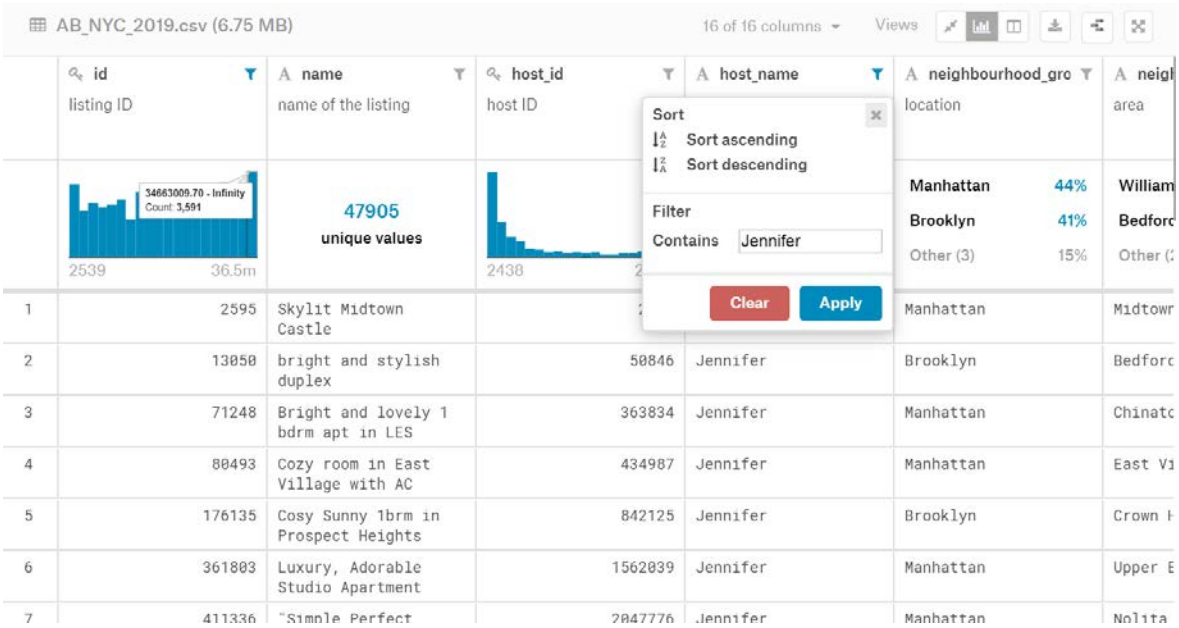
INF 551 – Fall 2019
Project Guidelines
Theme: Using a Cloud Database to Manage a Kaggle Dataset

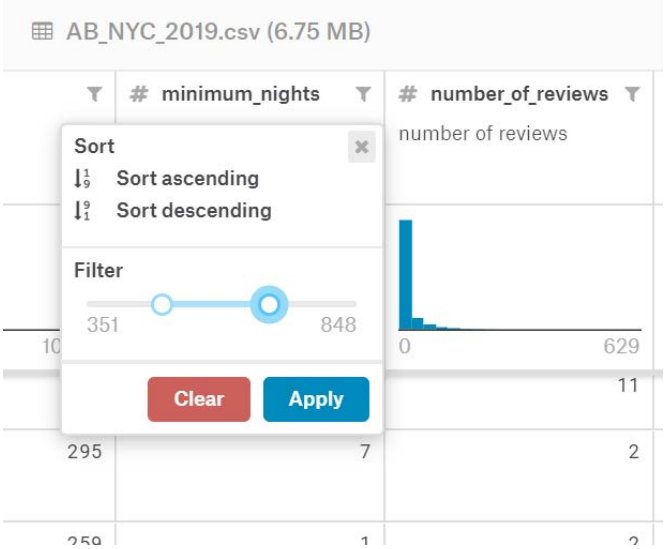
Kaggle.com is a hosting site for open data sets for data science. Kaggle has an interesting exploration facility for the dataset. For example, consider the New York City Airbnb Open Data set. The data set is in CSV format and the Web site presents the dataset in a tabular format where it also shows the statistics for each column/attribute and allows users to sort the records by the attribute. It also allows users to filter the dataset. For example, the following shows the records whose host name contains the word "Jennifer".



In this project, we ask to choose an interesting data set from Kaggle, import the dataset into Firebase, and develop an app that allows users to explore the dataset similar to what is available at Kaggle. You may choose other database than Firebase, but you need to make sure it is a cloud database where data are stored remotely on the Internet in the public accessible database.

Requirements:

- The interface of your app should show all attributes in the dataset.
- You should implement sorting function on at least two attributes: one numerical; the other non-numerical.
- You should implement at least two filters: also one numerical; the other non-numerical. For the numerical, the filter should allow users to narrow the search by specifying ranges. The following screenshot is an example.



- You may develop a Web browser based app, or mobile app.
- Visualization is not required, but may be considered for an extra credit (see below).

Extra credits:

The following are considered as advanced features of the app and may be eligible for up to 10% of extra credit.

- Utilize Firebase cloud function in your app. More details can be found here: https://firebase.google.com/docs/functions
- Develop visualization facility similar to that in Kaggle.
- Other features that we think may enhance the user experiences and/or provide insights in exploring the data set.

Phases:

The project consists of 3 phases: proposal, midterm report, final report & demo. The total point of the project is 100, broken down as follows.

- Proposal: 10 points
- Midterm report: 10 points
- Final report: 20 points
- Demo: 10 points
- Project implementation: 50 points

Proposal (1-2 pages):

Your proposal should include the following content. Please also prepare 1-3 slides for a short presentation (1-2 minutes) of your project idea.

- Dataset you choose, including some example data.

- Initial design of your app: what features will it have? How you plan to implement the features.
- Group formation: who are in your group? What will be each member's responsibility? Is your group equipped to implement the application by the end of the semester?
- Milestones: a project timeline with milestones. (no need to present this, but include it in your proposal report).

Midterm progress report (1-2 pages):

- Provide a checklist showing the items in your timeline and the status on each time (complete, on-going, etc.).
- Provide a screenshot of the components you have completed.
- Are you on track to achieve your milestones?
- Any challenges you have encountered? Any helps that you will need?
- Any other things you think should be reported in the midterm?

Final report (5-10 pages):

It should be a comprehensive report. You may include the contents from your proposal and midterm report, with changes to reflect the final implementation of your project. The final report should have the following parts.

- Project idea.
- Screenshot for each working component with a description.
- Implementation details.
- Responsibility and work of each group member.

Final Demo:

- Demo of your app (10-minute) will be done in the last week.
- Show the working of each component of your system.
- All group members should be present at the presentations.

Deliverables:

Your phase & final reports and project codes.