

Assignment 2 - Shell Scripting

CS253 - 2024

March 1, 2024

Through this assignment, you will learn how to automate a complex process using shell scripting. For this assignment, you can use any shell command. We encourage you to use `sed`, `grep`, and `awk` commands in a shell script to do the desired tasks. You can use `man` pages to know more about these commands. You will also need to use output redirection (`>`, `>>`, or some equivalent method) to write to a file.

Problem Statement:

[1] Write a shell script and name it as ***main.sh***, which will take exactly two filenames as arguments. If exactly two names are not provided, the script prints a usage message and exits, otherwise the script does what is described next. The usage message tells the user how to use the script.

[2] Till this step, you have provided two filenames to the script. The first file is an input file, and the second one is an output file. You need to check whether the input file exists or not. If the input file does not exist (How do you test that in bash?), the script generates an appropriate error message and exits. If the input file exists, then proceed to the next step.

[3] You are given a CSV file named ***data.csv*** with the following structure: Name, Age, City, Salary. The CSV file contains headers and data. Use this file as the input file. Now, your script should print the unique cities from the input file to the output file (the second argument).

[4] Next, Find the top 3 individuals with the highest salary. Save their Name, Age, City, and Salary to the output file. Be careful that the script should not overwrite the output file.

[5] The script should compute the average salary for each City in the dataset and save it to the output file (second argument). Be careful that the script should not overwrite the output file. The output should be saved as follows:

```
City: city1, Salary: average salary of city1
City: city2, Salary: average salary of city2
...
```

[6] The script should identify individuals with a salary above the overall average salary. Save their Name, Age, City, and Salary to the output file. Be careful that the script should not overwrite the output file.

Submission Instructions:

- Create a single shell script file named *main.sh*.
- You can use one or more `awk` script files, or any other auxiliary file if needed. Submit all the source files in the zipped format. The directory must have *main.sh*, any auxiliary `awk` or other script files you used.
- Comment each section with a brief explanation of what the code is doing.
- Assume the CSV file is present in the same directory as the script.
- There is no need to submit data.csv or any generated output file.
- You are given a sample output file. Your output file should follow the format of this file strictly.

Contact: Please contact Shatroopa Saxena (ssaxena@iitk.ac.in) in case you have any questions.