

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.

In [1]:

```
pip install pandas
```

Defaulting to user installation because normal site-packages is not writeable
 Requirement already satisfied: pandas in /home/pict/.local/lib/python3.6/site-packages (1.1.5)
 Requirement already satisfied: pytz>=2017.2 in /usr/lib/python3/dist-packages (from pandas) (2018.3)
 Requirement already satisfied: python-dateutil>=2.7.3 in /home/pict/.local/lib/python3.6/site-packages (from pandas) (2.8.2)
 Requirement already satisfied: numpy>=1.15.4 in /home/pict/.local/lib/python3.6/site-packages (from pandas) (1.19.5)
 Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from python-dateutil>=2.7.3->pandas) (1.11.0)
 Note: you may need to restart the kernel to use updated packages.

In [2]:

```
pip install seaborn
```

Defaulting to user installation because normal site-packages is not writeable
 Requirement already satisfied: seaborn in /home/pict/.local/lib/python3.6/site-packages (0.11.2)
 Requirement already satisfied: numpy>=1.15 in /home/pict/.local/lib/python3.6/site-packages (from seaborn) (1.19.5)
 Requirement already satisfied: pandas>=0.23 in /home/pict/.local/lib/python3.6/site-packages (from seaborn) (1.1.5)
 Requirement already satisfied: scipy>=1.0 in /home/pict/.local/lib/python3.6/site-packages (from seaborn) (1.5.4)
 Requirement already satisfied: matplotlib>=2.2 in /home/pict/.local/lib/python3.6/site-packages (from seaborn) (3.3.4)
 Requirement already satisfied: pyparsing!=2.0.4,!=2.1.2,!=2.1.6,>=2.0.3 in /usr/lib/python3/dist-packages (from matplotlib>=2.2->seaborn) (2.2.0)
 Requirement already satisfied: python-dateutil>=2.1 in /home/pict/.local/lib/python3.6/site-packages (from matplotlib>=2.2->seaborn) (2.8.2)
 Requirement already satisfied: cycler>=0.10 in /usr/lib/python3/dist-packages (from matplotlib>=2.2->seaborn) (0.10.0)
 Requirement already satisfied: pillow>=6.2.0 in /home/pict/.local/lib/python3.6/site-packages (from matplotlib>=2.2->seaborn) (8.4.0)
 Requirement already satisfied: kiwisolver>=1.0.1 in /home/pict/.local/lib/python3.6/site-packages (from matplotlib>=2.2->seaborn) (1.3.1)
 Requirement already satisfied: pytz>=2017.2 in /usr/lib/python3/dist-packages (from pandas>=0.23->seaborn) (2018.3)
 Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from python-dateutil>=2.1->matplotlib>=2.2->seaborn) (1.11.0)
 Note: you may need to restart the kernel to use updated packages.

In [4]:

```
pip install numpy
```

Defaulting to user installation because normal site-packages is not writeable
 Requirement already satisfied: numpy in /home/pict/.local/lib/python3.6/site-packages (1.19.5)
 Note: you may need to restart the kernel to use updated packages.

In [5]:

```
import numpy as np
import seaborn as sns
import pandas as pd
```

```
In [14]: data = pd.read_csv('./titanic.csv')
print(data.head(5))
```

	Passengerid	Age	Fare	Sex	sibsp	zero	zero.1	zero.2	zero.3	\
0	1	22.0	7.2500	0	1	0	0	0	0	
1	2	38.0	71.2833	1	1	0	0	0	0	
2	3	26.0	7.9250	1	0	0	0	0	0	
3	4	35.0	53.1000	1	1	0	0	0	0	
4	5	35.0	8.0500	0	0	0	0	0	0	

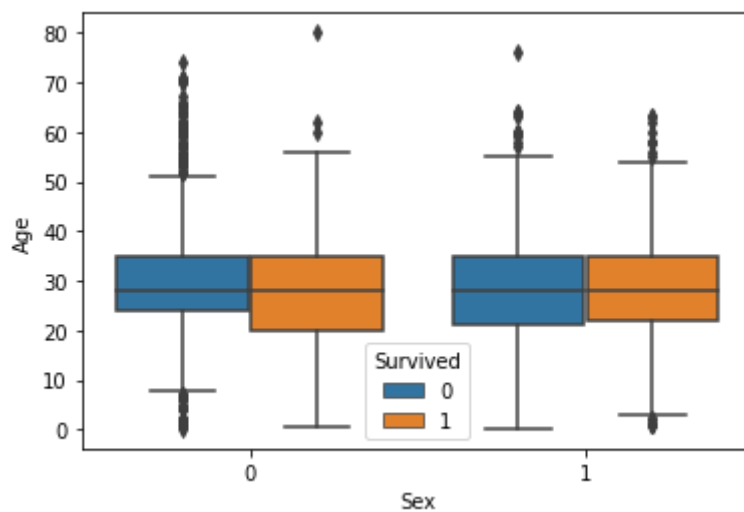
	zero.4	...	zero.12	zero.13	zero.14	Pclass	zero.15	zero.16	Embarked
\									
0	0	...	0	0	0	3	0	0	2.0
1	0	...	0	0	0	1	0	0	0.0
2	0	...	0	0	0	3	0	0	2.0
3	0	...	0	0	0	1	0	0	2.0
4	0	...	0	0	0	3	0	0	2.0

	zero.17	zero.18	Survived
0	0	0	0
1	0	0	1
2	0	0	1
3	0	0	1
4	0	0	0

[5 rows x 28 columns]

```
In [15]: sns.boxplot(y = data['Age'], x = data['Sex'],
                    hue = data['Survived'])
```

Out[15]: <AxesSubplot:xlabel='Sex', ylabel='Age'>



In []: