



## CS251\_Inlab3 - SED and AWK

Updated automatically every 5 minutes

# CS 251 Inlab3 - SED and AWK

## Useful instructions for the Assignment

1. **Remember, you need to change the permission of the .sh file before executing it:**  
`chmod u+x <script file> or chmod 777 <script file>`
2. **Follow the naming convention of the submission files strictly.**
3. **Remember to follow the exact output formats specified. Extra characters will lead to incorrect evaluation by the autograder.**
4. **Inside the script if you use any temporary file, write a command to remove it after the operation within that script only.**
5. **Do the lab inside cs251 Docker container ONLY.**



## CS251\_Inlab3 - SED and AWK

Updated automatically every 5 minutes

repeated words. Lets try using AWK for the same, in this question, you are supposed to:

- a. Find the number of unique words present in a “**data**” directory. [ **5 points** ]
- b. Report each **word** along with its **frequency** for every text file in a “**data**” directory in **lexicographical order**. The script will be run from the same location where the “**data**” directory is located. [ **10 points** ]

Note: we have provided a sample data directory. Your codes will be autograded over various text files kept in the **data** directory. The data directory should be a command line argument for your bash script

Refer sampleq1a\_output.txt and sampleq1b\_output.txt for the output of the scripts.

Your script will be checked by executing the command:

```
./q1a_awk.sh <data_dir>
./q1b_awk.sh <data_dir>
```

2. Create a function which takes in input a set of numbers from a text file in base **8+2\* (line\_number%3)** and returns the same number in **base 10**. Note that the digits will be space separated on each line. For digits greater than 9, **lowercase alphabets** will be used. [ **20 points** ]

Example (sampleq2\_input.txt also shared on moodle)



## CS251\_Inlab3 - SED and AWK

Updated automatically every 5 minutes

```
251-TA/sed-awk/inlab$ ./Q2_awk.sh sampleq2_input.txt
1234
34
8
25
139
```

**SED [ 10 points ]**

3. Refer to below *sample.txt*.

**sample.txt**

From then on, Madame Loisel knew the horrible life of the very poor. But she played her part heroically. The dreadful debt must be paid. She would pay it. They dismissed their maid; they changed their lodgings; they rented a garret under the roof.

She came to know the drudgery of housework, the odious labors of the kitchen. She washed the dishes, staining her rosy nails on greasy pots and the bottoms of pans. She washed the dirty linen, the shirts and the dishcloths, which she hung to dry on a line; she carried the garbage down to the street every morning, and carried up the water, stopping at each landing to catch her breath. And, dressed like a commoner, she went to the fruiterer's, the grocer's, the butcher's, her basket on her arm, bargaining, insulted, fighting over every miserable sou.

Each month they had to pay some notes, renew others, get more time.



## CS251\_Inlab3 - SED and AWK

Updated automatically every 5 minutes

Madame Loisel looked old now. She had become strong, hard and rough like all women of impoverished households. With hair half combed, with skirts awry, and reddened hands, she talked loudly as she washed the floor with great swishes of water. But sometimes, when her husband was at the office, she sat down near the window and thought of that evening at the ball so long ago, when she had been so beautiful and so admired.

What would have happened if she had not lost that necklace? Who knows, who knows? How strange life is, how fickle! How little is needed for one to be ruined or saved.

Task is that the end of every paragraph should be replaced with :) and the period at the end of every line should be replaced with \$.

Text file will be provided as input argument to the q3\_sed.sh file, the output should be printed on the Terminal.

***Solution for first paragraph will look like this:***

From then on, Madame Loisel knew the horrible life of the very poor\$ But she played her part heroically\$ The dreadful debt must be paid\$ She would pay it\$ They dismissed their maid; they changed their lodgings; they rented a garret under the roof:)



## CS251\_Inlab3 - SED and AWK

Updated automatically every 5 minutes

**<roll\_no>\_inlab3.** Compress it into a tarball using the following command:

```
tar -cvzf <roll_no>_inlab3.tar.gz <roll_no>_inlab3
```

**The following is what your submission directory must look like. Please follow the submission format strictly.**

```
<roll_no>_inlab3
├── q1a_awk.sh
├── q1b_awk.sh
├── q2_awk.sh
├── q3_sed.sh
└── references.txt
```

**Note:** references.txt should contain the exact links of webpages you referred to (apart from the ones linked in the resources shared before).

***... May the Bash be with you! ...***