Lab 9: Reading Material

Introduction to Python

Lab 9 addresses Python. So, you are required to:

1. Read Chapters 1 and 2 from Python tutorial
2. Browse through chapter 3 and try to get familiar with Python Interactive GUI.

Additional Reading - Python(not obligatory)

- Python v2.7.6 documentation
- Dive Into Python
- Wikibook

Shell Commands

- Plumbum library for integrating shell commands in python.
- awk: parse CSV files. working with CSV files, adding conditions.
- sed: replace values.
- sort: sort alphabetically.
- wc: word count.

Lab 9 - Python and Shell Commands

This lab may be done either solo or in pairs.

In this lab, you are going to use Python and shell commands to parse a CSV file and print some statistics about its content. The first part of this lab is about using shell commands in python (see: awk, sed, sort, uniq). The second part is about implementing a program to manage the Covid-19 vaccine distribution, it is all up to you!

Task 0

In this task, you are going to download a third-party library for integrating shell commands into Python and use it to extract a list of all the students in the file, and their number. You also need to extract a list of all error-codes (no repetitions) and their number.

Task 0a - Preparation

We are going to download a library that adds shell commands to python and use it during the lab. Do the following:

```
Download plumbum library:
$>wget
https://pypi.python.org/packages/50/15/f26f60e1bb82aabed7ff86f3fd297678404
7f9a291c63ac9019086a69559/plumbum-
1.6.3.tar.gz#md5=e0c588ba9271711fae3beb8c0511e8a9

Uncompress the file:
$>tar -xzvf plumbum-1.6.3.tar.gz

Change directory:
$> cd plumbum-1.6.3/
```

Your python code should be written in this directory.

Task 0b

The file to be parsed contains error-codes describing errors made by different students and how much to reduce for each errorcode. The format of the file is: `student\tab terror_code1:1|error_code2:1|error_code3:0.5...` where the number after each error_code is used to give partial reduction. 1 means full reduction. Any number less than 1, means a partial reduction for its relevant error_code.

```
Example:
Danny   no_README:1|wrong_file_name:1|code_repetition:0.5|no_task2:1
```

Use the following file grades_error-codes.

Write procedures in python that receive a grade file in the given format, see above, and calculates the following, one procedure per task:

- A list of all students mentioned in the file.
- The number of students mentioned in the file.
- A list of all error-codes mentioned in the file together with how many times each error-code was mentioned.
- The number of unique error-codes found in the file.

All these tasks must be done using shell commands in python, see: awk, sed, sort, uniq, wc. Each calculation must be a line of shell commands. Shell commands in python return a 'n' separated list of strings.

For example: in order to view the first field of every line of a file 'test', you can use `awk -F '\t' '{print $1}' test` where `'\t'` is the separator. In the reading material you can see several links that contain information and additional examples for using awk, sed, sort, wc, and uniq.

Task1 and Task2 can be done either with shell commands in python or regular python commands.

Task 1

You are about to implement the program managing the Covid-19 vaccine distribution, it is all up to you! You need to provide data and statistics. You have the liberty to implement the tasks as console arguments, as a menu, or to add a specified GUI (see Task 3).

The first file is Vaccine_Distribution which stores the data regarding the vaccines in the following format:`company_name\tab date_of_arrival,number_of_vaccines,dose`

The second file is Cities which stores the data regarding the cities in Israel in the following format:`city\tab population`

Task 1a

Calculate the number of vaccines to be distributed by each company and output it to *vaccine.stats* in the following format: `company_name|number_of_vaccines`.

Task 1b

Calculate the number of vaccines to be in Israel after a specific date, add the option to filter by a company.

Task 1c

Draw a histogram of the number of vaccines to be in Israel by date, see this page.

If you receive an error when trying to import matplotlib, run the following command in the shell: pip install –user matplotlib

Task 2

Now, that you know everything about the vaccine you can start to distribute it around the country.

Task 2a

Calculate the number of people that can get a vaccine according to the recommended dosage by each company, output it to *vaccine.stats* in the following format: `company_name|number_of_vaccines`.

Task2b

For a specific date, calculate which cities can get a vaccine such that all of the population is covered. This is similar to SubsetSum which is NP-hard, you may calculate it accurately (long computation..), you may use heuristics or any other reasonable idea.

Task 3 - Bonus task

Add GUI. You may use this package tkinter

Submission