Practical 9: 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.

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

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

# Display basic information about the dataset
print(titanic.info())
```

```
# Plot a box plot for age distribution by gender and survival status
plt.figure(figsize=(12, 8))
sns.boxplot(data=titanic, x='sex', y='age', hue='survived')
plt.title('Age Distribution by Gender and Survival Status')
plt.xlabel('Gender')
plt.ylabel('Age')
plt.show()
```

Observations and Inferences:

1. **Age Distribution by Gender**:
    - **Males**: The age distribution for male passengers is wider, showing a larger spread of ages.
    - **Females**: The age distribution for female passengers is more concentrated, with fewer outliers compared to males.
2. **Survival Trends**:
    - **Females had a higher survival rate than males**:

■ The boxplot for **female survivors** is significantly larger than that for **female non-survivors**, indicating that a higher proportion of women survived.
■ The boxplot for **male survivors** is smaller than that for **male non-survivors**, indicating that more men did not survive.

3. **Presence of Outliers**:
   ○ There are a few extreme age values (outliers), particularly for older passengers.
   ○ Some elderly passengers survived, but their numbers were very few.

4. **Conclusion**:
   ○ The **"women and children first"** policy seems to have played a major role in survival rates.
   ○ Younger males had a slightly better survival chance than older males, but overall, **males had a lower survival rate** compared to females.

Key Takeaways from the Boxplot:

● **Women were more likely to survive** than men, across all age groups.
● **Young passengers (both male and female) had better survival rates**, possibly because they were prioritized for lifeboats.
● **Older passengers, particularly males, had lower survival rates**.