Bash One-Liners

Created by Jason Madar
Modified by Ivan Wong

Background

Many Linux tasks involve manipulating text files (configuration files, log files, text-based content, etc.). Linux shell commands are designed to work with text streams, and are very convenient in dealing with text based data.

In this lab, you will be practicing using linux commands to extract and transform data from a local csv file and from an external web site.

You will use the data file from the BC open data website

https://catalogue.data.gov.bc.ca/dataset/6e317257-54fe-44d1-89a6-1fcd81f5a484/resource/f991757f-26f3-459a-96c4-66b3dc3f5d10/download/tui2_international_tuition_fees_at_public_post_secondary_institutions_by_economic_development_re.csv

which is a csv file containing international tuition for all BC post-secondary institutions. If you are interested, you can find additional information for this dataset at https://open.canada.ca/data/en/dataset/6e317257-54fe-44d1-89a6-1fcd81f5a484

Preparation

1. Do this assignment in your AWS instance.
2. Start by cloning your git repository to you home directory. cd into the newly created directory. [Research on how to clone the repository to your Linux.]
3. Download the tuition fees csv file into the newly created home for this assignment. You can use either curl or wget for this, but make sure you rename data file to tuition.csv
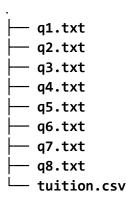
Steps

For questions 1 to 8, I'm looking for a one-line bash command. You will put each bash command into its own text file. For example, answer to question 1 will be put into a text file called q1.txt

1. A command to show how many lines (just the number of lines) are in the file tuition.csv.
2. Show only the data lines in the file (skip the header lines)
3. Show the first and last data line in the dataset (hint: you need to execute 2 commands on the same line)
4. Output all unique institution names

5. Find all data lines that contain either 'Capilano University' or 'Langara College'
6. Find all data lines for the 2019/20 year, and show only school names and tuition amounts
7. Following from question 6, sort the output in ascending order by college name
8. (Challenging!) Following from question 6, sort the output in ascending order by tuition amount, only output the insitution names

Hand-in

At the end of the lab, you will have the following additional files in your assignment directory.

```
.
├── q1.txt
├── q2.txt
├── q3.txt
├── q4.txt
├── q5.txt
├── q6.txt
├── q7.txt
├── q8.txt
└── tuition.csv
```

Make sure you run **pytest** to check your score. [Research on how to run use **pytest**.]

Once you are satisified with your score, run the following commands to submit your assignment:

cat q1.txt q2.txt q3.txt q4.txt q5.txt q6.txt q7.txt q8.txt

Copy and paste the output to the answer area for submission.