Group Projects Guideline
BDA 503 - Fall 2017

Brief

- Groups will consist of 3-4 people.
- You should work on an interesting and fairly large data set. You can find your own data set or ask for help from your instructor. Data sets related to Turkey are recommended.
- Your data cannot be confidential. At the end of the project, you will open up both your work and your data to the public.
- There will be three important phases: Proposal, Minimal Working Report, Final Report/Presentation.
- Reproducibility will be important. From the very beginning of input commands to the output of your report you will be asked to adhere reproducibility rules.
- Peer review will be important. For peer review, you will be asked to report on another group's work and bring suggestions for improvement.
- Style will be important. You should add comments to your code, do proper indentation, follow a naming style. You will be asked to check your plots appearance, do good color scheming, proper and clear connections to the context. Style will only be evaluated at the final stage.
- On demand, your instructor can arrange you an advisor. The nature of help from your advisor will depend on the advisor and your relationship.
- *To be determined:* You might be asked to maintain a GitHub repository for your code and report.

Proposal Phase (Deadline Oct. 20, 23:59)

- Your proposal should include your group name, group members, data (with description), your objective (what you want to achieve) and your tentative plan. State your references.
- Your proposal should not exceed 2 pages (except references). Start with an abstract.
- Your data should satisfy the following conditions
    - at least 1,000 observations
    - contain at least one categorical variable (you may create one)
    - contain at least 8 different variables
    - be in a tidy format (you may need to clean and reshape the data as part of your exploration) refer to http://vita.had.co.nz/papers/tidy-data.pdf
    - the data set should be in a commonly used format such as .csv, .tsv, .txt, .xlsx or .xls

Peer Review of Proposal Phase (Deadline Oct. 25, 23:59)

- Prepare a report on your assigned group of maximum one page. Check quality of the data. Check the objective if it is clear. Bring suggestions to the project plan. Recommend extra material if you can.

Minimal Working Report (Deadline Dec. 1, 23:59)

- Your code, your graphs and your core report should be ready at this stage.
- You should achieve basic reproducibility capability for your project. Your RMarkdown html outputs should be traceable.
- You should be able to say something about your objective at this stage.
- Styling is not that important at this phase.

Minimal Working Report Peer Review (Deadline Dec. 8, 23:59)

- Try to reproduce your assigned group's work with their code. Check the relevance of their work with their objective. Suggest improvements and insights about the data (e.g. "You can plot a histogram here.", "You can check the relationship between customer churn and overall satisfaction.").

Final Report / Presentation (Deadline Dec. 15, 23:59)

- You should be completely ready at this phase.
- You should submit your full report in pdf and Rmd formats.
- You should also prepare your reproducible report in html and present it somewhere accessible from internet.
- Your styling should be impeccable. No incomprehensible axis names or legends. No gibberish code (provide code comments).
- You are also going to make a presentation of 10 minutes. Preparing it on RMarkdown instead of Powerpoint is bonus.
- Your report should not exceed 20 pages (including code/plots, excluding appendix/references), desirably around 7-10 pages. You can put your preliminary exploratory analysis in the appendix.

FR/Presentation Peer Review (Deadline Dec. 22, 23:59)

- Reproduce all the work. Report the errors.
- Comment on the overall report. State good parts. State improvement points. Is the objective covered well?
- What is the extra effort that is not shown in class? (e.g. an R package, R Shiny interface)
- Evaluate the presentation.
- Your peer review report (except reproducibility part) should not exceed 2 pages.

Some Helpful Links

GitHub Pages

GitHub Desktop

Markdown editor

Markdown Cheat Sheet

Markdown Tutorial 1

Markdown Tutorial 2