Boston Colleges Location Rating and Analysis

Lin Li, Yiding Ou, Tianqi Xu

Introduction:

We are exploring the rating of living around the schools in Boston by calculating the safety rate, the comfort rate, and the convenience rate. And, clients can set their own preference to find the most desired school for themselves. Another objective of us is to use echart to visualize the data in bar chart and bubble chart for the user to better compare the difference between data. We also provide the optimal needed hospital location to improve the total scores. Our safety rate will include data from crime, hospitals and crash. Our comfort rate will include data from entertainment and restaurants. Our convenience rate will include data from crash, hubway, traffic signals and MBTA.

Tools and Data sets:

Programming tools:

- Python
- MongoDB
- D3.js
- Echart

Data sets:

- Crash: http://datamechanics.io/data/eileenli_xtq_yidingou/crash.json
- MBTA: http://datamechanics.io/data/cyung20_kwleung/mbta-t-stops.json
- Hubway: http://datamechanics.io/data/eileenli_xtq_yidingou/Hubway_Stations.geojson
- Schools: http://datamechanics.io/data/eileenli_xtq_yidingou/Colleges_and_Universities.geojson
- Restaurants: http://datamechanics.io/data/eileenli_xtq_yidingou/Restaurant.json
- Crime: http://datamechanics.io/data/eileenli_xtq_yidingou/crime.json
- Hospitals: http://datamechanics.io/data/eileenli_xtq_yidingou/hospital.json
- Entertainment: http://datamechanics.io/data/eileenli_xtq_yidingou/new.json
- Traffic signals: http://datamechanics.io/data/eileenli_xtq_yidingou/Traffic_Signals.geojson

Data Transformations & Method Used:

Name of files	Data Used	Data Output
Safety	Crime Hospitals Crash	Extraction of coordinates data used
Comfort	Entertainments Restaurants	Extraction of coordinates data used
Convenience or Traffic	Crash Hubway Traffic Signals MBTA	Extraction of coordinates data used
Schoolfinal	Data output from safety, comfort and convenience Data set: schools	Schoolfinal data: names and coordinates of schools with coordinates of safety, comfort, convenience locations which are within 2 miles from a specific school
		Schoolscore data: names of schools with numbers of crime, hospitals, crash, entertainments, restaurants, hubways, traffic signals and MBTA within 2 miles, following by rating for safety, comfort and convenience.
		New optimal hospital location by using k-means

Data Analysis and Visualization:

Bubble Chart:

We have provided a bubble chart which the x-axis is the convenience rate, y-axis is the safety rate and the area of bubble is the total rate of a specific school (For total score, we have included the rate for comfort level). To visualize the bubble chart, the user can go to bubble.html and see the total score of the location when you put the mouse tip on the bubble with the school name (as shown on the Figure 1). From this analysis, we can find that a school with higher convenience rate will have higher overall rating. Those with both lower safety rate or convenience rate will definitely end up with lower overall total score. One can also visualize the comfort rate by subtract safety and convenience rate from total rate. It is a quicker way to visualize the best school with our rating system in Boston.

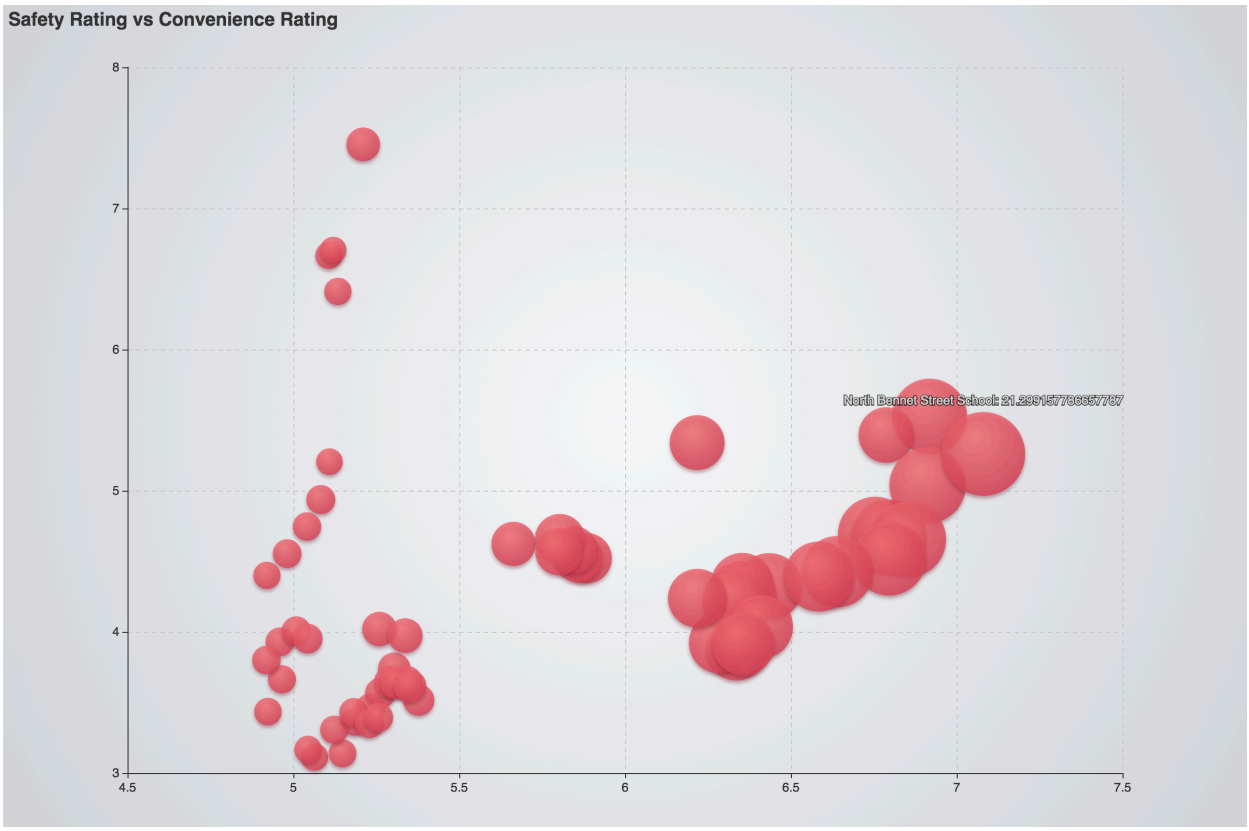

Figure 1. Bubble Chart

Bar Chart:

We also provide a visualization bar chart for people who are interested in top 10 best rated schools in Boston. One can check out information about these 10 schools by placing mouse tip on the bars (shown on Figure 2 about Boston University). To save users big amount of time, one can only compare the best schools' information and compare them with their own preference rate on different sections such as safety, comfort or convenience. Our figure does not show all

the names of schools on x-axis because there is not enough space for the names of school since they are very long, but feel free to check out from our website at bar.html.

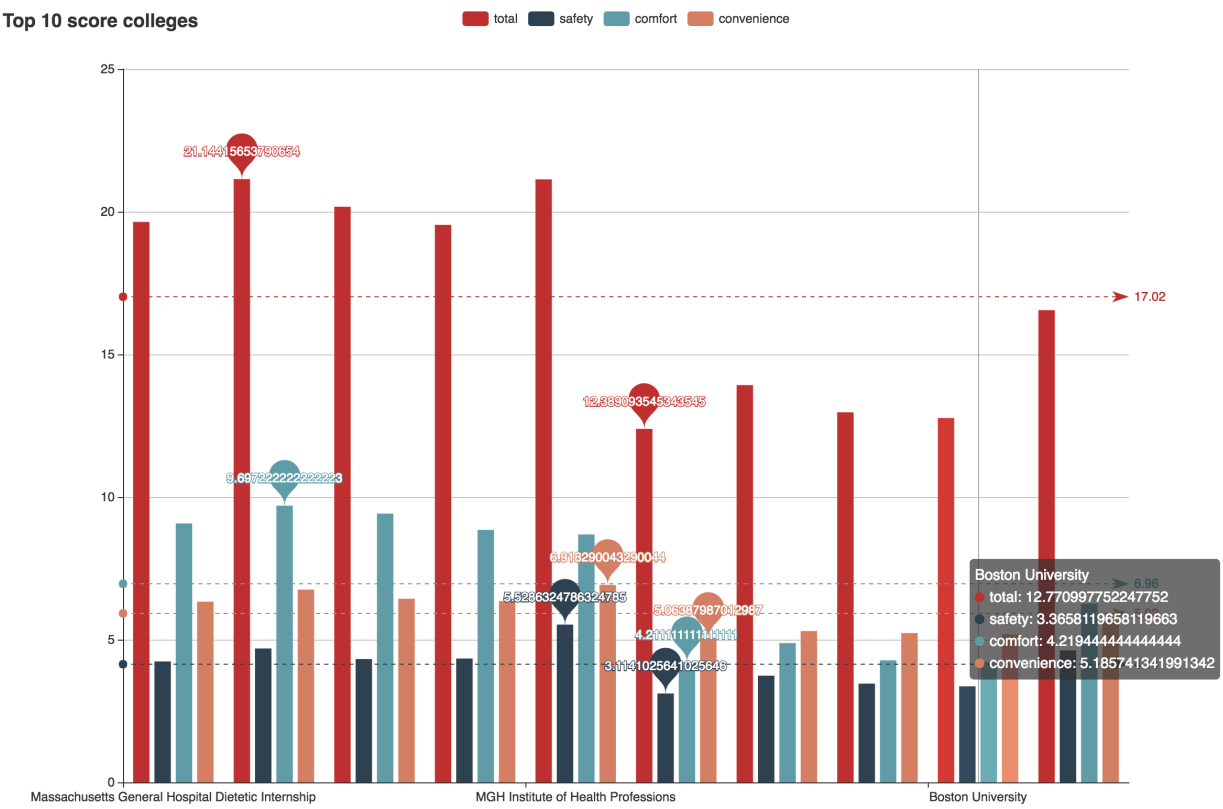

Figure 2. Bar Chart

Top Preferred Colleges:

For this interactive client-server application, we would like users to put their own preference for the rating of different scales such as safety, convenience and comfort, ranging from 1 to 3 (users are allowed to put equal rate for one or more scales) shown on Figure 3. It will return back top 60 colleges base on user's preference and their locations. When users click or place the mouse over the bar or map icon, it will display the total score for the school shown on Figure 4. We calculate the total base on another method, which get the math floor for different scales from users and times their actual rate. Total score is their math floor of their sum. It will provide a very clear view for users to select the best suited school for themselves.

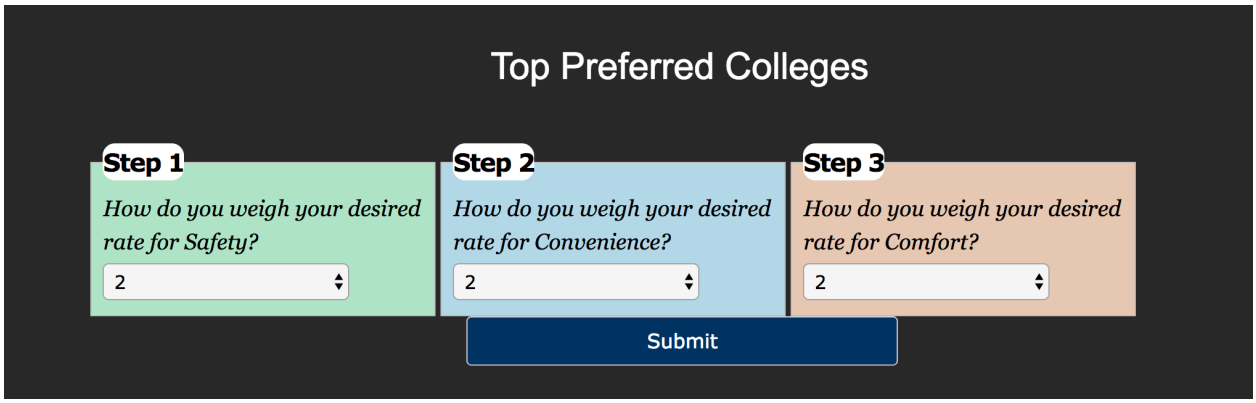

Figure 3. Preference Rating

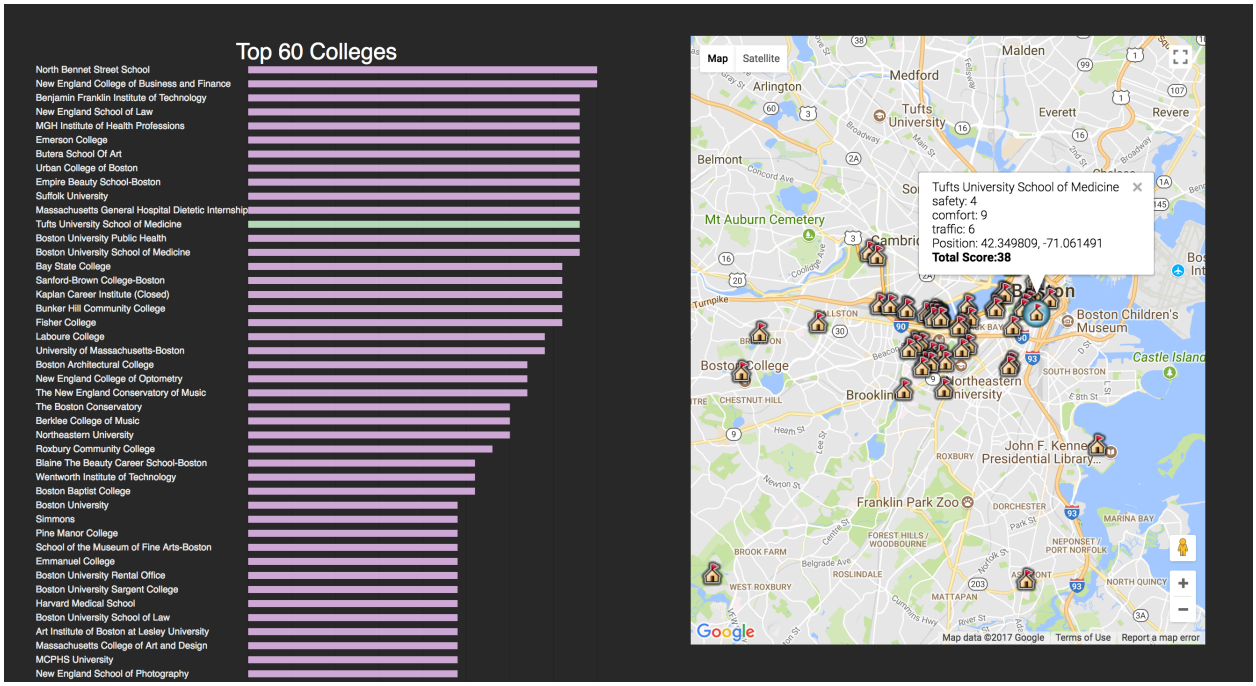

Figure 4. Top Preferred Colleges

Conclusion:

Our system is most suitable for users who would like to explore the beauty of Boston related to entertainments around Boston with consideration of safety and convenience, not a totally academic rating system. Users will find our system very helpful when they combine the results from our system with their own academic preference. It is useful that our visualization will save users big amount of time when considering a best suited school for themselves. We also use k-means to recommend for the best hospital location to improve the total score for the schools in results shown by schoolfinal.py.

Future Works:

In the future, we can make it better by using more detailed data sets. For example, we could take more data sets for each rating scale. We have not included the academic score data for different schools which yet is a very important score for choosing a school. We could also improve our scoring system with a general score system by most of the users used. Actually, the schools in Boston are not this many. We have included different colleges with same school name but they are in different region on the map. We should put them into consideration for one school, instead of two or more different schools. We have not make visualization for optimal hospitals, we may want to add this for users who can set their own preference.

Running Commands:

Transformation command:

python3 execute.py eileenli_xtq_yidingou

Website command (in folder visualization):

On terminal: python -m http.server 8000

On Chrome: localhost:8000

References:

Echart : http://echarts.baidu.com/index.html

Top preferred colleges: http://bl.ocks.org/juan-cb/faf62e91e3c70a99a306

Map api: AIzaSyCblrBBShv_9XMVeYMqNKfjW7FocjHEntY

Techniques: http://cs-people.bu.edu/lapets/591/s.php