Assessment for MOM - Data Analyst Intern, Sensing & Analytics (LRWD)

Hello! Thank you for applying for the role of Data Analyst Intern, Sensing & Analytics, OPPD (LRWD) in Ministry of Manpower. You are required to complete all the questions in Section 1 for this assessment.

**Section 1**

You would have received a Access\_data.zip file with 22 datasets in .csv format and 1 data dictionary in .docx format. The datasets are sample of the mock-data on individuals accessing two office sites of Company ABC, consisting of:

1. When
2. Profile
3. Dept
4. CardNum

You can assume that the total number of staff in the company is 2000 and the data is extracted from the company’s building access system. An individual can tap in and out several times within the same day. When the individual first clock in, that would be the earliest time slot and the only record you will base off the analysis. (You can also state your other assumptions if need be.)

Question 1.1

Write code preferably in R or Python to process and organise the raw data (‘Access\_data.zip’) to make it suitable for analysis. Identify and resolve the data quality issues in the raw data, if any. (The created code should allow user to efficiently and easily run it to ingest additional datasets of different period, beyond the given sample.)

Question 1.2

Using the prepared data from question 1.1, create dashboard(s) for end-users to self-serve to report and make decisions on the following issues on a regular basis.

1. The company’s target is to have more than 80% of staff working from home at any time.
2. For those staff who are not able to work from home, the company wants to know how many days per week are they coming in and to stagger their work hours and ensure an even proportion are coming in at the three different time slots. The proportion should roughly be 1/3 for each slot.
3. 0700 to 0829
4. 0830 to 0929
5. 0930 to 1030

1. For individuals issued with Temp or Visitor Pass, the company wants to know the number at each office site.

Please upload your findings, responses and code to **your** Google Drive or Dropbox, and send us the link as part of your answer by 4 April (Thursday), 6pm (within 5 days from date of releasing the questions).

* Your code (with code comment) in ipynb or Rmd (html), output, Tableau twbx, etc. in response to the questions

Thank you, and good luck!