

Worked on this assignment with Ji, Ryan C, Steve Y, Anthony

### **Part One: Shell Commands**

1. Assume all of the directories below are valid directories (e.g., they exist and we have access to them). What would the following print out?

```
cd /usr/src/INFO3401/./homework/./problem1
```

Pwd

```
cd /usrsrc / homework / problem1
```

2. What set of commands would you use to achieve the following:

You've found yourself in a situation where your python program is trying to parse files in a directory and is throwing an error. You think it might be a permissions issue (i.e., you don't have permission to access the files in that directory) and need to verify the permissions on all files in that directory. The directory is in your home directory under the "datasets/activedata" directory. Make no assumptions about what directory you are currently working in.

```
cd ~/ datasets / activedata
```

```
pwd
```

```
use sudo cd ~ / datasets / activedata
```

**OR**

```
cd ~
```

```
cd datasets
```

```
ls -l activedata
```

```
pwd
```

3. Briefly describe what the following set of commands would achieve. What process would happen and what would be printed to the command line?

```
cd ~ - Change to the home directory
```

**mkdir ./problem\_set\_1** - Make a new directory named 'problem\_set\_1' in the ./ portion. At this point there is nothing printed out.

**touch submission.txt** - create a blank text file named submission.txt in the home directory. Command line will not print anything out.

**cd ..** - Go to the parent directory, which is the 'users' directory. The command line wont print anything

**Pwd** - print the working directory

**First, change our directory to the home directory, create new directory "problem\_set\_1" then create an empty txt file named "submission.txt" inside of the 'problem\_set\_1' directory, go back to parent directory print working directory.**

4. What set of commands would you use to achieve the following:

Copy a document called config.txt from your home directory to the root directory. Then, create a new directory in the root called preferences. Within preferences, copy prefs.txt from a matching directory called preferences in your home directory. Finally, determine that the start and end of the config.txt document in both root and home match.

**cd/** - change your directory to root directory

**Cp ~/config.csv** - start at home directory: copy the config.txt to the root directory

**Mkdir /preferences** - In the root directory, making a directory named 'Preferences'

**Cp ~/preferences/prefs.txt** - Copying pefs.txt from the root directory to the preferences directory

**head/config.txt** - find out the start of the config.txt in the root directory

**tail/config.txt** - find out the end of the config.txt in the root directory

**tail~/confi.txt** - find out the end of the config.txt in the home tirectory

**head~/config.txt** - find out the start of the config.txt in the root directory

5. You have accidentally moved a file (~/Documents/datafile.csv) to your current directory rather than copying it. What command did you use to do that? What command should you have used to do that? What command might you use to undo it?

**mv~/Documents/datafile.csv**

**Mv datafile.csv~/Documents/**

6. Download and unzip [asciify-master.zip](#) from Canvas. Then, navigate to the directory and use the command line to run `asciify.py`. What does this script do? Include a screenshot of your results.

It visualizes what looks like the github logo

[illegible]

7. Install the Delorean Python package. What command did you use? Then, from the command line, launch Python. Use it to execute the following commands:

```
from delorean import Delorean
EST = "US/Eastern"
d = Delorean(timezone=EST)

print(d)
```

What prints to the command line?

**Delorean(datetime=datetime.datetime(2020, 1, 29, 12, 38, 25, 694604),  
timezone='US/Eastern')**

8. We can scrape webpage content or download collections of data files using wget. NASA, the USDA, and the NCBI all recommend using wget to download data from their repositories because it is faster and allows you to download data in bulk. We'll experiment with that on a smaller scale: use the command line to download data on the World Bank's budgetary expenditures at

<https://finances.worldbank.org/api/views/yu93-ayrw/rows.csv?accessType=DOWNLOAD>

[\(Links to an external site.\)](#)

[Links to an external site.](#)

. What command did you use? Make sure to include this file as part of your submission.

**Wget**

<https://finances.worldbank.org/api/views/yu93-ayrw/rows.csv?accessType=DOWNLOAD>

9. The above budgetary data contains both individual items and aggregated totals across different sectors. Let's distill the data down to different aggregated totals. In this dataset, you can use the keyword "Total" to identify only those rows containing aggregate budget items. Use the command line to identify these rows. What command did you use? How many rows did this find (hint: the -c option can be helpful here)?

**Use grep Total World\_Bank\_Expenditures\_by\_Organizational\_Unit.csv to identify these rows. It prints out random lines that include 'Total' inside**

**Use `grep Total World_Bank_Expenditures_by_Organizational_Unit.csv -c` to count the total number of lines that have 'Total'**

**There are 33 lines**

10. Now, print the lines containing the word "Total" to a file to "distilledExpenditures.csv". What command did you use? Include this file as part of your submission.

**`grep Total World_Bank_Expenditures_by_Organizational_Unit.csv > distilledExpenditures.csv`**