

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

sns.set(style="whitegrid")

df = pd.read_csv("titanic_data/train.csv")
```

Observations:

- Most passengers **did not survive** (Survived = 0 is dominant).
- There were **more male passengers** than female.
- **3rd class (Pclass = 3)** had the highest number of passengers.
- Most passengers were between the age range of **20–40 years**.

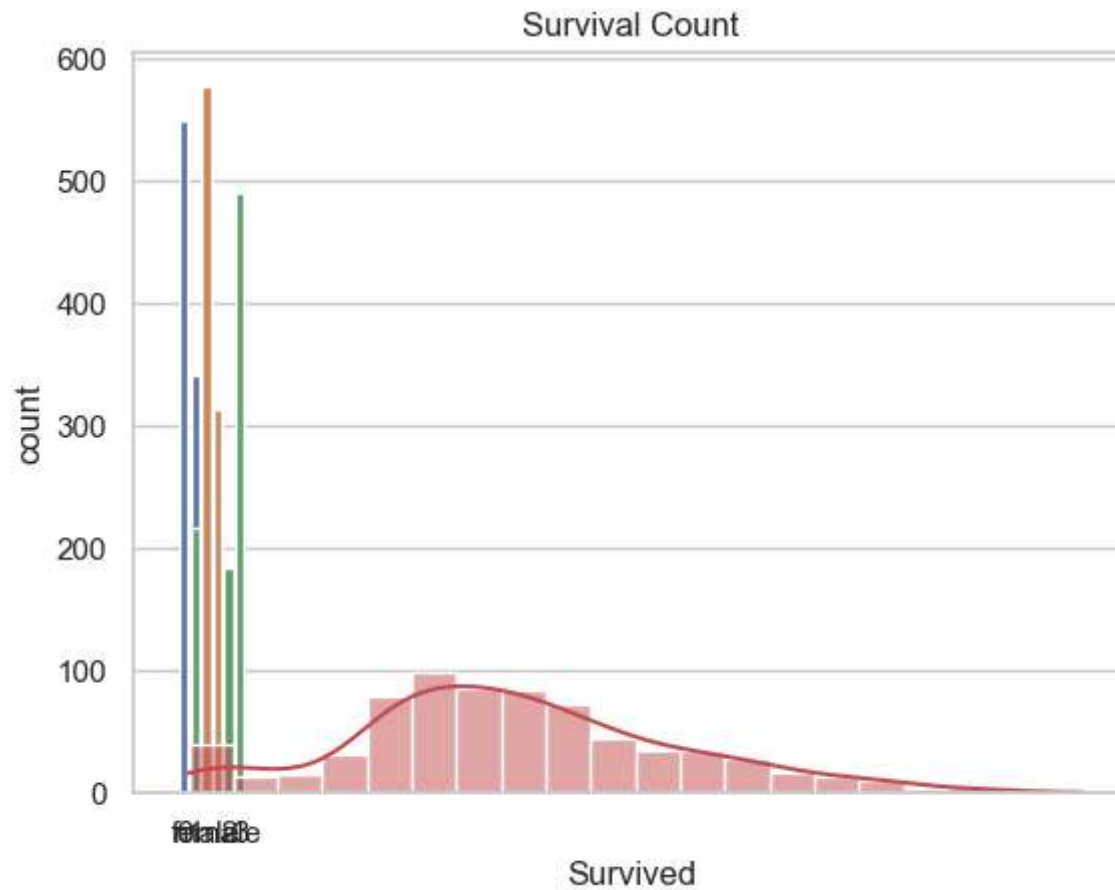
```
In [12]: sns.countplot(x='Survived', data=df)
plt.title("Survival Count")

sns.countplot(x='Sex', data=df)

sns.countplot(x='Pclass', data=df)

sns.histplot(df['Age'].dropna(), kde=True)
```

```
Out[12]: <Axes: title={'center': 'Survival Count'}, xlabel='Survived', ylabel='count'>
```



Observations:

- **Females had a higher survival rate** compared to males.
- Passengers in **1st class had significantly better chances of survival** than those in 2nd or 3rd class.
- Survivors tend to be **younger**, while many older passengers didn't survive.

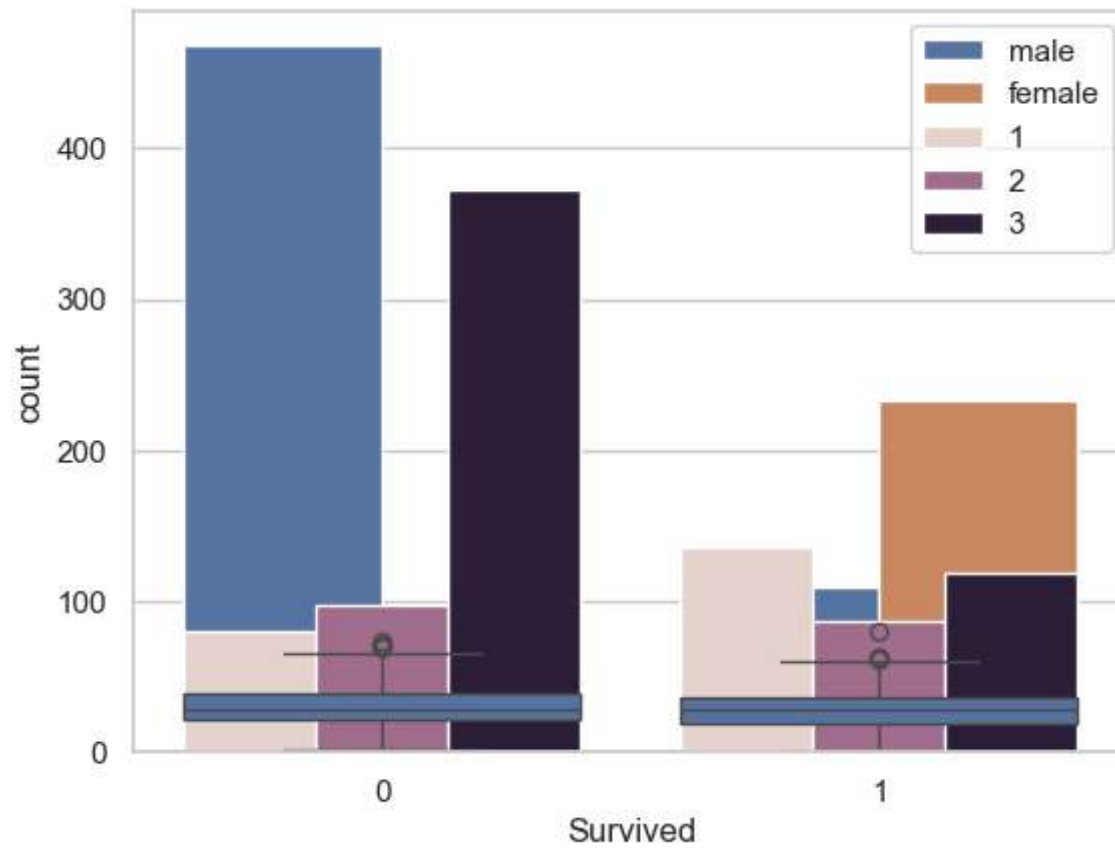
```
In [14]: sns.countplot(x='Survived', hue='Sex', data=df)

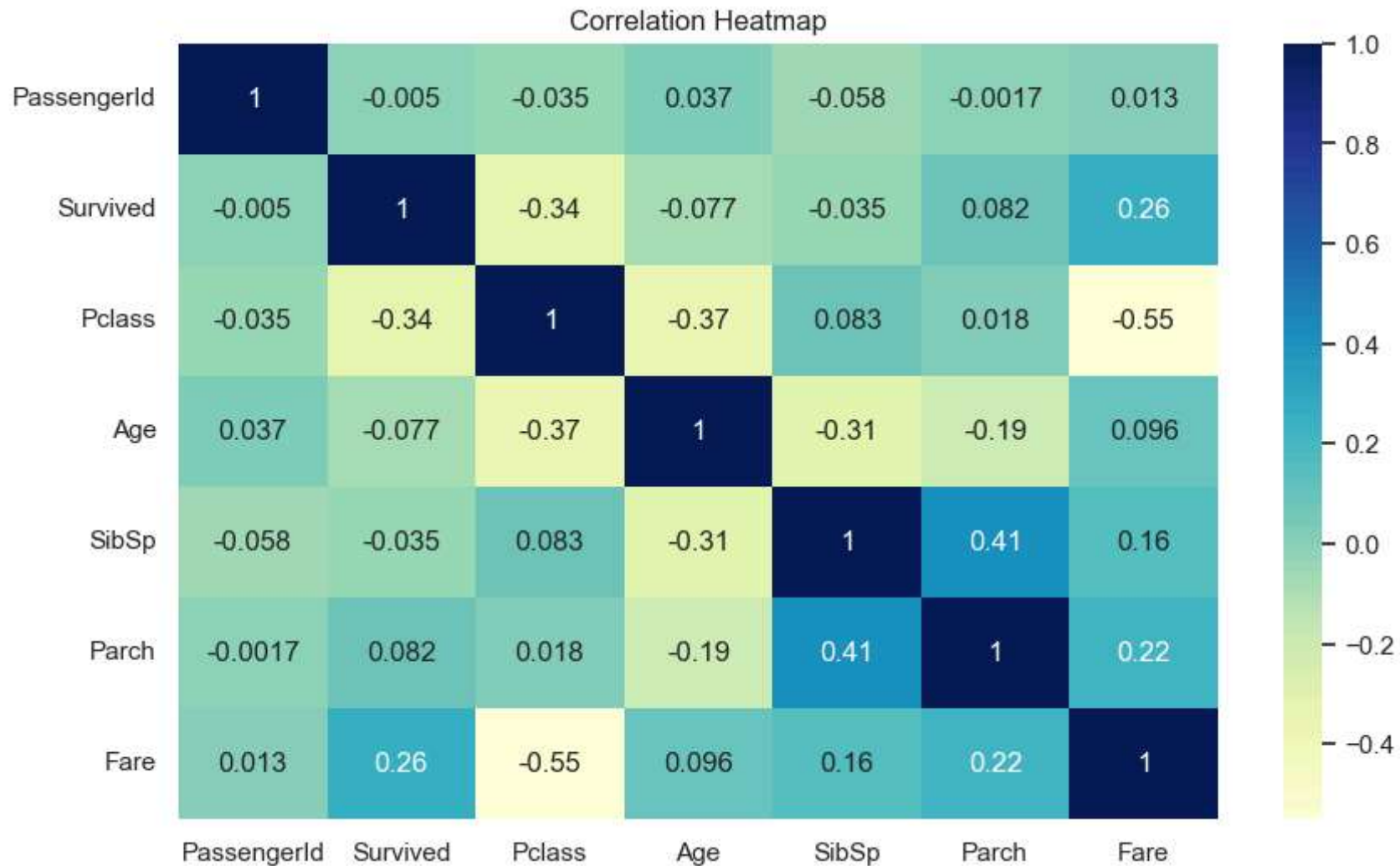
sns.countplot(x='Survived', hue='Pclass', data=df)

sns.boxplot(x='Survived', y='Age', data=df)
```

```
plt.figure(figsize=(10,6))  
sns.heatmap(df.corr(numeric_only=True), annot=True, cmap="YlGnBu")  
plt.title("Correlation Heatmap")
```

Out[14]: Text(0.5, 1.0, 'Correlation Heatmap')



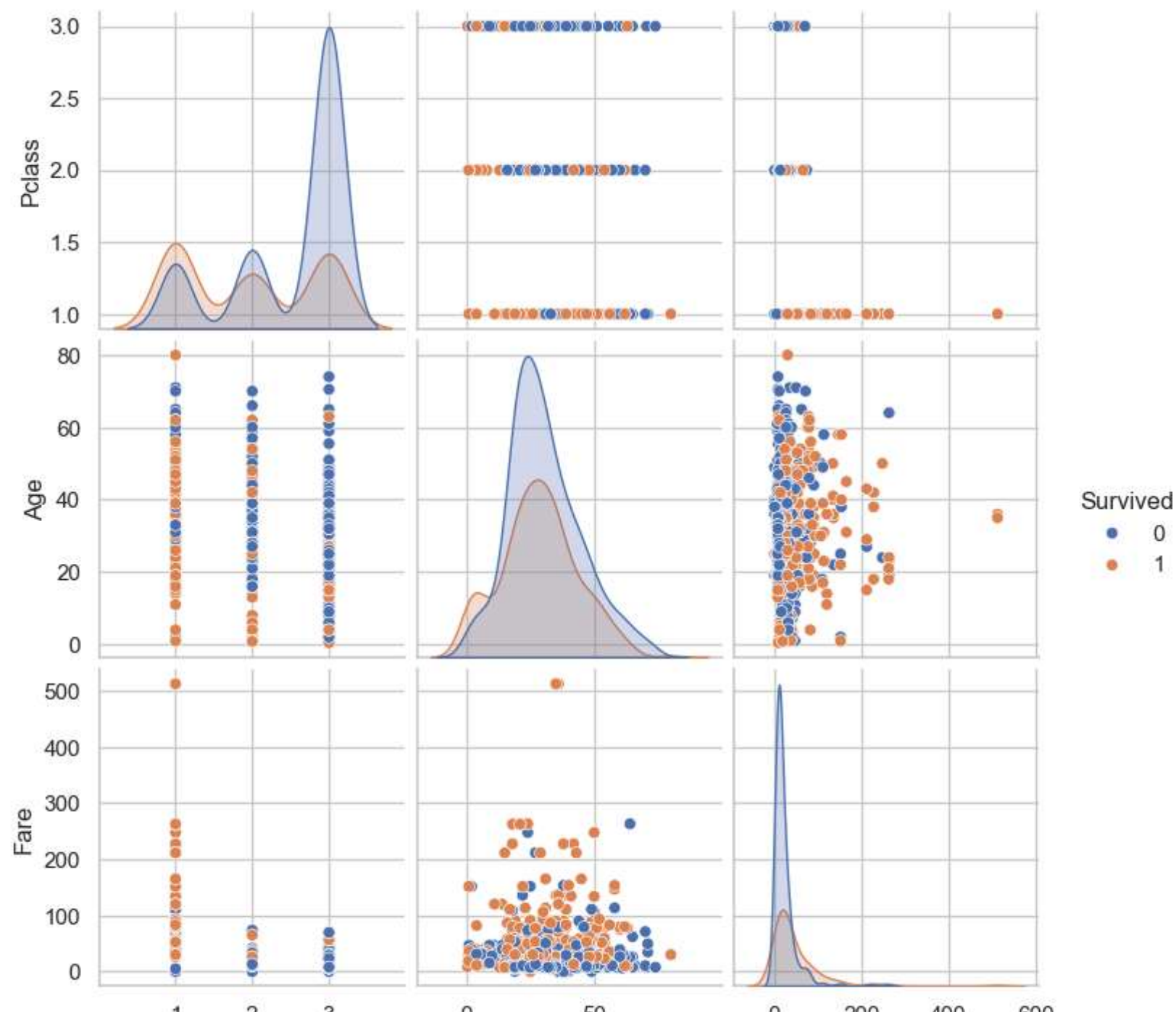


Observations:

- Fare has a **positive correlation** with Survived, indicating that passengers who paid more were more likely to survive.
- Pclass is **negatively correlated** with Survived, meaning lower class number (higher class) had better chances.
- Age and Survived show a **slight negative correlation**, suggesting younger passengers had better odds.

```
In [15]: sns.pairplot(df[['Survived', 'Pclass', 'Age', 'Fare']], hue='Survived')
```

```
Out[15]: <seaborn.axisgrid.PairGrid at 0x2c9a5d4b230>
```





Observations:

- The **pairplot** reinforces previous insights:
 - Survivors are often from higher classes and paid higher fares.
 - There's a clear separation in `Pclass`, `Fare`, and `Survived`.
 - There's no strong trend between `Age` and `Survived`, but younger age still slightly favors survival.