**CS-504 Intro to Programming / Scripting**

**Activity 8 – Lists**

1. Write a program to remove the duplicates from the following list and create a new one without duplicates.
   * 1. Create a list with the numbers below.
     2. Create another empty list.
     3. Create a loop to check whether each member of the first list is in the second list.
     4. Add the member to the new list if not.
     5. Print the new list.

list1 = [1, 2, 3, 3, 3, 4, 4, 5, 6, 6, 6, 6, 7, 7, 8, 8, 8, 9, 10, 11, 11, 11, 11, 12, 13, 13, 14, 15, 15, 15]

1. Create a list of words from the randomtext.txt.
   1. Open a portal to the randomtext.txt file.
   2. Create an empty list and name it.
   3. Create a loop to go through each line and remove the whitespaces on the right side of the lines.
   4. In the same loop, split the lines into a list of words.
   5. Construct another (embedded) loop to check if the word is already in the list, and if not, add it.
   6. When your list is complete, sort it in alphabetical order and print it.
2. Write a program to give the user the shortest, the longest and the average length of the messages in the mbox1.txt file (Content-Length).
   1. Start by creating an empty list and opening the file.
   2. Construct a loop to look for lines with ‘Content-Length’ in them.
   3. Split those lines into a list of words and create a variable for the second member of those lists (index position 1). The value needs to be an integer.
   4. Add the numbers to the empty list.
   5. Sort the list.
   6. Calculate the average by using sum() and len() functions.
   7. Print out the shortest, longest and average message lengths with an appropriate message.

**e.g. “The shortest message is 357 characters.”**

**“The longest message is 1000 characters.”**

**“The average message length is 500 characters”**

1. Create an empty list.
   1. Get some numbers from the user as integer input by creating a while True loop and add the numbers to the list.
   2. Construct a safety net so that if the user enters anything but an integer, the program will print out;
      1. the maximum number entered so far (max() function)
      2. the minimum number entered so far (min() function)
      3. the average of the numbers so far

with corresponding messages and end the loop (break statement).

1. Write a program to find out who the emails in mbox1.txt are sent by. Print them out, without the duplicates.

* Open a portal to mbox1.txt file.
* Create an empty list.
* Construct a loop to see whether there is a pattern of ‘From: ‘ in line and the line does not start with ‘Resent’ (to remove the lines in the file that start with the pattern ‘Resent-From:’
* Split the lines into words.
* Get the email part from the list of words (the last element in the list) by creating a new variable and using indexing.
* Create another variable and use indexing to get rid of greater than and less than signs around the email address strings.
* Append those emails in the empty list and sort them.
* Create another empty list to put the email addresses in without the duplicates and print it.