CS384 2020 End Sem - Class Group Allocation

Mayank Agarwal

November 28, 2020

Assignment Given on 28th Nov 2020, 1730 hrs
Assignment Deadline on 30th Nov 2020, 1730 hrs

You are given Btech_2020_master_data.csv file that consists of 417 students who were admitted to IITP. The names and emails have been changed but the roll numbers are kept intact.

IITP divides the students into various groups for ease of practicals. So lets say, if the 417 students are divided into 12 groups, each group will have an average of 417/12 = 34.75 approx 35 students. Now each group of 35 students are permuted such that in each group, the students from each branch have a minimum representation (typically, **floor**(branch strength/total groups), eg., there are 40 students in MM, so if wee take 12 groups, each group should have at least 3 MM students.

Given the Btech_2020_master_data.csv that contains 417 students data info, you need to do the following tasks:

1. Read the roll numbers and extract unique branches along with their batch strength in a csv file. The csv file will have the following format branch_strength.csv BRANCH_CODE, STRENGTH The entries must be sorted by STRENGTH, if there is a clash, it should be branch code.

2. Make individual csvs depending on the branches (should be auto-generated) that you find in step 1. I am showing here for one branch, you need to follow the nomenclature for others CS.csv [Branch name in CAPS, extension lower case]
   Roll, **Name**, **Email** (basically all the 3 cols that are there in the input file)
   Entries must be sorted in ascending order by roll num

3. The input needs to be taken from the user regarding the number of groups. The total number of students needs to be determined automatically by reading the file "Btech_2020_master_data.csv". The logic for the division of the groups was explained on webex already. You can refer to the "Logic.xlsx" for the logic.
   In short the process is to read the "branch_strength.csv" made in step 1, get the branch strength, divide by the number of groups inputted in step 3, and then take a floor value always, and uniformly distribute among all the **n** input groups. Make a list of left over candidates along side. After all the groups are filled up by using the floor value, take the count of the left over(don't sort left over count, just have them in the same order as you have in Step 1, branch strength wise), and distribute one by one in each of the **n** groups till the left over students are filled up. For n groups there be **n** csv files: Group_G01.csv, Group_G02.csv, Group_G0n.csv, ensure that the group file names are padded with 2 digits intergers, so if we have 12 groups, we will have Group_G01.csv, Group_G02.csv, .., Group_G12.csv

4. Finally there should be statistics file that should tell us about the group info of each group. See "stats_grouping.csv"

So, if your "Btech_2020_master_data.csv" has **n** branches and we have **m** groups, so there would be in the output, **n** csv files for each branch, having only those branch students, **m** group csvs and a stats_grouping.csv file displaying the statistics. So for **n** branches, **m** groups, we will have **n+m+1** CSV files.

Formatting: BRANCH code must be always **CAPITALS** where-ever you are using.