Assignment 1

Overview

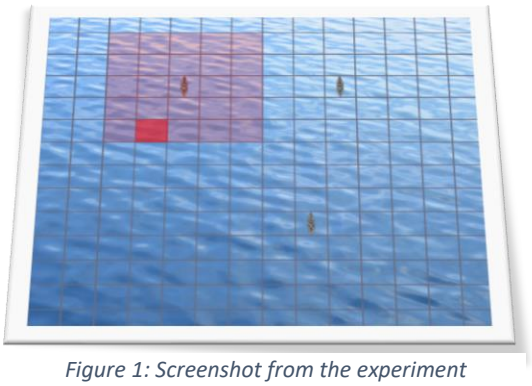

Figure 1: Screenshot from the experiment

Your task is to write a program that will parse, reformat and summarise results from a decision-making experiment.

You will be parsing actual results from an experiment that investigated the role of skill or perceived skill in the near miss effect. In a nutshell, the experimental question was whether a near miss (i.e. an outcome that comes very close to success) produces excitement and increases motivation even in conditions where the user has no control over the outcome. If you want more information, you can have a look at the attached abstract. You can also try the experiment yourself by visiting goo.gl/4JVvEk

Below you will find a detailed description of the result files that your program will analyse and the specification for the tasks that your program needs to perform. Each task is independent of the others which means that you can program them in any order.

Full marks will be awarded to programs that not only accomplish the tasks, but do so using readable, efficient and easy to maintain code. Critically, your program should work not only with the provided dataset but with any dataset that has the structure described below. For example, your program should not crash or omit results, if files are added or removed.

Submit your program in Moodle as a single python (.py) file, named after your **candidate number** (not your name or your student id), e.g. VXCZ9.py.

Note: No marks will be awarded **for code that was not discussed in the lectures**.

Description of experiment result files

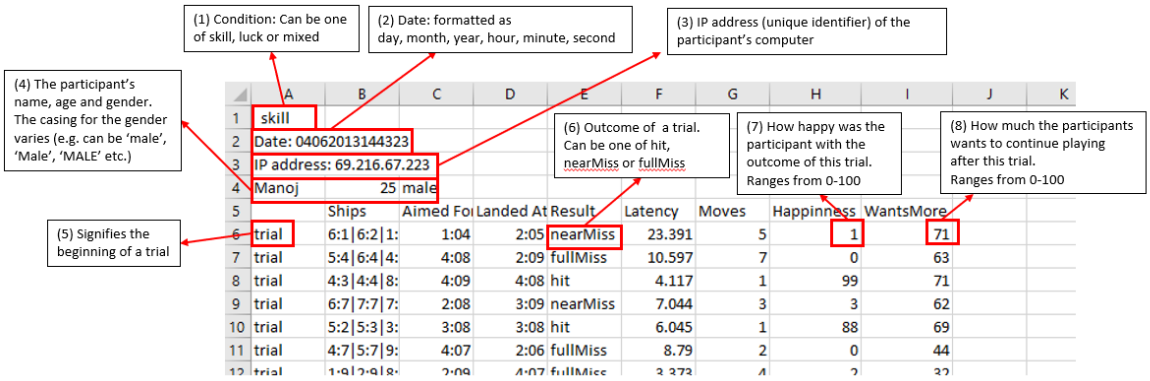

Figure 2: Example of a results file

In the compressed (zipped) file results.zip[1] you can find the result files, one per participant. Figure 2 above shows how such a file (csv) looks when opened in Excel (the file "File Structure.pdf" contains the same figure in higher definition for your convenience).

The experiment took the form of a battleship game. Participants played a variable number of rounds and in each round, they tried to hit a ship by launching a missile. After each attempt, participants were asked to report how happy they were with the outcome (fig. 2, label 7) and how willing they were to continue playing (fig. 2, label 8).

There were 3 different conditions (skill, mixed, luck) varied between participants that differed in the skill required (fig. 2, label 1). Condition "luck" was essentially a game of chance, in "skill" the participant was in control of the outcome and condition "mixed" was somewhere in between. In each round the possible outcomes were hit, nearMiss, and fullMiss (fig. 2, label 6).

The experiment also recorded the participant's name, age and gender (fig. 2, label 4). Note, that due to a programming error, the letter casing for the gender varies (i.e. can be "Male", "MALE", "MaLe" etc). Finally, the IP address (fig. 2, label 3) and the date of participation were recorded. The date is recorded as day, month, year, hour, minute and second without spaces. For example, if the date was 30/1/2019 and the time was 11:05:01, the date will have been recorded as 30012019110501.

1 The files are compressed for convenience only. You are not expected to work directly with the zip files. Rather, you should unzip them manually in a folder. If you don't know how to unzip / decompress files, have a look here: https://www.wikihow.com/Unzip-a-Fileds

Specification

1. Create a comma separated (csv) file suitable for import into R/SPSS/Minitab. Each row should **represent one participant** and contain 17 columns with the following:
    a. Condition (1 column) (10)
    b. Name, Age, Gender (3 columns) (10)
        - For the gender, you should write 1 for male and 2 for female instead of the strings Male/Female.
    c. Proportion of Hits, Near misses, and Full Misses over the total number of trials (3 columns) (5)
    d. Mean happiness (fig 2, point 7), mean willingness to continue (fig 2, point 8), per outcome type (6 columns) (5)
    e. The maximum and the minimum reported happiness levels (irrespective of outcome) and the trial in which they occurred (if multiple trials share the min/max value for some participant, report the first of those trials) (4 columns) (10)

2. Sometimes an individual took part more than once. However, the unique identifier of each participant's computer (known as IP address) has been recorded in the results file. Your program should ignore (i.e. not add in the csv) participants whose IP address has already been found, i.e. record their first but not their second participation. (10)

3. Create another csv file summarising the results per day. The file should consist of the following 6 columns:
    - Date, formatted as dd-mm-yyyy (eg 11-06-2013)
    - Total number of participants on that date
    - Number of participants in the skill, luck and mixed condition for that date (3 columns)
    - Average number of trials for that date (rounded to 2 decimal places) f-string,,slide 25

    (20)

Note: The numbers in brackets next to each part, correspond to the maximum marks that will be awarded. Notice, however, that they sum up to 70. Another 30 marks will be awarded based on the readability, efficiency and robustness of your code.