CSCE 3600: Systems Programming

Minor Assignment 1 – sed and gawk

Due: 11:59 PM on Friday, February 18, 2022

PROGRAM DESCRIPTION:

In this assignment, you will write `sed` and `gawk` commands to accomplish certain requested functionality. Given the many powerful features of `sed` and `gawk`, you are provided with links to manuals and tutorials for `sed` and `gawk` to assist you in completing this assignment.

Resources (links to manuals and tutorials for `sed` and `gawk`):
https://www.gnu.org/software/gawk/manual/gawk.html

https://www.gnu.org/software/sed/manual/sed.html

https://www.grymoire.com/Unix/Sed.html

https://tldp.org/LDP/Bash-Beginners-Guide/html/sect_06_01.html

Using sed

a) Consider the following file called **`dates.txt`** containing some dates in US format, some of them palindromes (i.e., words that read the same backwards and forwards) and some not:

```
02/02/2020
03/02/2030
12/02/2021
06/19/1960
01/10/2010
03/07/2100
```

Write a one-line `sed` command that prints out only the lines containing date palindromes so that after running the appropriate `sed` command, the following would be output to the terminal:

```
02/02/2020
03/02/2030
12/02/2021
```

You may write the `sed` command-line in the space provided below and ensure that this Word document (with your solutions) is submitted to Canvas.

Answer: `sed -E -n '/([0-9])([0-9])\/([0-9])([0-9])\/\4\3\2\1/p' dates.txt`

b) A company has a set of letters that need to be modified. As an example consider the following file called **`letter.txt`** containing a company letter:

```
ABC Corp
1234 7 Oaks Lane
Denton; TX 76509-4321
(682) 543-7890

February 10th, 2022

Ms. Sue Jones
1515 Mockingbird Lane
Apartment @701
Denton; TX 76205


Dear Mrs. Jones,

This letter is to inform you that you have won the super coder
contest. To Claim your winnings, you may visit our office or
call our claim hotline (654) 817-4321. Mention your claim
number @687.

Sincerely,
Boris Lane
```

Write a complete `sed` script called **`minor1.sed`** that will modify the company's letters in the following ways. Note that this letter is an example. Your program will be tested against several similar letters. Note that spacing may be different in each letter.

1. *Preprocessing*:
    a. Change the area code of the phone number in the heading to 800.
    b. Remove the parentheses from each phone number.
    c. Add a dash between the area code and the rest of the phone number (i.e. 789 123-4567 would become 789-123-4567)
    d. Change the zip+4 code in the heading from 4321 to 1234. The zip+4 is the last 4 digits of the zip code after the dash.
    e. Change Boris' last name from Lane to Kent, since she became married.
    f. Change Lane in the addresses to Ln.
2. *Transposition* (i.e., rearrange the order of individual characters):

a. Swap the area code and exchange for the phone number in the body of the letter.

3. *Substitution* (i.e., replace characters by other characters):

   a. Substitute '#' for '@' to correct a type where '@' were used instead of '#' and change ';' to ',' to correct the typos in the addresses.

4. *Append extra data*:

   a. Append a p.s. clause after the signature (the end of the letter). It should read

   `p.s. your winnings must be claimed by February 29th.`

In the example, your `sed` script should print the following:

```
$ sed -r -f minor1.sed letter.txt
ABC Corp
1234 7 Oaks Ln
Denton, TX 76509-1234
800-543-7890

February 10th, 2022

Ms. Sue Jones
1515 Mockingbird Ln
Apartment #701
Denton, TX 76205


Dear Mrs. Jones,

This letter is to inform you that you have won the super coder
contest. To Claim your winnings, you may visit our office or
call our claim hotline 817-654-4321.
Mention your claim number #687.

Sincerely,
Boris Kent
p.s. your winnings must be claimed by February 29th.
```

This sed script file will be submitted to Canvas.

Using gawk

a) Consider a file containing grade test grades for students called **`grades.txt`** file:

```
Last,First,A/I,Test1,Test2,Test3
Smith,John,A,90,100,99
Ballard,Sue,I,15,,50
Clark,Sally,A,100,50,
Koen,Jack,I,90,100,99
Ball,Lucy,A,100,100,100
Bently,Mark,A,100,,100
```

The file is comma delimited. The first line is a header. The records contain Last Name, First Name, Active/Inactive file and several test scores. Ignore the heading line. For each Active student calculate their average grade for each student printing out the Students First Name followed by Last Name and average score. Display a heading at the beginning with a name for the report and headings for the columns of Name and Average.

After all students have been processed display the number of active students and the average class test score.

Format the output so that only 1 digit after the decimal point is shown and the averages scores line up on the decimal points.

In this file, for example, the `gawk` program should print the following:

```
$ gawk -f minor1.gawk grades.txt
Student Average Test Scores
Name            Average
John Smith       96.3
Sally Clark      50.0
Lucy Ball       100.0
Mark Bently      66.7
4 Students with average score of 78.2
```

Formatting properly in columns as shown is required. This `gawk` program file will be submitted to Canvas.

b) Consider the following file called **`gasoline.txt`**:

```
Station,Regular,Premium
Exxon,2.39,2.50
Mobile,2.19,2.25
Shell,3.19,3.29
CircleK,3.01,3.15
QT,2.11,2.15
```

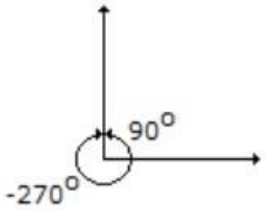
Each field in this file is separated by a comma and each record is separated by a newline character. For this file, you will write a one-line `gawk` command-line to filter records where the premium price is less than $3 and the premium price difference between regular and premium is less than 10 cents. Specifically, you will print out only the station name and the price for premium. Add a $ in front of the price. You may write the `gawk` command-line in the space provided below and ensure that this Word document (with your solutions) is submitted to Canvas.

Answer:
```
gawk -F',' ' $3 < 3.0 && ($3 - $2) < 0.1 {print $1 " $" $3}' gasoline.txt
```

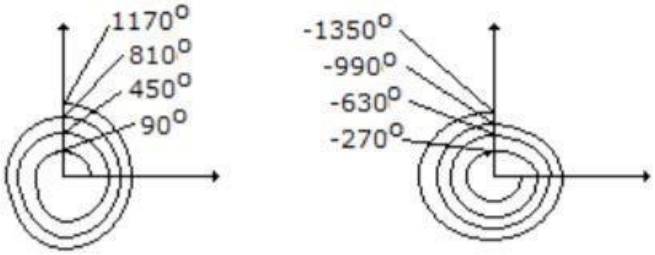
c) Consider a list of angles in degrees, such as the following angles.txt file:

```
Degrees
135
90
1860
-45
0
-270
1170
-180
-795
-630
```

When two angles have the same initial and terminal sides, they are called coterminal. This means that angles of 90° and -270° are coterminal as the following figure shows:



In fact, all of the following angles shown in the figure below are coterminal.



We want to make sure that all angle values in degrees are between 0° ≤ x < 360°. For any file containing a list of angles in degrees with the `Degrees` header line at the top (as in **`angles.txt`**), write a complete **`gawk`** program that prints out the original value of the angle

in the file as well as its coterminal value between 0° ≤ x < 360°. In this file, for example, the `gawk` program should print the following:

```
$ gawk -f minor1.gawk angles.txt
135       135
90        90
1860      60
-45       315
-270      90
1170      90
-180      180
-795      285
-630      90
```

Formatting properly in columns as shown is required. This `gawk` program file will be submitted to Canvas.

REQUIREMENTS:

- Your `sed` script and `gawk` program files should **include your name and EUID at the top of the file**. No other comments are needed in these files.
- Your programs should work correctly with any file formatted as per the directions.
- For the `gawk` and `sed` commands, test out your results on real files on our CSE machines (e.g., `cse01`, `cse02`, …, `cse06`), to make sure that they indeed work.
- Your solution to the one-line `sed` script and `gawk` program can be typed (or copied and pasted) to this document and will be submitted to Canvas.
- Your program will be graded based largely on whether it works correctly on the CSE machines (e.g., `cse01`, `cse02`, …, `cse06`), so you should make sure that your program runs on a CSE machine. Please include any special instructions required to run your `sed` script and `gawk` program.
- This is an individual programming assignment that must be the sole work of the individual student. Any instance of academic dishonesty will result in a grade of "F" for the course, along with a report filed into the Academic Integrity Database.

SUBMISSION:

- You will electronically submit this file with your typed one-line solutions for `sed` and `gawk` along with your `sed` script **`minor1.sed`** and `gawk` program **`minor1.gawk`** to the **Minor 1** dropbox in Canvas by the due date and time.