



Software Developer Coding Exercise

Introduction:

You have been given two CSV files. The first csv contains offline brushing data of all participants in a toothbrushing study while the second one groups the participants into groups.

Your task is to extract information for each participant (identified by their PlaybrushID) and provide meaningful information that can be communicated to the user via a newsletter.

General Rules:

1. Each row in the raw data csv represents a brushing session with the time spent brushing left, right, bottom and unidentifiable (none) sides in seconds. Each row is also time stamped
2. Merge brush sessions that are less than 2minutes apart into a single brush session.
3. The total length of a brush session is the sum of the up, down, left, right and none times.
4. Discard brush sessions that are less than 20 seconds in total
5. When a user brushes multiple times in a morning or an evening, record the longest brush and discard the others.
6. To simplify the task, assume every brush before 2pm is a morning brush while every brush after 2pm is an evening brush.

Task 1 - User Information: Please extract the following data for each user:

1. How many times the user brushed in the morning, and in the evening for each day of the week. 0 for no brush, 1 for morning or evening brush, and 2 for morning and evening brush. Multiple brushes in the same morning should only be counted once.
2. How many days in the week did a user brush twice a day. Again, twice a day represents morning and evening, and not just multiple brushes.
3. The total number of valid morning and evening brush sessions in the week.
4. The average time spent brushing per valid session in the week.

Your output should be a csv with these headers:

"group,PBID,mon,tue,wed,thu,fri,sat,sun,total-brushes,twice-brushes,avg-brush-time"

Task 2 - Group Dynamics: Please extract the following summary data for each group.

1. How many valid brush sessions were observed in total
2. What is the average number of brushing sessions per user in that group
3. What is the average brushing duration per user in that group
4. Which group performed the best? Rank the groups in terms of performance.

Task 3 (Optional): Make the code reusable and accessible:

Create a website that can accept rawdata and group csvs and display the table with user information to the user. You can use any web framework for this.

Please send your results to stephanie@playbrush.io and tolu@playbrush.io by 2300 GMT on Sunday 17th February. Upload all your code to a public/private repository and attach the link to your email.