# Display the first few rows of the dataset

print(df.head())

# Display the structure of the dataset

print(df.info())

# Display the statistical summary of the dataset

print(df.describe())

# Handle missing values

df['Age'].fillna(df['Age'].median(), inplace=True)

df['Embarked'].fillna(df['Embarked'].mode()[0], inplace=True)

df.drop('Cabin', axis=1, inplace=True) # Too many missing values

# Remove duplicates

df.drop\_duplicates(inplace=True)

import matplotlib.pyplot as plt

import seaborn as sns

# Set the aesthetic style of the plots

sns.set\_style("whitegrid")

# Visualization 1: Distribution of Ages

plt.figure(figsize=(10, 6))

sns.histplot(df['Age'], bins=30, kde=True)

plt.title('Distribution of Ages')

plt.xlabel('Age')

plt.ylabel('Frequency')

plt.show()

# Visualization 2: Survival Rate by Gender

plt.figure(figsize=(10, 6))

sns.barplot(x='Sex', y='Survived', data=df)

plt.title('Survival Rate by Gender')

plt.xlabel('Gender')

plt.ylabel('Survival Rate')

plt.show()