IIT PALAKKAD

Indian Institute of Technology Palakkad
Department of Computer Science and Engineering
Programming Lab - Jul to Nov 2020
30 September 2020

Instructions:

- Run the command `$ make` to generate an executable `generator`. This program will generate input files for this assignment.
- Run the generated executable named `generator` with appropriate command line arguments to generate a master csv file. This master file should be used as input data that your shell script should manipulate
- The data contained in the csv file is of the format given below

```
file.csv
============
1118010099,CS3010,23
1218010099,EE3010,39
1017010098,CE3010,12
1116010079,CS3030,40
```

  The first column in each of these files contain names of students, the second column contains the subject, and the third column contains their marks in the corresponding subjects.
- *Read the entire question, identify a high-level strategy to tackle this problem, break each of it down to smaller sub-problems, and write the outline of these steps as comments inside your shell script. Upload the completed script.sh to moodle.*

Write a single BASH script named ROLLNO.sh (example 152002099.sh) that does all of the following:
NOTE: Your final script should be able to process csv files with 100 million (or more) entries in reasonable time (< 10 mins).

1. Generate a CSV using the `generator` program (whose source file is given with this assignment). The name of the output file thus generated (example `master.csv`) should be taken as a command-line argument to the script.

2. The master file thus generated (say `master.csv`) should be split into course-wise files. For example, if the user wishes to see the records of course `CS3010`, then only the student records corresponding to their performance in the course `CS3010` should be placed into the file `CS3010.csv`.

a. The course-wise file thus generated should be placed inside a folder named `courses/`.
b. If `all` is entered in place of course code by the user, the course-wise files corresponding to all the course codes in the master file should be generated and placed in the folder `courses/`.
c. The course-wise csv files should not have any duplicates.

NOTE: If a roll number is repeated multiple times in a course, only the maximum of these marks should be retained.

3. The master file (say `master.csv`) should be split into branch-wise csv files. For example, if the user wishes to see the records of the CS branch, then only the records of students in the CS branch should be into the file `CS.csv`. The output csv file should be to the following specifications:
   a. The header row should contain course code of all courses offered by the department.
   b. There should be only one entry per student. The marks obtained by the student in a particular course should be shown in the column corresponding to the course code. If a student has not credited a particular course, then that column should not have any value in it.
   c. There should not be any duplicate entries. In case a roll number has the same course code, but different marks, then only maximum marks should be retained.

   NOTE: You can look at the source code of the generator (generator.c) to see how you can identify the branch of a student from their roll number.

4. The command-line options that your script should take is as follows:

`-s` or `--setup`	Create the executable `generator`
`-g` or `--generate`	Run the executable `generator` to generate the master file. You can assume that the default output filename is master.csv, but the user should be able to change it (with additional options if required). Example: `-g 100` should generate master.csv with 100 lines
`-c` or `--course`	Generate the course files for the course mentioned Example: `-c CS3010` generates `CS3010.csv` `-all` generates files for all courses and generate separate files for each course such as `CS3010.csv, EE1020.csv`, etc..
`-b` or `--branch`	Generate the branch files for the branch mentioned Example: `-b EE` generates `EE.csv` `-all` generates files for all branches `CE, CS, EE & ME`