

[CS209A-24Fall] Final Project Grading Scheme (100 points)

Data Collection (10 points)

- Use the Stack Overflow API to collect at least 1000 Java-tagged threads. (5 points)
- Store the data in a database (e.g., PostgreSQL, MySQL) or in local files to ensure that subsequent data analysis can access data from the local source. (5 points)

Java Topic Analysis (10 points)

- Analyze the common Java topics from the collected data. (5 points)
- Visualization. (5 points)

User Engagement Analysis (15 points)

- Calculate user engagement. (5 points)
- Analyze the common topics with more engagement from high-reputation users. (5 points)
- Visualization. (5 points)

Common Mistakes Analysis (15 points)

- Extract error and exception information properly. (5 points)
- Analyze commonly discussed errors and exceptions. (5 points)
- Visualization. (5 points)

Answer Quality Analysis (30 points)

- Analyze if an answer is "high-quality" based on these factors:
 - Time elapsed between the question and answer creation. (5 points)
 - Reputation of the user who provided the answer. (5 points)
 - Another meaningful user-defined factor. (5 points)
- Visualization (5 points per factor, total of 15 points).

RESTful Service (20 points)

- Topic Frequency Query Service (10 points)
 - Users can query the frequency of a specific topic. (5 points)
 - Users can query the top N topics sorted by frequency. (5 points)
- Error/Exception Frequency Query Service (10 points)
 - Users can query the frequency of a specific error or exception. (5 points)
 - Users can query the top N errors or exceptions sorted by frequency. (5 points)

Further Notes

Visualization

A good visualization should convey information accurately and effectively, in a clear and concise manner. You should choose proper visualizations based on the nature of the data to ensure that users can correctly interpret the charts and understand the meaning of the data. Meanwhile, visualization should also consider

aesthetics and user interactivity. The grading for the visualization section will consider all of these aspects, rather than awarding full marks for just functional implementation.

Candidates for demonstration on week 16

Projects selected for demonstration on week 16 may exhibit the following characteristics:

- High-quality, in-depth data analysis, with meaningful and interesting insights.
- Nice visualization that effectively conveys the analysis results; great UI/UX design.
- Anything that makes the project stand out from the rest.