

## Assignment 10

**Due:** Task 1: Sunday, April 26<sup>th</sup>, 11:59am PST  
Task 2: Tuesday, April 28<sup>th</sup>, 11:59pm PST.

### 1. Generation of Data in the webpage – Please wait for Piazza post to start!

The beta demo version of the platform is now live!

<http://crowds.Sharad.com/web/#/>

In order to test the platform and begin exploring data analysis methods, we will run an in-house experiment on five domains. We will all answer the questions on the five domains to generate the data we will be analyzing. This has to be done by **Sunday, April 26<sup>th</sup>**. Please take half an hour to do this, so that we can generate a decent enough data-set to begin working with.

**Submission:** Each team must submit a document with a screenshot of the final page on the platform after all tasks have been completed, for each team member. This will be submitted on crowdgrader. Please include your team's name, the screenshots and the name for each team member. Please submit answers to the best of your ability – do not generate “junk” data, so as to provide yourselves with realistic data for analysis.

### 2. R script to analyse the data

Write an R script to make a basic analysis of the data collected. This script has to be submitted by Tuesday, April 25<sup>th</sup>. You need to develop two functions:

1. Aggregate crowd answer: This function will evaluate how the crowd performed overall for each task. There will be two cases for this function:

- Case 1 - multiple choice, binary: mode
- Case 2 - point estimate: median, mean, geometric mean, truncated mean, truncated geometric mean

This function will have the task answers as input and will output 1 answer for multiple choice and binary questions, and 5 answers (1 for each aggregation method) for point estimate questions.

Remember: THIS FUNCTION OPERATES WITH TASKS.

2. Ranking: How did participants perform on a domain overall ? This function will rank participants PER DOMAIN. You will use this function to know where does the “crowd” or “crowds” rank with respect to the users per domain. There will be two cases for this function:
  - Case 1 - multiple choice, binary: You will compute the number of right tasks in the domain and rank the individual according to the number of tasks that she got right in the domain.
  - Case 2 - point estimate: There will be two rankings for this case:
    - Case 2.1 - Relative performance: For each task you will compute the participants' relative performance as: (the participant's guess - the true answer) / the correct answer. These quantities will then be averaged over the tasks in the domain.
    - Case 2.2 - Average Ranking: Compute the percentile of the individual per tasks according to how far they are from the correct answer; then average the percentiles over the tasks in the domain.

#### Example

- In the case of 3 tasks (point estimate) with 2 participant responses, considering the “mean” crowd as the aggregation method. Let A be the matrix with rows representing the responses and columns representing the tasks. The entry  $A_{ij}$  is the answer of the  $i^{\text{th}}$  agent on task  $j$ .

1	2	4
0	2	10

- Let T be the vector of correct answers where  $T_j$  is the correct value of task  $j$ :

1	2	7
---	---	---

- First, compute the mean crowd. For each task, compute the mean over the answers of the users:

.5	2	7
----	---	---

- Compute the relative performance of each individual as follows:
  - $R_i = (1/3) * (((A_{i1}-T_1)/T_1) + ((A_{i2}-T_2)/T_2) + (A_{i2}-T_2)/T_2)$
- The relative performance for the two agents would be

3/21
10/21

- The relative performance for the crowd would be 1/6
- This means that the crowd would have a rank of 2. It ranks in second place with respect to the other users in relative performance.
- For Average ranking we observe that the crowd is better than or equal to 1/3 of users for the first task, better than or equal to all users for the second one, and better or equal than all users for the third one. So the average percentile would be  $1/3 * (1/3 + 1 + 1) = 7/9$ .

### Submission

For this part of the assignment you need to submit the following tables, indicating the crowd ranking.

Table 1.

Point Estimate	Mean	Median	Geometric Mean	Truncated Mean	Truncated GM
Ball Trick					
Penalty Direction					
Movie Song					
Calories					
Historical Monuments					

Table 2.

Multiple Choice / Point Estimate	Mode
Ball Trick	
Penalty Direction	
Movie Song	
Calories	
Historical Monuments	

### 3. Have fun with the data. Explore, Plot, Analyze.

Submit to crowd grader any extra analysis you performed. The data will be rich enough for you to explore your own hypothesis. Get to know the data. Try using dplyr and ggplot to present your results. This is the part when your creativity will be mostly appreciated.

### 4. Finalize tasks

All github folders are close to being finalized, and the checker file should have helped in detecting any last inconsistencies:

<https://docs.google.com/spreadsheets/d/1RiJDdbh1QFkizJpxW7nuMO1zqJRmbdpK8Xg9yVvC1YA/edit#gid=1584722480>

The checker teams for each column will be responsible to ensure that all folders are in perfect condition by **April 17** (next Friday). Please take the lead on this to ensure all fixes are complete.

- Checker teams may wish to make the changes themselves, or work together with the original team.
- For some of the columns, the checker teams will not be able to fix the issues (eg. if the corpus is missing or incorrect), so the original team will have to make these changes.
- Feel free to communicate and work together via Slack, email, or any other method of your choice to ensure that all the issues are fixed.