**Live Session Unit 12 Assignment**

**Due: April 18, 2017 (Tuesday)**

1. Create a list named my\_list in python with the following data points -

45.4 44.2 36.8 35.1 39.0 60.0 47.4 41.1 45.8 35.6

* 1. Print the 5th element in the list
  2. Append 55.2 to my\_list
  3. Remove the 6th element in the list
  4. Iterate over the list to print data points greater than 45

1. Introduction to numpy –
   1. Import the numpy library using the following command – **import numpy**
   2. Declare numpy array with the same data points as in my\_list using **numpy.array()**
   3. Compute the mean and standard deviation using numpy.mean() and numpy.std() of the above array
   4. Use logical referencing to get only those values that are less than 45
   5. Compute the max and min of the array using numpy.max() and numpy.min()
2. Introduction to pandas –
   1. Import the pandas library – **import pandas**
   2. Read the IRIS dataset into **iris** using pandas.read\_csv(). Data file –



* 1. Using iris.head(), display the head of the dataset
  2. Use DataFrame.drop() to drop the **id** column
  3. Subset dataframe to create a new data frame that includes only the measurements for the setosa species
  4. Use DataFrame.describe() to get the summary statistics
  5. Use DataFrame.groupby() to create grouped data frames by Species and compute summary statistics using DataFrame.describe()
  6. Use DataFrame.boxplot() to plot boxplots by Species
  7. Plot a scatter matrix plot using the seaborn library. Use the following to load and plot
     1. Import seaborn
     2. Seaborn.pairplot(dataframe,by=’column\_name’)

Here is the code for part (i)

**import seaborn as sns**

**sns.pairplot(iris,hue='Species')**

**Deliverable is a link to Jupiter notebook file on GitHub.**