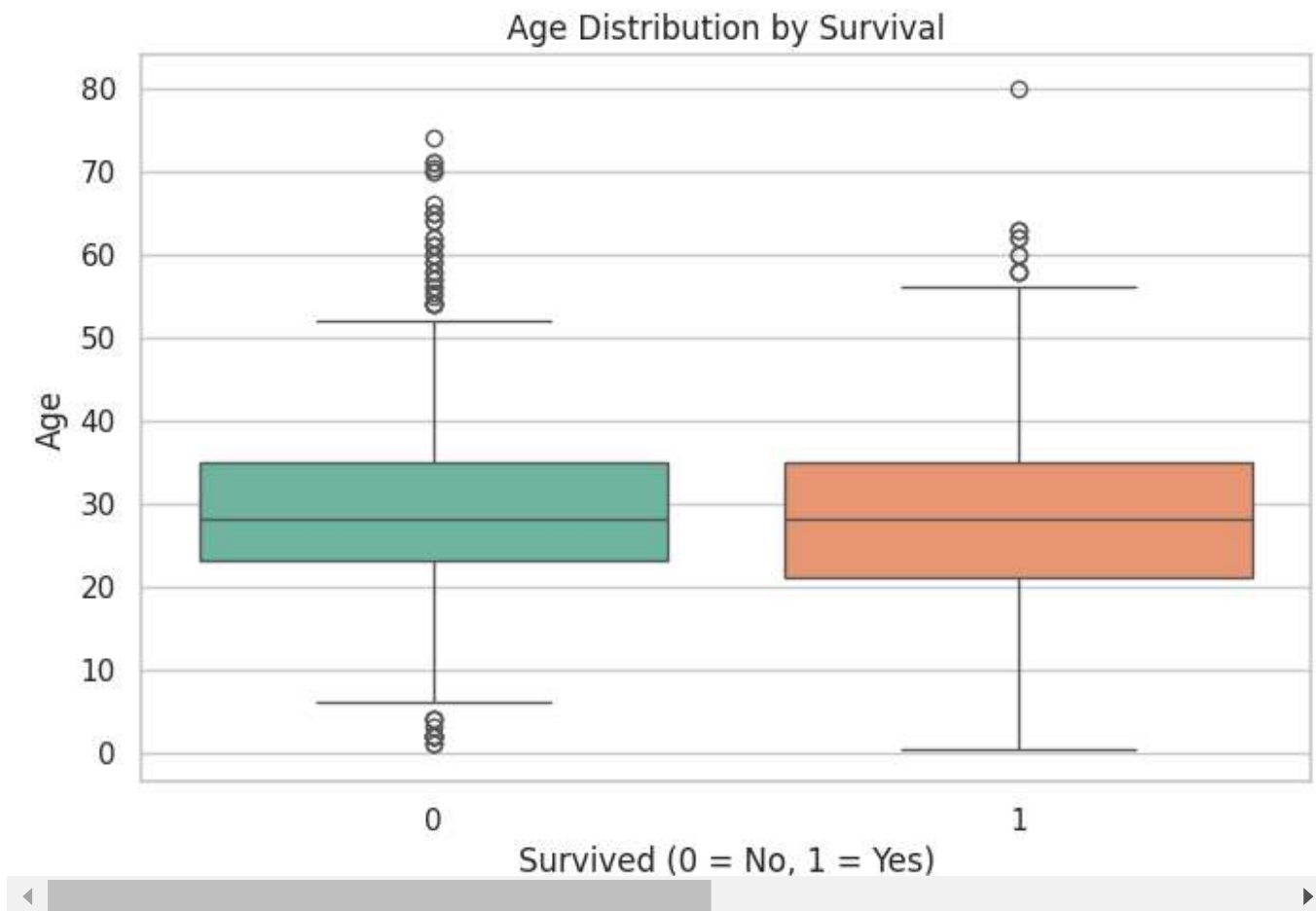

The box plot shows the age distribution for passengers who survived and