# LAB 04

## SUBMISSION INSTRUCTIONS

Submit 2 python files using the naming convention below (replace JaneDoe with your first and last name respectively):

- JaneDoe4_1.py - For question 1.
- JaneDoe4_2.py - For question 2.

## QUESTIONS

1. Download **input.txt** to your Python project and write a script to:
    a. Look for lines in **input.txt** that begin with **X-DSPAM-Confidence:**
    b. Send those lines to an output file named **output.txt**.
    c. Extract and add the floating-point number on the lines to a running total.
    d. When done processing **input.txt**, display the total and average values of the dspam confidence to 2 decimal places at the very end of **output.txt**.
    e. Your output file should have data identical to what you see in the **sample.txt** file I provided.

    Please note:
    - Write your program as a script (i.e. include the if __name__ == '__main__': block).
    - Use the input and output file names provided in the instructions.
    - Don't forget to close any files you open.

2. Using **input.txt**, write a script that:
    a. Create a csv file named **output.csv** whose first row contains the headers: Email, Subject, and Confidence.
    b. Go through the lines in **input.txt** looking for lines that begin with (please note the colons are part of the text)

```
To: source@collab.sakaiproject.org
From: stephen.marquard@uct.ac.za
Subject: [sakai] svn commit: r39772 - content/branches/sakai_2-5-x/content-impl/impl/src/java/org/sakaiproject
X-Content-Type-Outer-Envelope: text/plain; charset=UTF-8
X-Content-Type-Message-Body: text/plain; charset=UTF-8
Content-Type: text/plain; charset=UTF-8
X-DSPAM-Result: Innocent
X-DSPAM-Processed: Sat Jan  5 09:14:16 2008
X-DSPAM-Confidence: 0.8475
X-DSPAM-Probability: 0.0000
```

        i. From: - Extract the email address (e.g. stephen.marquard@uct.ac.za)
        ii. Subject: - Extract the commit number (e.g. r39772)
        iii. X-DSPAM-Confidence: - Extract the numeric value (e.g. 0.8475)

    c. Send the results obtained in (b) to **output.csv** for each occurrence of the 3 items in **input.txt**.
    d. When done, your csv file should have data identical to what you see in the **sample.csv** file I provided.

    Please note:
    - Write your program as a script (i.e. include the if __name__ == '__main__': block).
    - Use the input and csv file names provided in the instructions.
    - Don't forget to close any files you open.