Project

Due Friday August 04 2023.

The project will be completed in groups of 2 – 3 members. Please create your groups in CourSys. Feel free to use the discussion forum to look for group members.

All group members are expected to contribute equally to the project, including implementation and data-science work. Failure to do so will result in mark adjustments.

Topic Options

Each group should pick a topic. There is a page of project ideas that may inspire you, or you can work on a topic of your own choice.

If you would like to choose your own topic, please post a private message to instructors in the the discussion forum with a brief summary, and get approval before proceeding.

Requirements

You are expected to apply the concepts and techniques covered in the course (as relevant to the topic).

Of course, we expect more from larger groups. A group of three should produce a final result that is about 1.5 times as "good" as a group of two.

The project is worth 32% of the final mark: this should be some guide to the scope. That's as much as 8 weekly exercises, or a little more than the quizzes. The results we expect to see will be scaled accordingly.

Code

You will be submitting a tag to a Git repository containing your code for the project. All group members are expected to commit their own work to the repository.

Please also commit **one or two sample input files** in the format your code expects (if relevant).

In your repository's README (or README.md) file, you should document your code and how to run it: required libraries, commands (and arguments), order of execution, files produced/expected. You should do this because (1) you should always do that, and (2) to give us some hope of running your code.

Make sure you add the instructor and TA (ggbaker, ada104, msa332, zhitianz) as developers on the project as well so we can see your code to mark it.

Report

A group of n should submit a report of approximately $2n$ pages. When writing the report, imagine you have been asked to do this analysis as part of a job, and your audience is your coworkers and manager: you should address technical aspects of the project and how you got your results in a way that's accessible to a technically-literate person. On the other hand, it shouldn't be too jargon-heavy.

In that report, you should address (as relevant):

› The problem you are addressing, particularly how you refined the provided idea.
› The data that you used: how it was gathered, cleaned, etc.

› Techniques you used to analyse the data.
› Your results/findings/conclusions.
› Some appropriate visualization of your data/results.
› Limitations: problems you encountered, things you would do if you had more time, things you should have done in retrospect, etc.

Project Experience Summary

In your report, include (in addition to the $2n$ pages) an overview of the project as your would include it as project experience on your resumé. Co-op calls this a "accomplishment statement" and it might go on a resumé under a heading like "Project Experience".

One of these summaries should be included **for each group member**. Of course, some of the project overview will be shared by each member, but such a statement should include individual contributions which will differ.

When writing content for your resumé, focus on your skills, education, experience, and knowledge as *accomplishments*, rather than duties and tasks. Accomplishment statements could also be phrased by *beginning* with the result, rather than ending with it.

SFU Co-op has some guidelines on how to write accomplishment statement. You may also want to look at Co-op's blog post on accomplishment statements *(https://olc.sfu.ca/index.php/blog/killer-content-make-your-resume-stand-out-accomplishment-statements)* .

Greg's Notes

For the exercises, it should be clear that you're being guided very carefully: I have worked through the problems and have a very good idea how you will solve them.

That is not true for these projects. I have worked enough on each project idea to know that it's possible to get *some* results. My gut tells me that more work will get better or more interesting results. After that, the direction you take is up to you: we may be able to help you, or we may have no idea what problem you're encountering.

Grading

Criteria for marking the projects will include:

› acquiring/cleaning the data;
› defining/refining the problem;
› data analysis;
› how well you explained the whole thing.

Submitting

Submit your files through CourSys for the Project activity.

Updated Tue May 02 2023, 12:55 by ggbaker.