import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

titanic = sns.load\_dataset('titanic')

print(titanic.info())

print(titanic.head())

print("Count of Null values in various columns:")

print(titanic.isnull().sum())

# Fill missing values in 'age' with the mean

titanic['age'].fillna(titanic['age'].mean(), inplace=True)

# Fill missing values in 'embarked' and 'embark\_town' with the forward fill method

titanic['embarked'].fillna(method='ffill', inplace=True)

titanic['embark\_town'].fillna(method='ffill', inplace=True)

# Fill missing values in 'deck' with the mode

titanic['deck'].fillna(titanic['deck'].mode()[0], inplace=True)

# Confirm no missing values remain

print("After filling Null values using various techniques:")

print(titanic.isnull().sum())

# Select relevant columns for analysis

filtered\_data = titanic[['sex', 'age', 'survived']]

# Display the first few rows of the filtered dataset

print(filtered\_data.head())

# Boxplot for age distribution by gender and survival status

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

sns.boxplot(x='sex', y='age', hue='survived', data=filtered\_data)

plt.title('Age Distribution by Gender and Survival Status')

plt.xlabel('(Gender, Survived)')

plt.ylabel('Age')

plt.xticks(rotation=45, ha='right')

plt.grid(False) # Remove gridlines for better clarity

plt.show()