

Assignment 1 - 11520G Data Captures and Preparations

Description

The group for performing this project consisted of four members whose names and Student IDs are mentioned below (by the order of Registration sheet). We represent ourselves by the name “The Data Cranes” as the team who collects and prepares mountains of data through our extended skills and capabilities.

- 1) Priyank Sharma - u3225285
- 2) Anu Maria George - u3228397
- 3) Ambika Huluse Kapanaiiah - u3227622
- 4) Yingqi Chen - u3215119

As none of the members were physically available in Canberra to perform the assignment (due to international border issues, etc.) our project was performed on the pre-collected dataset that was provided on UC (Part 2. Data Preparation and Analysis onwards). Canvas by the professor/unit convenor. The group consisted of members with varying levels of programming experience with a maximum of members programming as beginners. To make this as a learning opportunity we as a group decided that each of us will execute till certain portion of the questionnaire (say point 4 till 18th March then till point 6 till 20th March) so that everyone gets a learning opportunity and gets to try execute the code in order to understand the assignment entirely and not just portions of it. The idea was to also come up with new ideas in case one person got stuck or if some logic made more sense than the other. We also realised that assigning parts of the assignment to a specific member may not be productive as it would consume a lot of time and the next member would be dependent on the previous member to start the work which would further delay getting it done on time. This is why we took the other more open approach.

The R script was maintained on a google drive for ease of access to everyone. We evaluated our progress on private group chat and on zoom call each day so as to understand how we as a group are progressing with getting the desired output. By the end of week 1(21st March), Ambika (u3227622) managed to execute and obtain the most desired output without any hard coding or shortcuts by using for loops and other base R utilities effectively which we were still struggling to learn. Since then, the group was trying to take hurdles in that code one at a time with Ambika being our leader and mentor, communicating the drawbacks or hurdles in executing the code and the group trying to find solutions for it. The group met each day to discuss progress and hurdles with Ambika conducting the presentations. We would then decide way forward and short-term goals for each day.

Once the code was completed the group went on to search for ways to enhance the plots and parts of code and execute it on our machines to solve any recurring errors or inaccuracies.

Note: Instructions to execute the code delivered in zip file:

Once the zip file is extracted into any folder, double click on the R file and select the complete code to execute.

As the input data(csv) files are included along with the code in a subfolder called “InputData” under the main folder “TheDataCranes”, the code should execute without any changes to the folder structure after extracting.