

SSN College of Engineering

Department of Information Technology

UIT2201 — Programming and Data Structures

2022 – 2023

PSP Exercise — 01

April 11, 2023

-
- The purpose of this exercise is to practice PSP0 personal software development process
 - Instructions on PSP0 will be given separately
 - This PSP exercise is due by 10PM on April 17, 2023
 - Grace period may be given up to midnight of April 17, 2023
 - You can upload only one ZIP file
 - The naming convention is “<Your first name (first letter capital and all the other letters small)>-UIT2201-pp-ex-01.zip”
 - Judicious use of Python features and standard modules, version control using ‘git’, adhering to Python coding standards are expected

The purpose of this exercise is to get you started on the Personal Software process (PSP). Go through the PSP0 process scripts (uploaded as PSP-Assignment-Instructions.pdf) to understand the different phases of developing a software. You need to fill-up the required forms (uploaded as PSP-Forms.docx) during the development. Instructions for filling up these forms have been uploaded as ‘PSP-Forms-Instructions.pdf’.

1. Write a Python program that reads all the given text files and perform the following:
 - Prepares the set of words appearing in the given text files and stores them as “words.txt”
 - Prepares a bag of words that records how many times each word is appearing in the given text files and stores them as “words-histogram.txt”
2. Write another Python program that reads “words.txt” and “words-histogram.txt” files and does the following until ‘EOFError’ is raised:
 - Reads input text ‘prefix’ from the console and prints all the words that start with that ‘prefix’. The output words should be in the descending order of their appearance count. Your program should continue to do this until ‘EOFError’ is raised (which can be achieved by pressing ctrl-D)