**Exp 9**

**import pandas as pd**

**import numpy as np**

**import matplotlib.pyplot as plt**

**import seaborn as sns**

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

**dataset.head()**

**sns.boxplot(x='sex', y='age', data=dataset)**

**sns.boxplot(x='sex', y='age', data=dataset, hue="survived")**

Certainly! Let's delve deeper into each line:

1. \*\*Importing necessary libraries\*\*: This line imports the required libraries for data manipulation and visualization. `pandas` is imported as `pd` for data handling, `numpy` as `np` for numerical operations, `matplotlib.pyplot` as `plt` for plotting, and `seaborn` as `sns` for statistical data visualization.

2. \*\*Loading the Titanic dataset into a Pandas DataFrame\*\*: The `sns.load\_dataset('titanic')` function loads the Titanic dataset from the Seaborn library into a Pandas DataFrame named `dataset`. This dataset contains information about passengers on the Titanic, including details like their age, sex, fare, and whether they survived or not.

3. \*\*Displaying the first few rows of the dataset\*\*: The `dataset.head()` function displays the first few rows (by default, the first five rows) of the DataFrame `dataset`. This helps in getting a quick overview of the data and its structure.

4. \*\*Creating a box plot showing the distribution of 'age' by 'sex'\*\*: The `sns.boxplot(x='sex', y='age', data=dataset)` line creates a box plot using Seaborn. The 'sex' column is plotted on the x-axis, and the 'age' column is plotted on the y-axis. This visualization allows us to compare the distribution of ages between male and female passengers. Each box represents the interquartile range (IQR) of ages for a specific gender, with the median age depicted by a line inside the box.

5. \*\*Creating a box plot showing the distribution of 'age' by 'sex', with separate boxes for each 'survived' category\*\*: The `sns.boxplot(x='sex', y='age', data=dataset, hue="survived")` line is similar to the previous one but adds another layer of information. In addition to comparing the age distribution between male and female passengers, this plot also distinguishes between passengers who survived (hue = "survived" and value = True) and those who did not survive (hue = "survived" and value = False). This is achieved by using the `hue` parameter, which separates the data points based on the values of the specified column ('survived' in this case). Each box now represents the age distribution within each gender group, further segmented by survival status.

These visualizations help in exploring the relationship between passengers' age, gender, and survival outcome, providing insights into potential patterns or differences within the Titanic dataset.