

Assignment 11

Due: Thurs, May 7.

Select a research question of interest and perform a thorough analysis using the demo data. Please try to use dplyr package in R, as it provides a more interpretable and comprehensive style of code. Below is a list of 5 possible topics, however feel free to propose your own as well.

Topics:

- 1) Does the social condition have an effect on the performance of the crowd?
- 2) If we estimate the crowd with those users who had the first ten questions answered correctly does the crowd estimate improve?
- 3) If we estimate the crowd with those users who had high confidence when answering the question, does the crowd estimate improve?
- 4) How does the number of participants influence the crowd estimate ? Does the crowd estimate improve as we keep adding more observations, is there a threshold beyond which no further improvements are observed (ie. does it seem to converge?)
- 5) For multiple-choice questions, can we define the quality of a crowd by the distribution of the participants' answers? Does crowd agreement on an answer indicate that the crowd is correct?
- 6) Propose your own topic

Submission:

Submit a report on crowd grader including the code, results, and an interpretation of the results. Provide any relevant exploratory analysis, visuals, and /or tables that support your conclusions.