ygevduzat

April 29, 2025

[1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

[2]:
```
# Load the Titanic dataset
df = sns.load_dataset('titanic')
```

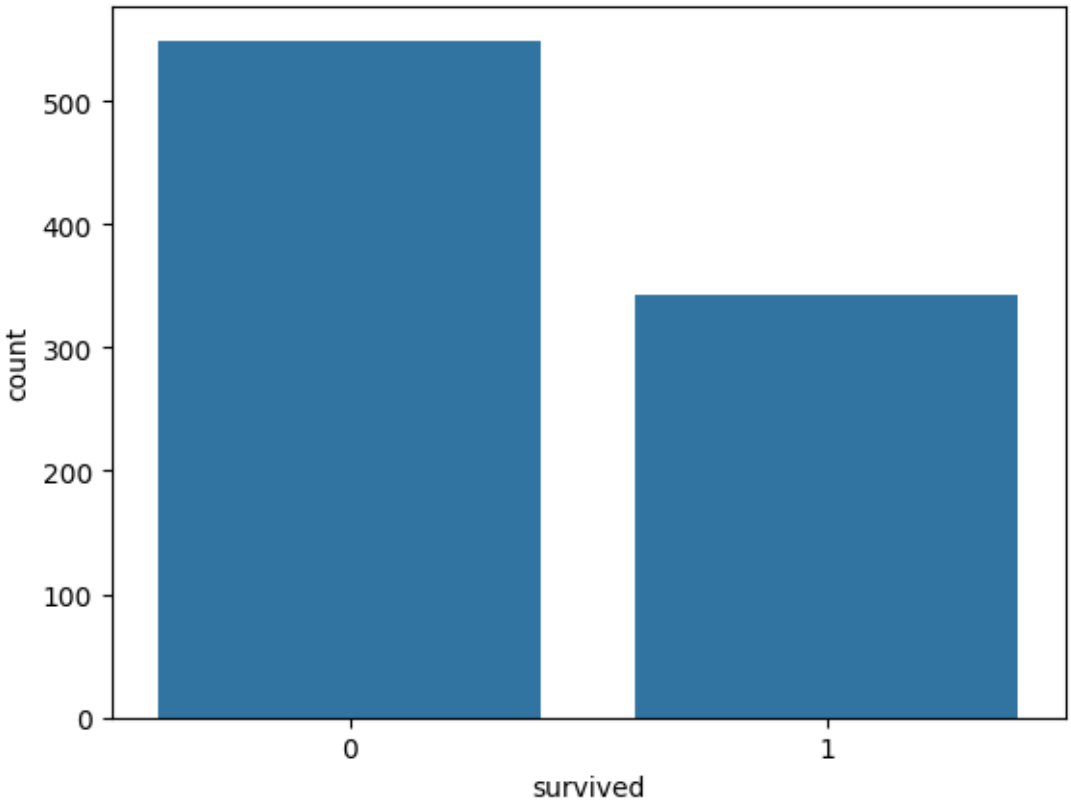
[3]:
```
# Display the first few rows of the dataset
print(df.head())
```

```
   survived  pclass     sex   age  sibsp  parch     fare embarked  class  \
0         0       3    male  22.0      1      0   7.2500        S  Third
1         1       1  female  38.0      1      0  71.2833        C  First
2         1       3  female  26.0      0      0   7.9250        S  Third
3         1       1  female  35.0      1      0  53.1000        S  First
4         0       3    male  35.0      0      0   8.0500        S  Third

     who  adult_male deck  embark_town alive  alone
0    man        True  NaN  Southampton    no  False
1  woman       False    C    Cherbourg   yes  False
2  woman       False  NaN  Southampton   yes   True
3  woman       False    C  Southampton   yes  False
4    man        True  NaN  Southampton    no   True
```

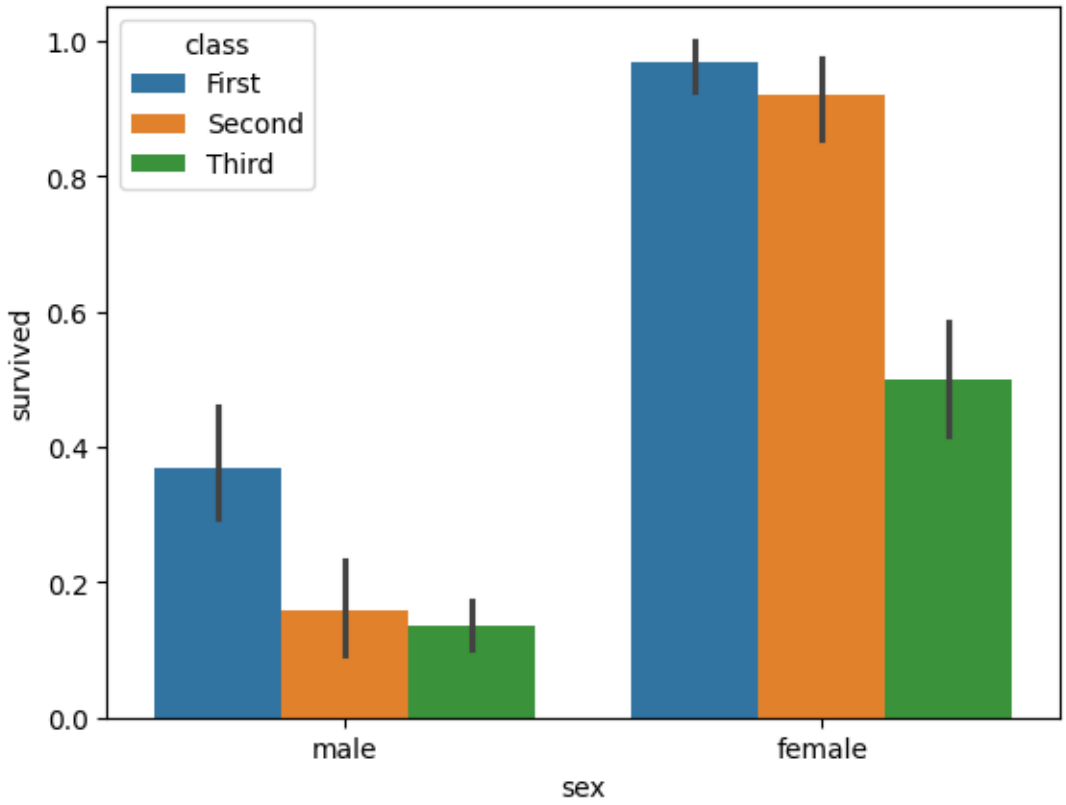
[4]:
```
# Countplot:
sns.countplot(x='survived', data=df)
```

[4]:
```
<Axes: xlabel='survived', ylabel='count'>
```



[5]:
```
# Barplot:
sns.barplot(x='sex', y='survived', hue='class', data=df)
```

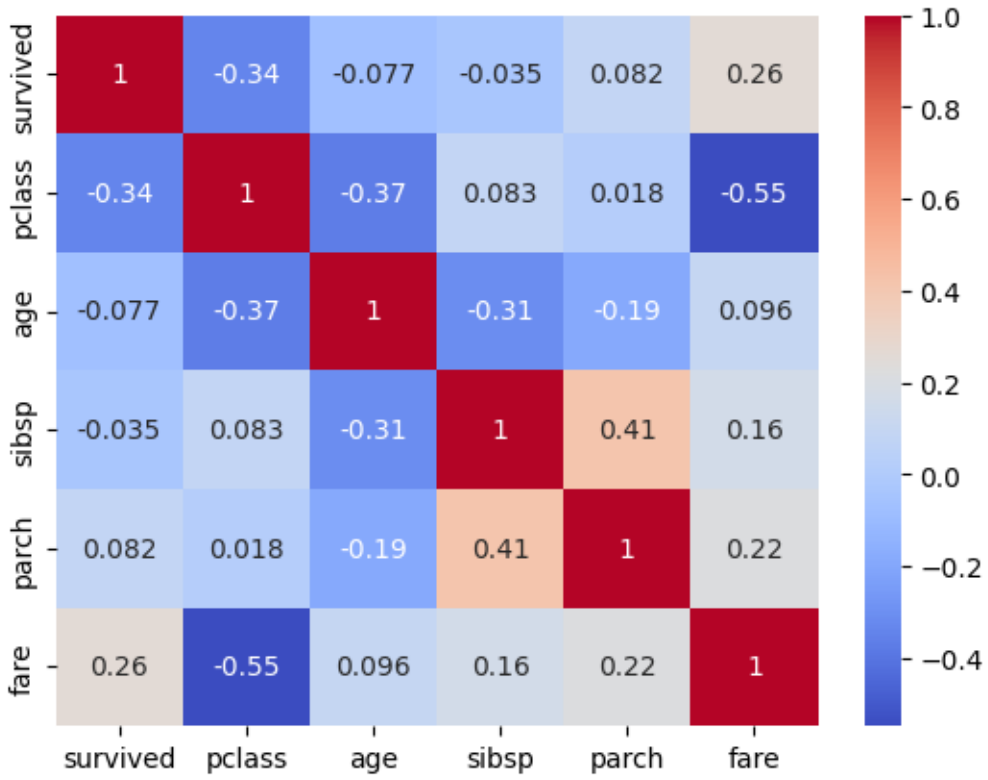
[5]: <Axes: xlabel='sex', ylabel='survived'>



[6]:
```
# Heatmap:
# Select only numeric columns
df_numeric = df.select_dtypes(include=['float64', 'int64'])

# Create correlation matrix and heatmap
sns.heatmap(df_numeric.corr(), annot=True, cmap='coolwarm')
```

[6]: <Axes: >



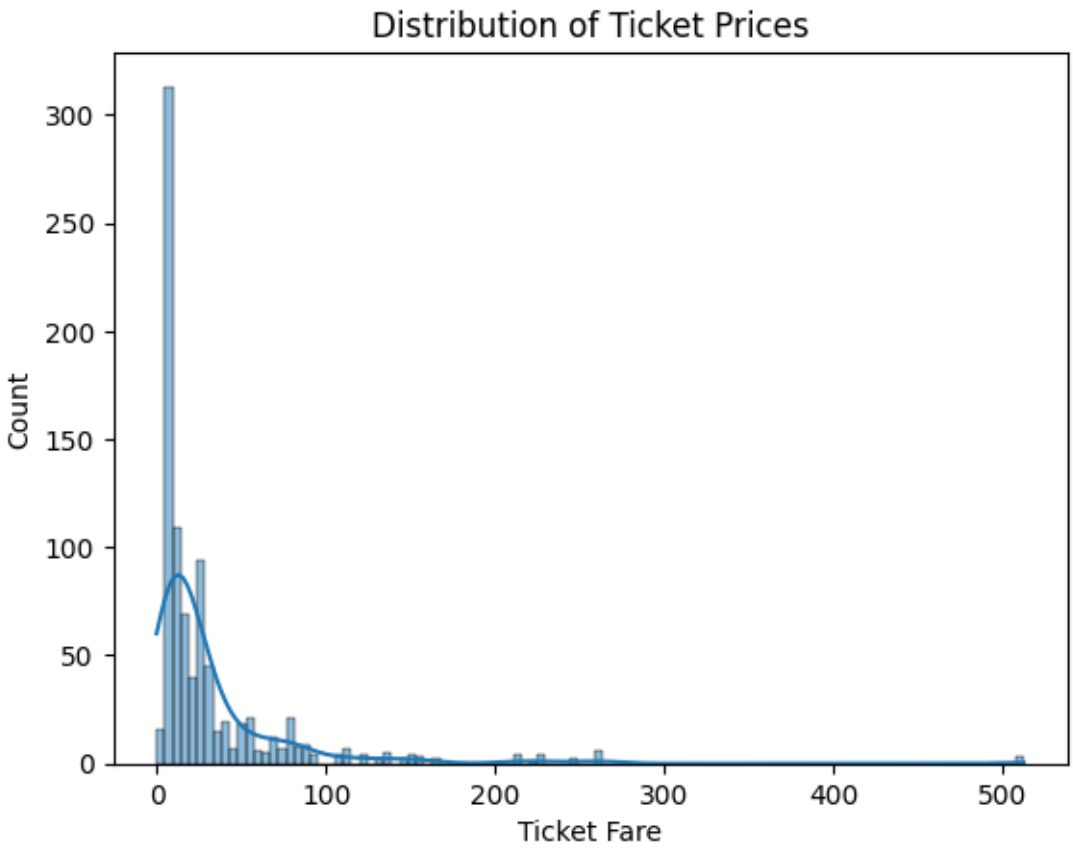
Part 2

[7]:
```
# Load the Titanic dataset
df = sns.load_dataset('titanic')

# Plot histogram of ticket prices
sns.histplot(data=df, x='fare', kde=True)

# Set plot title and labels
plt.title('Distribution of Ticket Prices')
plt.xlabel('Ticket Fare')
plt.ylabel('Count')

# Display the plot
plt.show()
```


Distribution of Ticket Prices
Count
Ticket Fare
0
50
100
150
200
250
300
0
100
200
300
400
500

[ ]: