Contributions

Dustin Seltz

- (Question 1) Scraped frequency information from https://scriptin.github.io/kanji-frequency/.
  Saved the frequency information into a csv and combined the information into the rest of the data
- (Question 4) Used the frequencies and level information to attempt to determine what method of learning (JLPT, Genki, WaniKani, or school grade) was best in order to learn the kanji found on Twitter, News, Wikipedia, and Aozora.
- (Question 4) Wrote queries for the level information: Given that a user has completed “N5” or “Grade 6” what are the next kanjis that user should learn?
- Communicated and organized meetings. Took the meeting notes, both for our team’s weekly meetings as well as meetings with Professor Lodha and the TAs, to ensure that our project stayed on track and everyone knew what their tasks were.

Mengdi Wei

- (Question 1) Scraped the jisho Kanjis’ readings, number of strokes, radicals, compounds, etc. information from Jisho for 2136 Kanjis.
- (Question 1) Scraped the WaniKani levels from wanikani for 2136 Kanjis
- Convert the final csv result to SQLite3 database.
- Created and maintain the Kanji website back-end server to implement SQL query operation.
- Created the kanji web page. And Integrated web page with back-end server and SQLite3 database
- Implemented some data visualizations on Kanji website.
- Implemented kanji searching and kanji table on Kanji website for 2136 Kanjis.

Jimmy Wang

- (Question 1) Scraped the Joyo Kanji list (most frequently used 2136 Kanjis)
- (Question 2) Scraped the most frequently used 20K words
- (Question 2) Scraped the readings and the meanings of that 20K words on the online dictionary Jisho

- (Question 1 and 2) Created a “Viewer” file that is human-friendly presenting data from csv file
- (Question 3) Made visualizations on Jisho-related information (radicals, number of strokes, grade levels, rank of frequency in newspapers, etc.)
- (Appendix Document) *Introduction to Kanji*

Xudong Guo

- (Question 1) Modify the code for scraping Jisho to have the main meaning of each kanji.
- (Question 1) Scraped lesson difficulty levels from Genki.
- (Question 1) Combined data from Jisho, Wanikani level and Genki level.
- (Question 1) Cleaned out extra data such as links and converting data structures to list for easy to use.
- (Question 3) Wanikani level visualization. Look into the relationship between Wanikani level and stroke numbers, frequency of use and grade levels.