

## Assignment No 9:

### Title: Data Visualization II

1. Use the inbuilt dataset 'titanic'. 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.

### Theory:

#### The Bar Plot

The barplot() is used to display the mean value for each value in a categorical column, against a numeric column. The first parameter is the categorical column, the second parameter is the numeric column while the third parameter is the dataset.

### Algorithm:

#### Step 1: Download the data set of Titanic Step

#### Step 2: Importing Libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

#### Step 3: Reading of dataset

```
df = pd.read_csv('titanic')
```

#### Step 4: Drop columns

```
df_cleaned=df.drop(['pclass','embarked','deck','embark_town'],axis=1)
```

#### Step 5: Draw Box Plot

```
sns.boxplot(data=df,x="sex",y="age", color="Purple")
sns.boxplot(data=df,x="sex",y="age", hue="survived")
```

**Question:** Write the observations from the box plot.

**Conclusion:** Implemented successfully Box Plot Data visualization techniques using Python on Titanic dataset.