**Hubei University of Technology**

Big Data Analysis Homework Report

Class: Name:

Tutor: Prof. Jianxia Chen

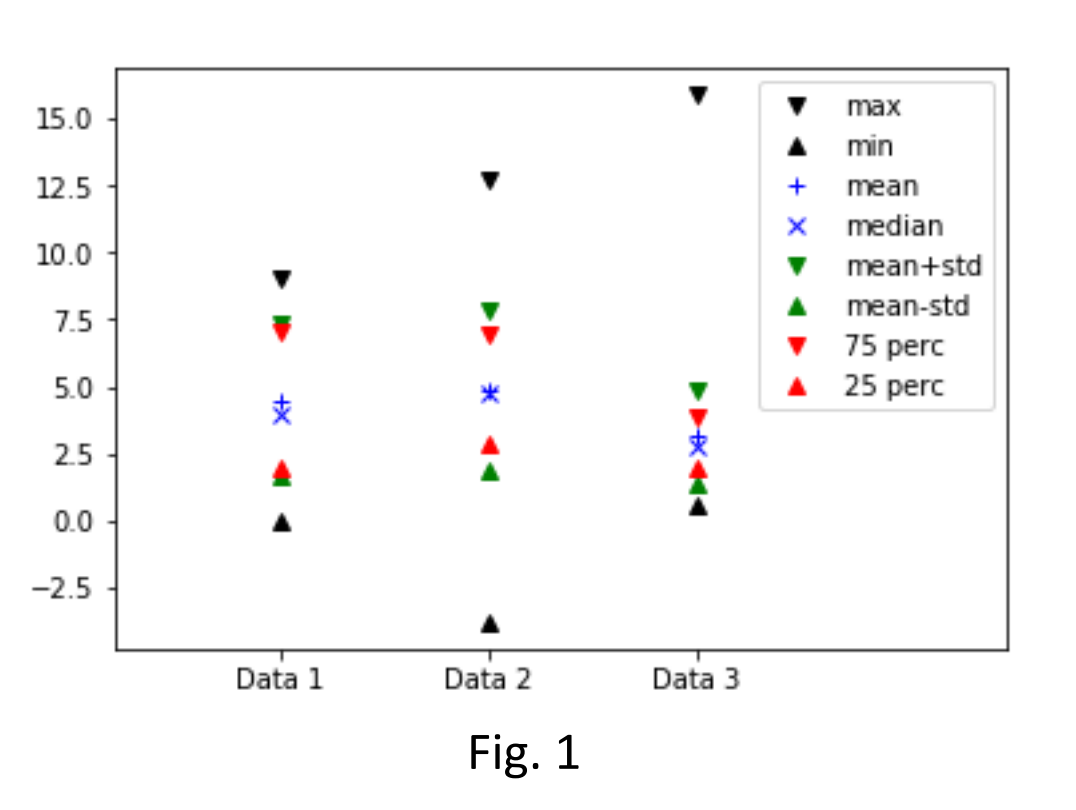
Date: No: 2

Assignment： Basics Plotting & Statistics

1. **Homework purpose and requirements:**
2. Familiar basic statistics.
3. Familiar with Matplotlib.
4. Master the operation of examples.
5. **Homework report:**

1. (50 points) Python programming. Please use the code skeleton provided on the course website.

1. Write a python function to merge sort a list of numbers. Test your function on a small random integer array of 10 elements.
2. Write a python function to calculate the summary statistics of an array of numbers: max, min, mean, standard deviation, median, 75 percentile, 25 percentile. Use the three sets of random numbers as generated in the code skeleton to test your code and output the summary statistics for each data set.
3. Write a python function that accepts a list of numbers in any range, then scales the numbers to integers in [0, 9], and then count the number of occurrences of each integer in the dataset. Test your function on the three datasets used in (2).
4. (50 points) Plotting.
5. For the three sets of data generated in 1-(2), plot the summary statistics in one figure. Try to reproduce all the details of the example Fig 1.



1. Using the counts you collected in 1-(3), plot the distribution of the three data sets in one figure. Again, try to reproduce all the details of the example Fig 2.

