Experiment – 2: Data Structures in Python

Aim: To understand and implement the data structures in python.

Theory:
Data Structures are a way of organizing data so that it can be accessed more efficiently depending upon the situation. The python data structures includes:

- Lists: List is an ordered collection of similar or different types of items separated by commas and enclosed within brackets [ ].
- Tuples: Tuple is an ordered sequence of items that are immutable which means that once created cannot be modified. The parentheses ( ) to store items of a tuple.
- Dictionaries: Dictionary is an ordered collection of items that stores elements in key/value pairs. Keys are unique identifiers that are associated with each value.
- Sets: Set is an unordered collection of unique items defined by values separated by commas inside braces { }.

Conclusion:

Task for submission:
(Write comments for every statement of the program)

1. Open the text file and make a list of unique words contained in the file.
2. Open the file mbox-short.txt and read it line by line. When you find a line that starts with 'From ' like the following line:
   From stephen.marquard@uct.ac.za Sat Jan 5 09:14:16 2008
   You will parse the From line using split() and print out the second word in the line (i.e. the entire address of the person who sent the message). Then print out a count at the end.
3. Write a program to read through the mbox-short.txt and figure out who has sent the greatest number of mail messages. The program looks for 'From ' lines and takes the second word of those lines as the person who sent the mail. The program creates a Python dictionary that maps the sender's mail address to a count of the number of times they appear in the file.