Data Science for Design  
IDEAL-How are people using our feedback?

**Introduction**

The data provided by data owner are the usage of an app, IDEAL and this app is used to collect feedback of energy use in people’s households, there are two csv files and show the information about new features and users’ login time from August 2016 to October 2017. The aim of this project is help data owners to improve the app, so the analysis should concentrate on finding target audience and their patterns of using app.

In first csv files, persondetails.csv shows information of releasing new features and it is including WEEK\_VIEW, MINUTE\_VIEW, ENHANCED\_VIEW, COMPARE\_CHART, DUAL\_CHART, and HEATING\_INTERVENTION respectively. Besides, some participants receive other features such as MINUTE\_VIEW and ENHANCED\_VIEW, the other people in the control group only receive WEEK\_VIEW. Developers often release new features around the 14th of each month to participants if they are not in the control group. This csv file also shows the age and gender of participants, the category primaryparticipant represents user is primary participant, the main account holder in each house or not. The value of primary participants shows 1, and the other value shows 0.

The second csv file, sessionlogs.csv, shows the times of participants’ login. The category time indicates the time of users log in to system, the session\_duration shows the time of the initial login until the last interaction with the app. The 0 seconds long represents that user log in and do not touch anything, also, developers assume if participants do not interact with app more than five minutes, they end the record.

**Process of the project/analyse data**

(我覺得這兩部分好像講起來差不多，你再看看要拆分還是混在一起寫)

For app developers, the definition of target audience is essential, it can help managers to develop marketing strategy. In persondetails.csv, the image of target audience and their preference can be related to the information of gender and age. There are 226 participants used the app and female participants are more than male. Male users take 46% and female users are 54%. Moreover, participants aged 30-34 occupies the largest number of all participants.

The persondetails.csv also shows the details of new features.

這部份交給你:D

Sessionlogs.csv indicates the information of login, most participant log in duration is below 500 second (8 minute 20 second) and there are only very few people logged in over 2000 second (33 minute 20 second). 簡述一下你怎麼算的? 用了那個spilt…etc

The login time can be counted by person personid in sessionlogs.csv, the user largest logged in times 265, the average log in time is 24, and the smallest log in time is 1. For the information of the peak of login, first, we sort the data by year because this app started to record data from August 2016 to October 2017 and the data cross two different year 2016 and 2017, it does not have the record of single year. According to the statistics, participants tend to login to system after summer.

In conclusion, the achievement of this assignment is help app developer to define target audience and improve their design and marketing strategy. For these reasons, we consider the content and layout should tend to commercial report and a guide brochure is the most suitable approach to give the analysis to data owners.