San Jose State University

CMPE 138/180B Project

Project Report Requirements

The purpose of this project is for the students to learn how to formulate a simple problem/task/application and to experience how to solve it using methods, tools and techniques taught in class. Students are encouraged to identify new problems/tasks/applications. We will be using the datasets from the "bigquery-public-data" repository, where each team will choose a dataset and implement the data layer of a possible application.

Grading of Course Project

- Project proposal (abstract): (10% of the project grade)
- Final project report: (55% of the project grade)
- Presentation: (25% of the project grade)
- Team work: (10% of the project grade)

Final Project Report

- The final report should be written in the style of a technical report. Describe your work such that it can be reproductible from the report description.
- Introduction/Motivation/Problem Definition (10%)
    - What dataset did you pick?
    - What application could use this dataset?
    - Why is it interesting?
    - Who would use the application?
    - And how would it be used?
    - Why is it challenging?
    - What algorithms did you choose and why?
- Related Work (10%)

    - How does your work(application) relate to those done by others in the field?
    - Provide a citation to the sources you have read, explain briefly what each source is about, what are the pros and cons of the approach, how does it compare and contrast to your approach? Did your method improve existing approaches on the problem?

- Methods Description (20%)
    - Provide an ER Diagram of your database based on the dataset selected
    - Provide a detailed description of your application and the features used.
    - What SQL queries are needed to implement each feature?
    - Provide diagrams and drawings if suitable.
    - Provide a link to your project github repository containing sql files.
- Analysis of queries (20%)
    - What are the most important queries in your application?
    - What feature are they responsible for?
    - What are the inputs and outputs of your queries?
    - Could these queries be optimized through addition of indexes or other methods discussed in class?
    - What is the total estimated IO cost with and without optimizations?
- Results and findings (20%)
    - Measure the running time of a few most important queries in your application. With and without optimizations.
    - Show graphs/tables with results
    - Discuss the impact of your optimizations
    - Suggestions for future improvements
- Conclusions and lessons learned (10%)
- Style and writing (10%)
    - Writing, grammar, organization and neatness.

Project Presentation

The goal of the presentation is to give you a chance to share with your classmates the exciting problem you have been working on, how you solved it and the interesting findings you discovered. Prepare maximum 8-minute presentation/demo, which will be followed by 2 minutes of questions from your instructor and your classmates. The presentation should include a couple of power point slides and a demo of your application.

Team work

The contribution of each team member will be evaluated via a Canvas survey and the commit history on Github.

Project timeline

- Project discussion: 4/08
- Abstract submission: 04/15
- Report submission: 05/01
- Presentations and Demo: 05/01; 05/06; 05/08