A blue parallelogram and a light green parallelogram are positioned on the left side of the slide, overlapping each other and the dark background. The blue shape is on the left, and the green shape is to its right, partially overlapping it.

Which Boston College/University to start my career as a Data Scientist?

by Tyler Lareau

Business Problem

Which college or university in the Greater Boston do I want to work for?

I am going to leverage the power of public datasets and location data to select a school which I think will best fit my new career as a data scientist.



School Must

- be located in the Greater Boston Area
- have over 5000 students. As a data scientist, I am going to need to intake as much data as possible and believe that 5000 is an acceptable minimum to retrieve valuable and diverse data.
- Sports Complexes/Arenas must be popular venues within the 500M radius of campus. Athletics is an industry with high demands for data scientists and I would like to work at a school with opportunities to study their athletic programs.
- School must have access to public transportation in the immediate radius (500M)



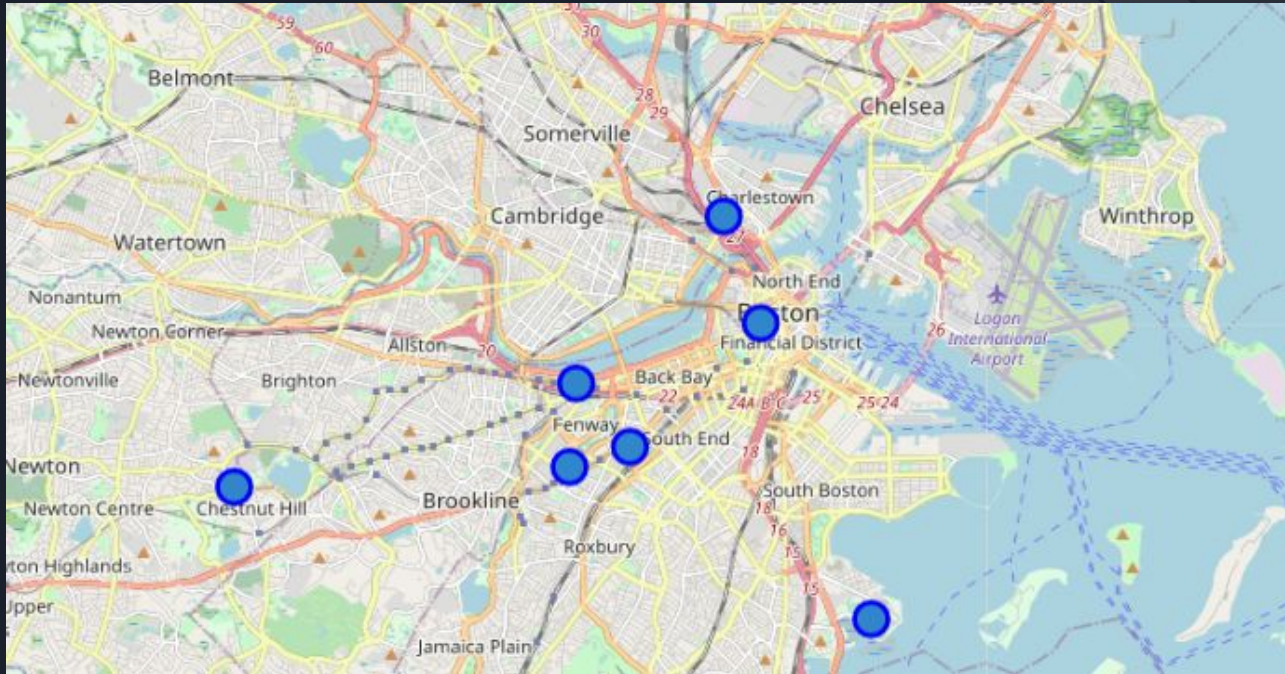
Data Acquisition and Cleaning

- School Name, Location Coordinates and Number of Students in 2013 (*most recent year) retrieved from <https://data.boston.gov/dataset>
- List included over 59 rows and 30 features
- Cleaned dataset to contain only 4 features

Boston Schools with 5000+ Students

| | Name | Latitude | Longitude | NumStudents13 |
|---|------------------------------------|-----------|------------|---------------|
| 0 | Boston University | 42.349560 | -71.099709 | 32411 |
| 1 | University of Massachusetts-Boston | 42.313809 | -71.039202 | 16277 |
| 2 | Boston College | 42.333833 | -71.169719 | 14309 |
| 3 | Bunker Hill Community College | 42.375117 | -71.069572 | 14023 |
| 4 | Suffolk University | 42.358905 | -71.061948 | 8675 |
| 5 | Northeastern University | 42.340048 | -71.088892 | 8479 |
| 6 | MCPHS University | 42.336880 | -71.101120 | 6548 |

Boston Schools with 5000+ Students Coordinates

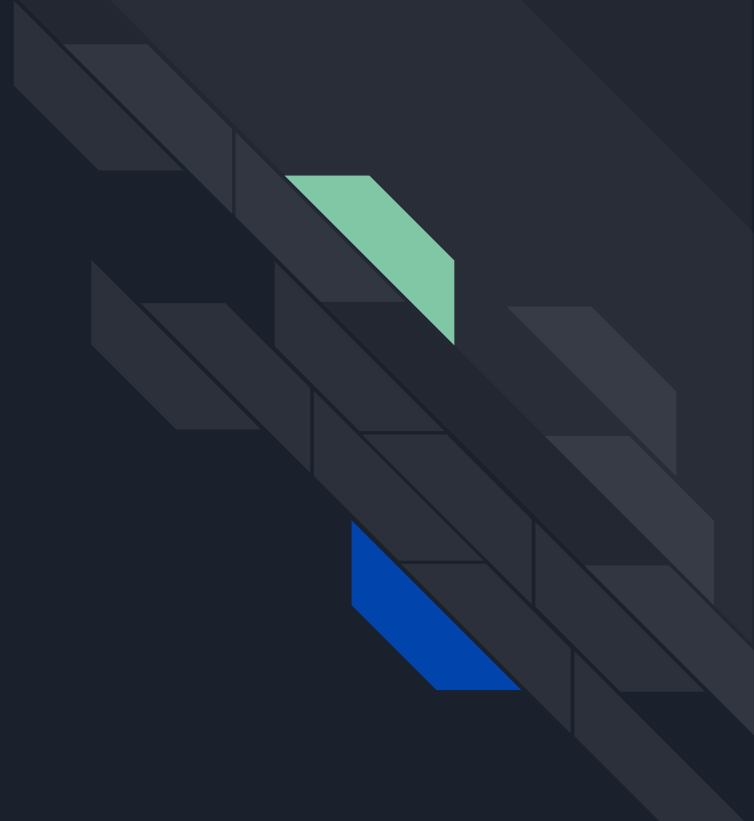


Most Frequently Checked-In Places within 500M of each college via FourSquare API

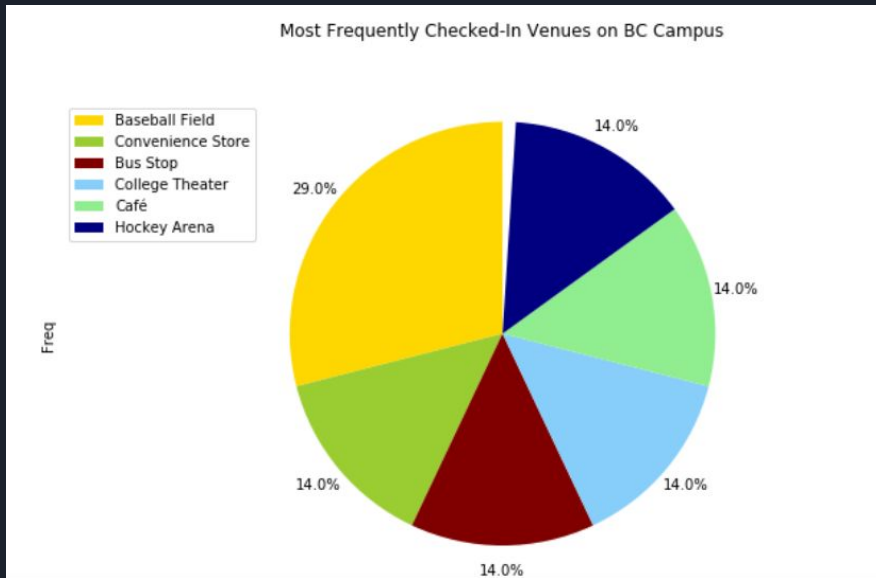
| | School | 1st Popular Venue | 2nd Popular Venue | 3rd Popular Venue | 4th Popular Venue | 5th Popular Venue | 6th Popular Venue | 7th Popular Venue | 8th Popular Venue | 9th Popular Venue | 10th Popular Venue |
|---|------------------------------------|---------------------|-------------------|--------------------|--------------------------|---------------------------|-------------------|-------------------|----------------------|----------------------|--------------------|
| 0 | Boston College | Baseball Field | College Theater | Convenience Store | Hockey Arena | Café | Bus Stop | Food Court | Cosmetics Shop | Deli / Bodega | Department Store |
| 1 | Boston University | American Restaurant | Lounge | Sports Bar | Coffee Shop | Hotel | Sushi Restaurant | Café | Pub | Gym / Fitness Center | Brewery |
| 2 | Bunker Hill Community College | Coffee Shop | Yoga Studio | Shopping Mall | Bank | Convenience Store | Donut Shop | Gastropub | Grocery Store | Light Rail Station | Liquor Store |
| 3 | MCPHS University | Coffee Shop | Sushi Restaurant | Sandwich Place | Convenience Store | Pizza Place | Pub | Donut Shop | Café | Falafel Restaurant | Gastropub |
| 4 | Northeastern University | Sandwich Place | Pizza Place | Grocery Store | Arts & Crafts Store | Middle Eastern Restaurant | Concert Hall | Café | Caribbean Restaurant | Burrito Place | Restaurant |
| 5 | Suffolk University | Coffee Shop | Historic Site | Seafood Restaurant | Mediterranean Restaurant | New American Restaurant | Hotel | Restaurant | American Restaurant | Market | Italian Restaurant |
| 6 | University of Massachusetts-Boston | Museum | Coffee Shop | Donut Shop | Fast Food Restaurant | Food Court | Concert Hall | Convenience Store | Cosmetics Shop | Deli / Bodega | Department Store |

Results

I will apply to become a
data scientist at.....



Boston College



- 3 venues in the Top 10 Most Frequently Checked in Places

- Baseball Field
- Hockey Arena
- Bus Stop



Conclusion

I analyzed the surrounding neighborhoods of seven colleges located in Boston, MA leveraging the FourSquare API to make the ultimate decision to apply to Boston College to begin my career as a Data Scientist.

- Popular sports complexes (hockey arena/baseball field) can be home to data science initiatives to predict future success of athletes and athletics programs.

- Location data can be further leveraged to select a neighborhood to move into.