CS 172 Twitter Post Collector (Spring 2020)
Group Members: Jeffrey Juan, Eric Chaing, Jose Guzman, Ji Hoon Choi
GitHub Link: https://github.com/dlwlrmaa/cs172project

Collaboration Details:

- Jeffrey Juan: fixed first phase of project to correctly collect tweets. Added more fields to json file and extended range to collect more tweets. Implemented threads and rewrote code to be efficient. Fixed code to retrieve title of html page.
- Eric Chaing: Wrote front end code with React and connected the front end interface with the backend search engine and index.
- Jose Guzman: Assisted in web-based interface development
- Ji Hoon Choi: Created indexer and searcher programs

Overview of System:

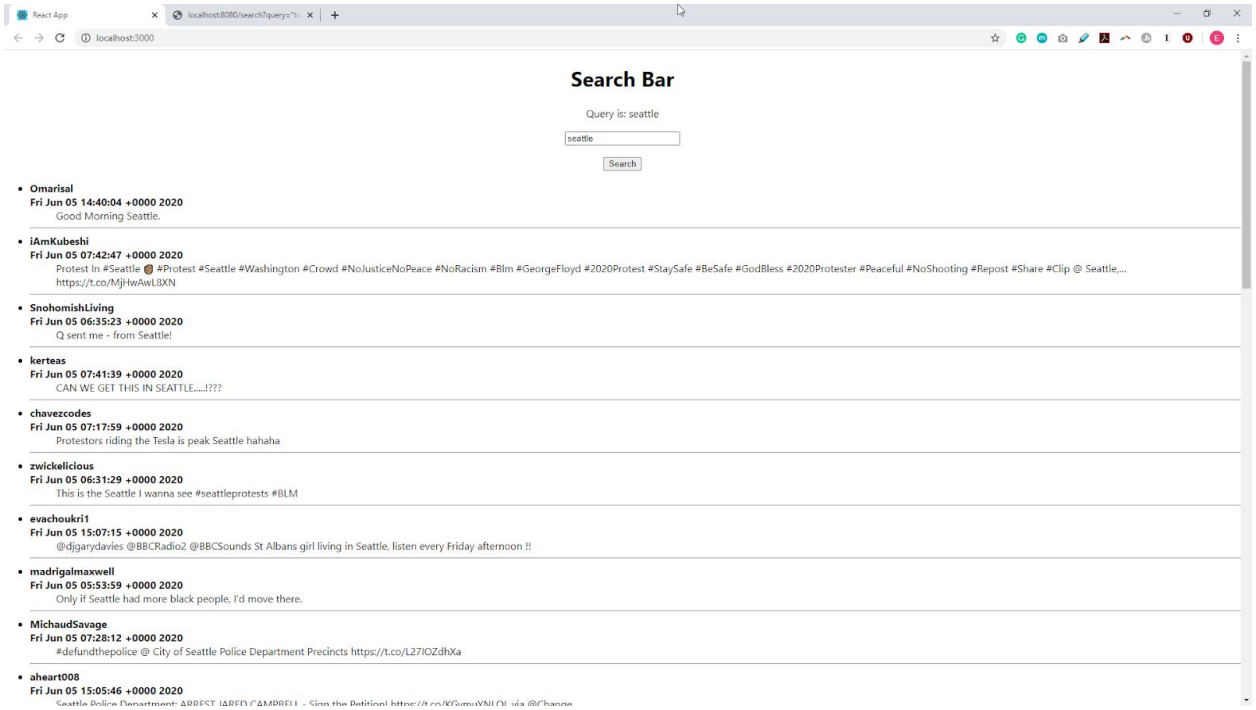
collect_tweets grabs tweets and dumps them to json files of around 10 MB each. index_tweets indexes the json files into a lucene index on disk. search_tweets reads the index and queries it. The frontend interfaces with search_tweets using REST.

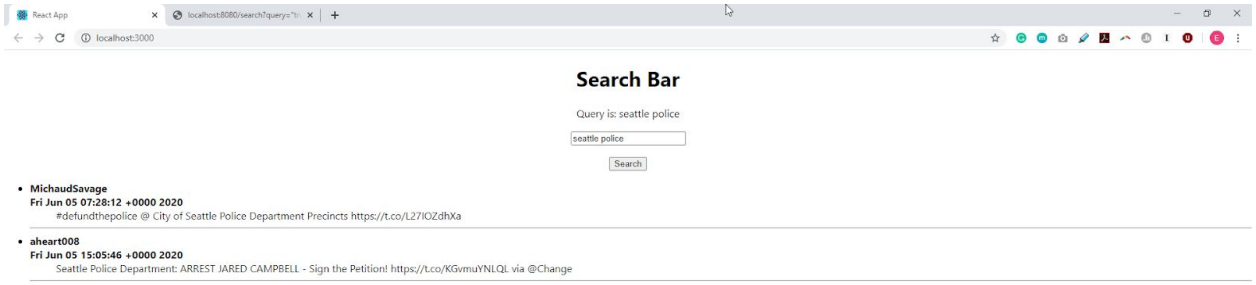
The query scoring is the default tf-idf score of lucene multiplied with a factor based on the tweet's age.

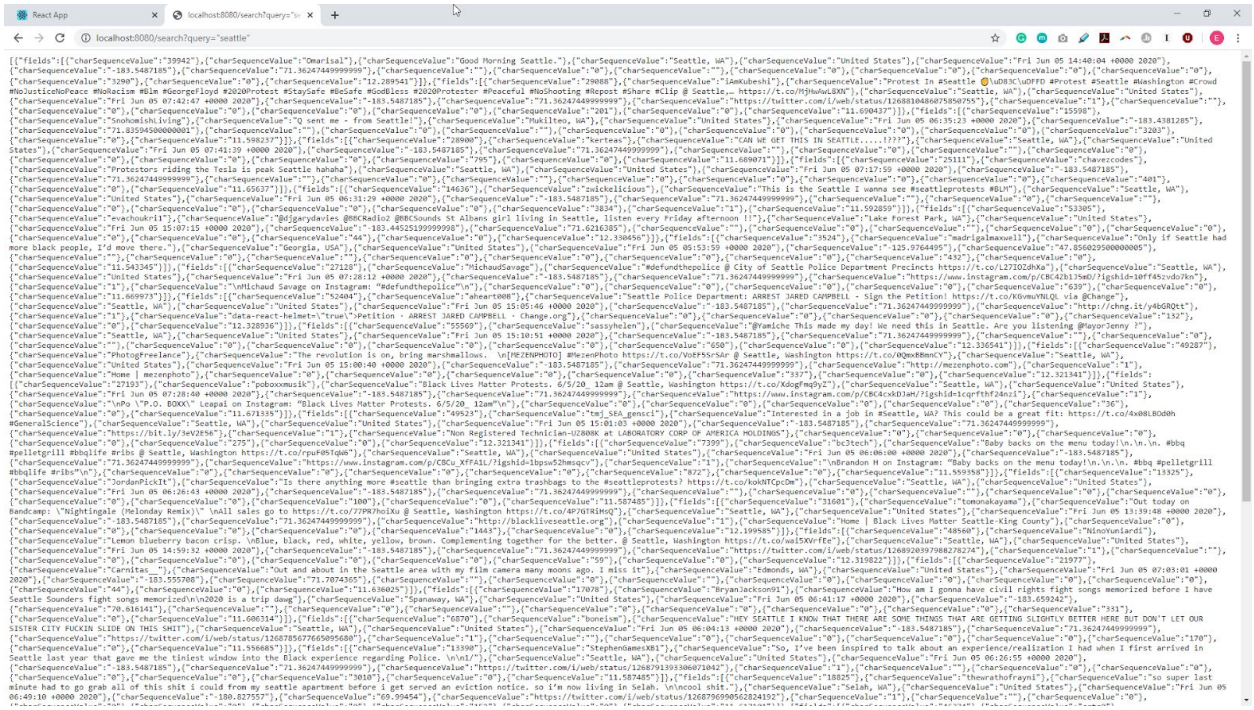
Running the program:

You will need several dependencies which are listed in the *instructions.txt* file located in the root folder.
The web interface is written in React so you will need npm and NodeJS to start the front end website. To start the website simply use the command `*npm start*`. If you run into an error message that describes missing react-scripts you may want to use `*npm install*` followed by `*npm start*`.
Before you request a search through the front end, you will need to run the back end service by building and running the *search_tweets* folder as a Java project. The included submission file should already contain the indexed tweets but you may also build and run the *index_tweets* project if you wish. Running the *search_tweets* project will boot up the local server for the search engine.
Once you have both the front end and back end servers running, you may start entering in queries.

Working Program:







Video Link:

https://youtu.be/Q1Cno2OGCVw