

Practical No-9

Date of Conduction :

Date of Checking:

Data Visualization II

1. Use the inbuilt dataset 'titanic' as used in the above problem. Plot a box plot for distribution of age with respect to each gender along with the information about whether they survived or not. (Column names : 'sex' and 'age')
2. Write observations on the inference from the above statistics.

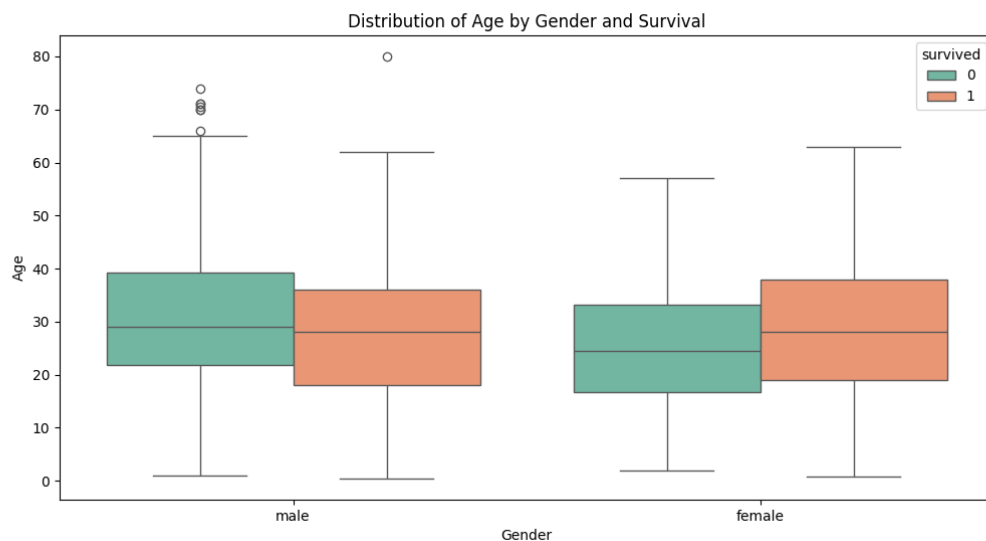
Python Code:

```
import seaborn as sns
import matplotlib.pyplot as plt

# Load the Titanic dataset
titanic = sns.load_dataset('titanic')

# Plotting a box plot for distribution of age with respect to each gender
and survival
plt.figure(figsize=(12, 6))
sns.boxplot(x='sex', y='age', hue='survived', data=titanic, palette='Set2')
plt.title('Distribution of Age by Gender and Survival')
plt.xlabel('Gender')
plt.ylabel('Age')
plt.show()
```

OUTPUT:



Observations:

- The box plot provides insights into the distribution of age with respect to each gender and survival status.
- For both genders, the boxes represent the interquartile range (IQR), and the line inside the box represents the median age.
- Outliers are shown as individual points beyond the whiskers of the box.
- The color hue represents survival status (survived or not survived).
- You can observe the variation in age distribution between males and females and how it relates to their survival.