INFO 3401
Jason Kibozi-Yocka

Problem Set Two

Note for Grader: All of my answers are coming from a Windows point of view.

Part One: Shell Commands

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

   cd /usrs/INFO3401/ ../homework/ ./problem1
   pwd

   cd /users/INFO3401/..homework/./problem1
   cd

   Answer:
   It would print out the current dir, which is 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.

   Answer:
   [go to home dir]
   cd /datasets/activedata
   for %i in (*) do icacls %i

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 ~
   mkdir ./problem_set_1
   touch submission.txt
   cd ..
   pwd

   [go to home dir]
   mkkdir ./problem_set_1
   echo "" submission.txt
   cd ..
   cd

   Answer:
   go to home dir (Windows has no equivalent)
   create a dir named problem_set_1
   create a text file called submission
   go to previous dir (home)
   print the current dir (home)

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.

Answer:
[go to home dir]
copy config.txt "C:/"
cd /
mkdir preferences
[go to home dir]
cd preferences
copy prefs.txt "C:/preferences"

(this next part is not possible in Command Prompt, so I had to use PowerShell)

[go to home dir]
compare-object (get-content "./preferences/config.txt") (get-content "C:/preferences/config.txt")

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?

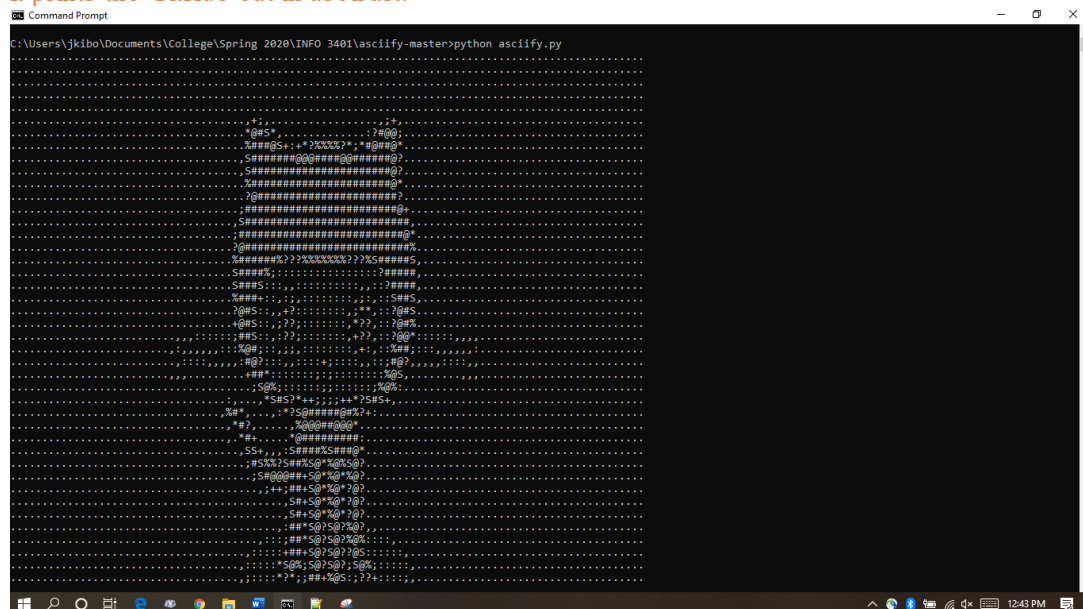Might Have Used:
move "~/Documents" "./" "datafile.csv" /mov

Should Have Used:
xcopy "~/Documents/datafile.csv" ./

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.

Answer:
It prints the GitHub cat in ascii art.

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?

Installation:
pip install delorean

Prints Out:
It prints the current datetime in the following format: (year, month, day, hour, minute, second, unknown 6-digit number)

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.

Answer:
wget https://finances.worldbank.org/api/views/yu93-ayrw/rows.csv

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)?

Answer:
find /c "Total" rows.csv
(This prints out the number of lines that have the substring "Total" in them.)

OR

Alternative Answer:
find /n "Total" rows.csv
(This prints all the lines that have the substring "Total" in them, preceded by the line number.)

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.

Answer:
find "Total" rows.csv > distilledExpenditures.csv