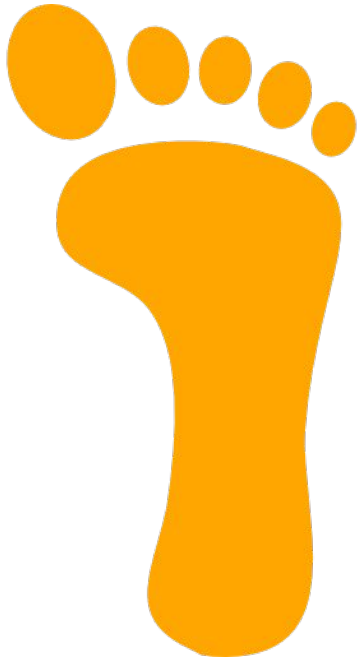
# WPI Foot Traffic

By JP Bulman, Michael Clements, DJ Alvarado

# Overview

Our project uses D3 to visualize WPI foot traffic. We surveyed almost 60 people to understand patterns of where students go throughout a typical C20 Monday.



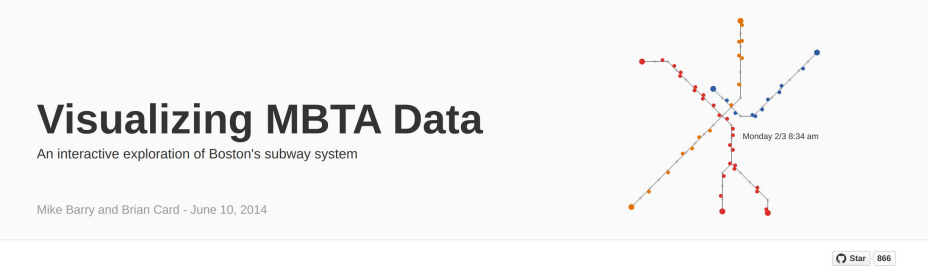
# Motivation

We had several things that motivated us, but the biggest factor was how original and close to home the project would be.

- **Original**
  This project would rely on a dataset that we would have to create.
- **Relatable**
  The data we would work with is closely tied with our school
- **Simple**
  We can readily survey students on campus

# Related Work

One of the works that inspired us was the visualizing MBTA data project. This creators of this made elegant visualizations with data that was native to just a small state in the U.S.. We wanted to follow a similar sentiment by creating a dataset that is somewhat niche and relatable to campus.



## Visualizing MBTA Data

An interactive exploration of Boston's subway system

Mike Barry and Brian Card - June 10, 2014



Star 866

Boston's Massachusetts Bay Transit Authority (MBTA) operates the 4th busiest subway system in the U.S. after New York, Washington, and Chicago. If you live in or around the city you have probably ridden on it. The MBTA recently began publishing substantial amount of subway data through its public APIs. They provide the full schedule in General Transit Feed Specification (GTFS) format which powers Google's transit directions. They also publish realtime train locations for the Red, Orange, Blue, and Green lines. The following visualizations use data captured from these feeds for the entire month of February, 2014. Green Line data became available in October, 2014 so is not shown here. Also, working with the MBTA, we were able to acquire per-minute entry and exit counts at each station measured at the turnstiles used for payment.

We attempt to present this information to help people in Boston better understand the trains, how people use the trains, and how the people and trains interact with each other.

### The Trains

In a typical weekday, trains make approximately 1150 trips on the red, orange, and blue lines starting at 5AM and continuing through 1AM the next morning. On Saturdays trains make 870 trips and on Sundays they make 760.

To better understand how the trains operate on a typical day, below are all trips that trains took on the red, orange, and blue lines on Monday February 3 2014. Each vertical line represents a station, and time extends from top to bottom. Steeper lines indicate slower trains. This visualization was first used by Étienne-Jules Marey to visualize train schedules and is typically called a "Marey Diagram."

Average Number of Trips per Day

| | Weekdays | Saturdays | Sundays |
|---|---|---|---|
| Red | 450 | 350 | 300 |
| Orange | 320 | 260 | 220 |
| Blue | 380 | 260 | 240 |
| Total | **1150** | **870** | **760** |

# Questions

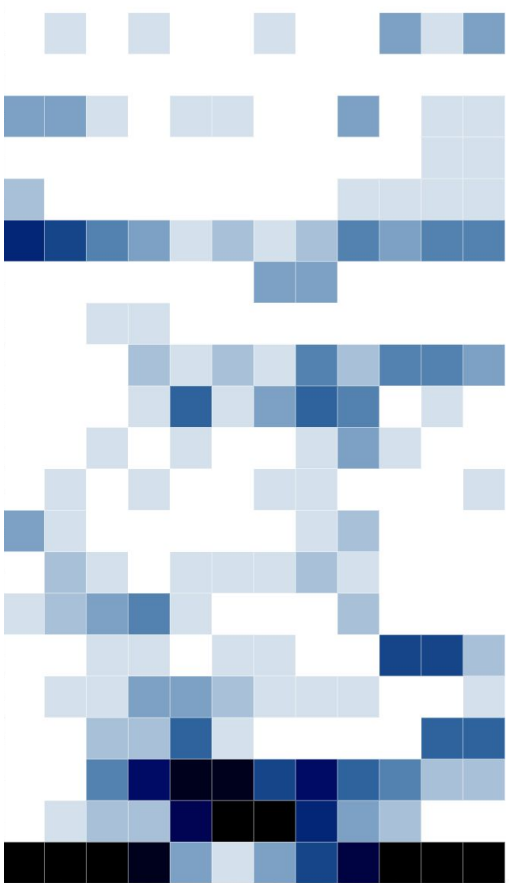
- When do people visit certain buildings on campus?
- Do people who are together at one time tend to stick together?
- Are there locations on campus that are more popular than others? At what times?

# Data

For our project, we were able to create our own data set by sending out a survey.

- **Ease**
  We wanted to get quality data and a good number of responses. So we made this survey moderately short so we could get more people to complete it while still getting usable data.

- **Who**
  Since this survey is about foot traffic on WPI campus, we sent this out to solely WPI community members.



# Exploratory Data.

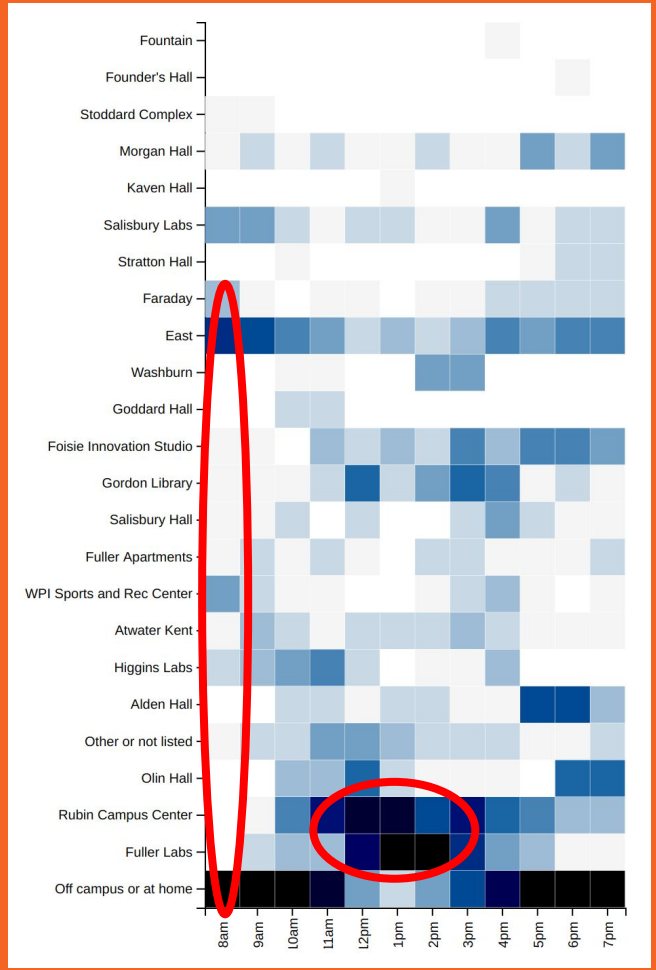
When we viewed our data as a heatmap, of location and time, there was no clear relationship between adjacent buildings in the Y-Axis. This made us think about what properties the building had that could relate them.

# Trends and Bias

Some of the trends we noticed followed what we expected going into the project.

1. Most people are at home during the earlier hours (8am)
2. Lots of people go to the campus center around 11-1 (lunchtime)

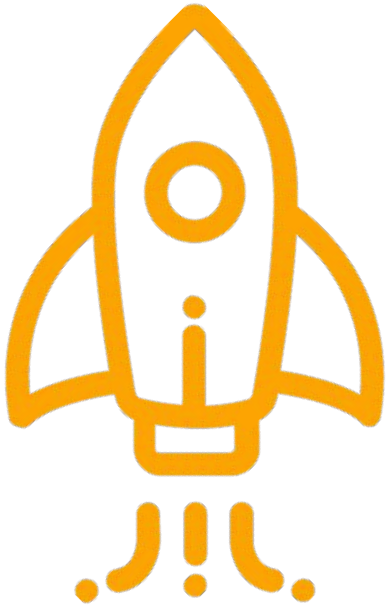
There were some biases with the survey as well. The most notable one being that there are a lot of people in Fuller from 1-3pm. This is not a trend, but rather just our data vis class.



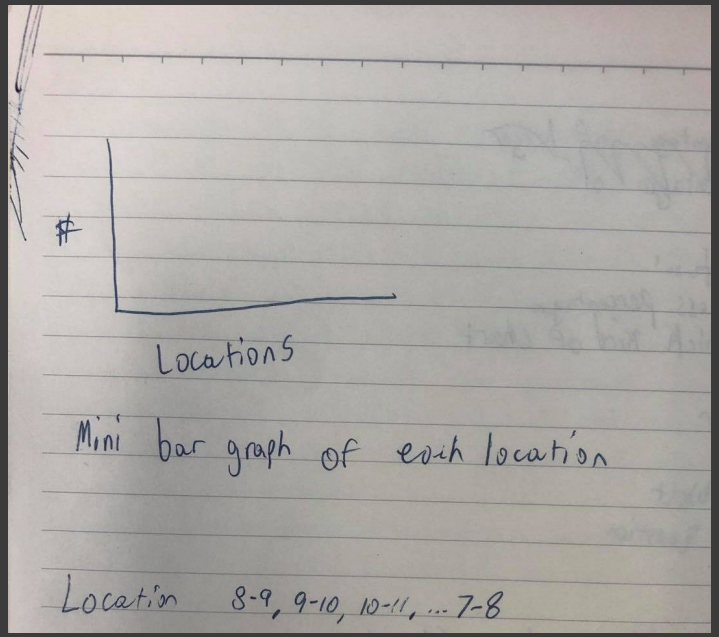
# Design Evolution

The following slides have some of our preliminary sketches for the vis.

We originally wanted to try and make this campus map as realistic as possible. At first, we thought it would have been cool for each node perform a search algorithm to find the best route to their next location. However, we quickly realized this would be too complicated and that we did not have enough data for that

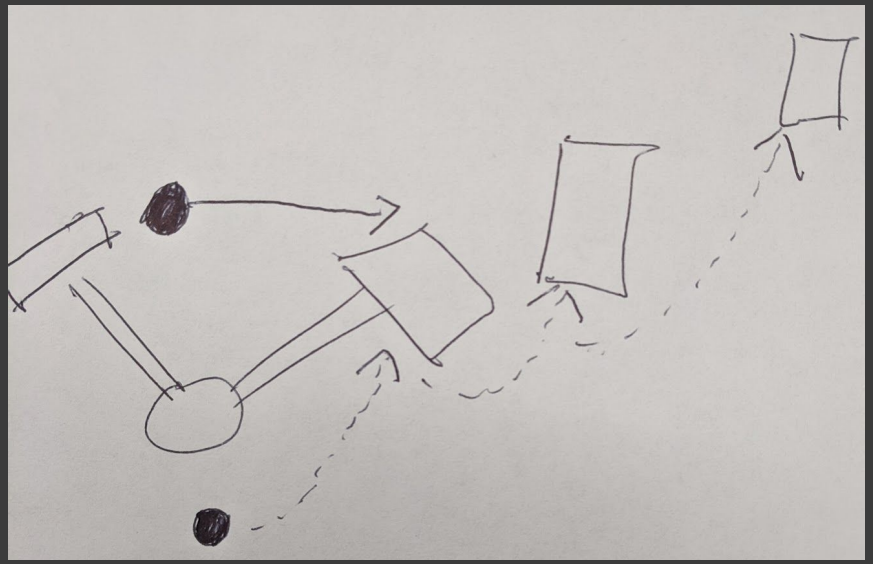


# Bar-Graph Ideas/Mockup




This mockup displays some early brainstorming on ways to visualize the data we would be collecting. A bar graph was decided as one of the ways but we needed to figure out how. We were deciding between having one large bar graph that displayed the amount of students at each location on campus at a given hour or having a bar graph for each location on campus and show the amount of students there for every hour. We opted to just try to include both graphs to have more ways to visualize our data.

# Node-Map Mockup

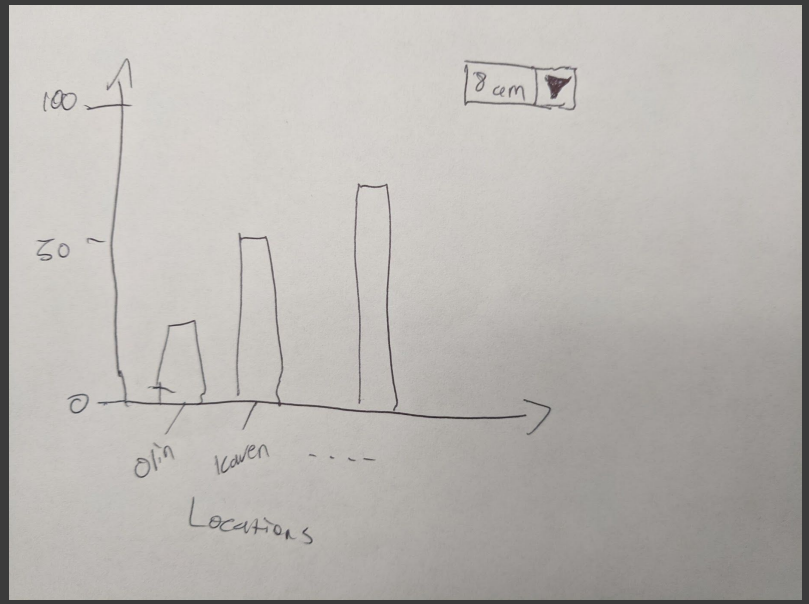


This mockup shows the idea of having dots representing each respondent, with the actual location of the building the respondent is in used to encode location.

This leverages a couple of perceptual abilities. First, high-traffic buildings become crowded, making it easy for the eye to quickly detect these clusters.

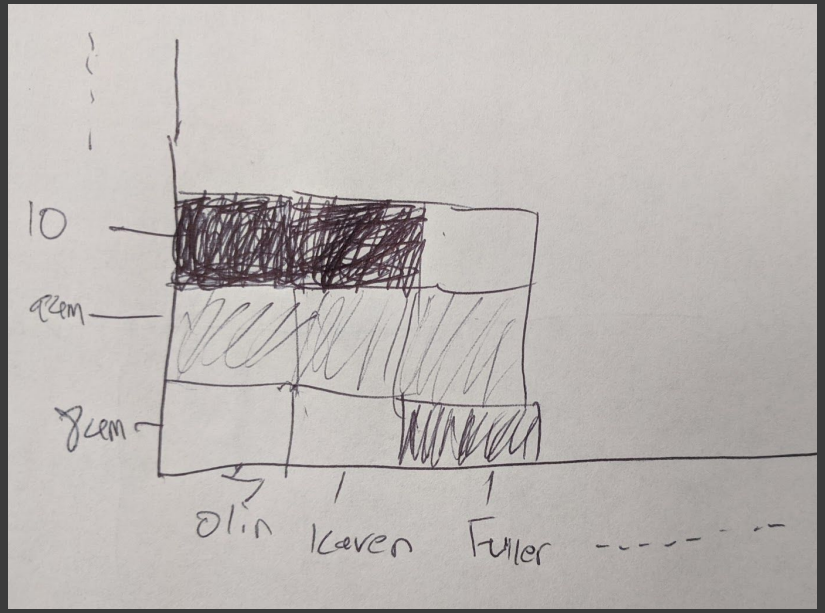
Second, it allows a user to potentially signal out one specific transition, by watching where the dots in one location end up.

# Interactive Bar Graph Mockup




This mockup shows our original sketch for the idea of an interactive bar chart. The x-axis would show the building names, the vertical axis would be the number of people, and there would be a time picker for what time the user wanted to view.

# Heatmap Mockup

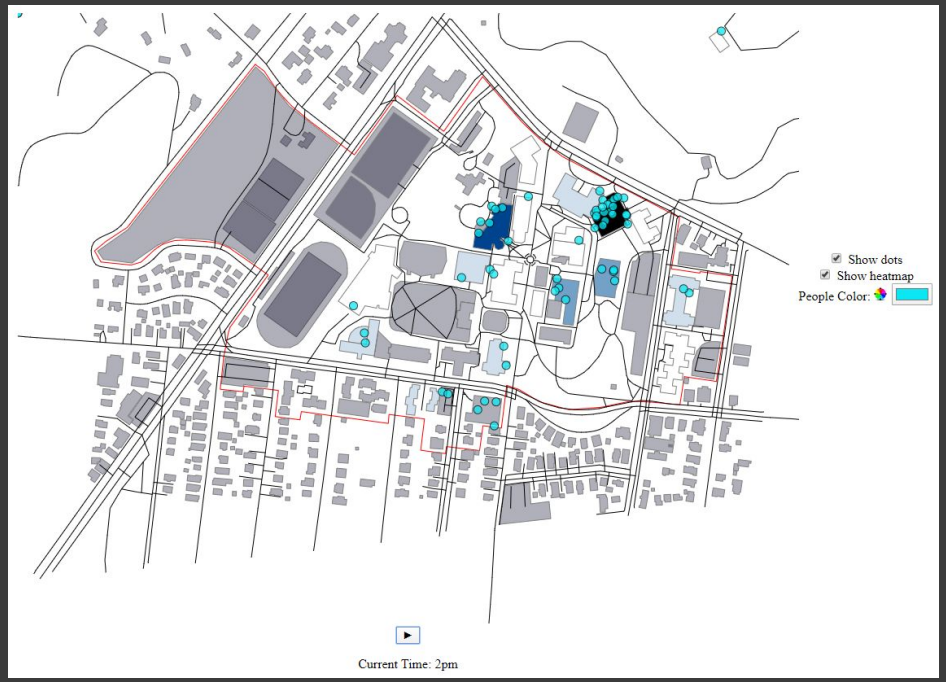



This figure is a our mockup of part of the heatmap we wanted to complete. Originally, we were not sure if it would be better to times or locations on the x-axis. We ended up deciding to implement it with times on the x-axis simply because of how many locations we had (>20).

# Mockups vs. Implementation

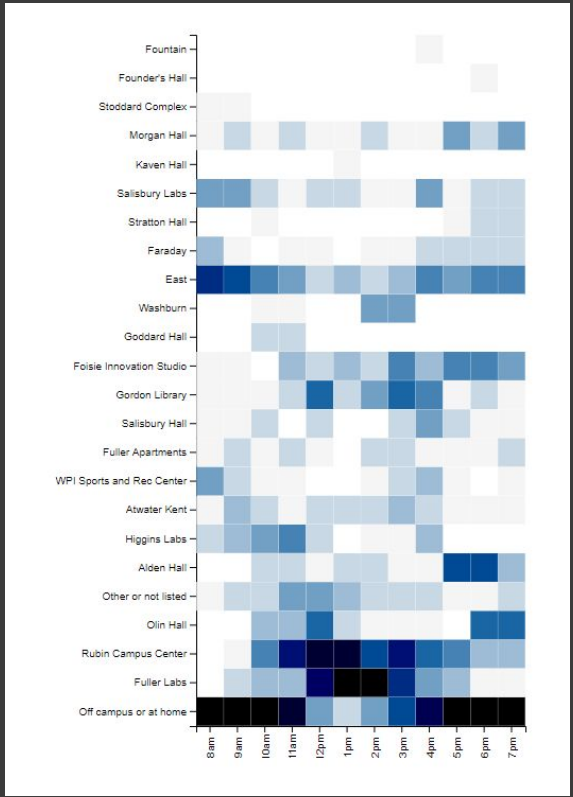
- The major differences were the additions
  - We added mini individual bar graphs for each building
- Node-map diagram
  - In the mockups we envisioned this to be slightly more complicated than we made it
  - However, we implemented enough features that we were fairly happy with it and thought that adding the additional features we thought of would be either too hard or impossible

# Heatmap and Animation on Map




For our main visualization we have generated a animated heat map of campus overlayed with nodes represented people. If you mouse over a building, a tooltip will display the name of it. The circles that represent individuals on campus. You can change the colors of all circles, and if you click on one, it changes it to red. Clicking the play button will start the animation and advance time by 1 hour every couple of seconds. The circles will move from location to location and the buildings will darken and lighten depending on the number of people at each location.

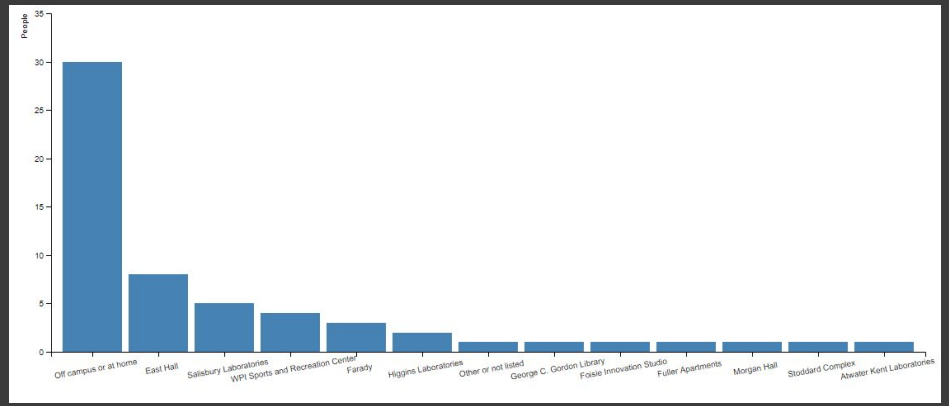
# Interactive Heatmap



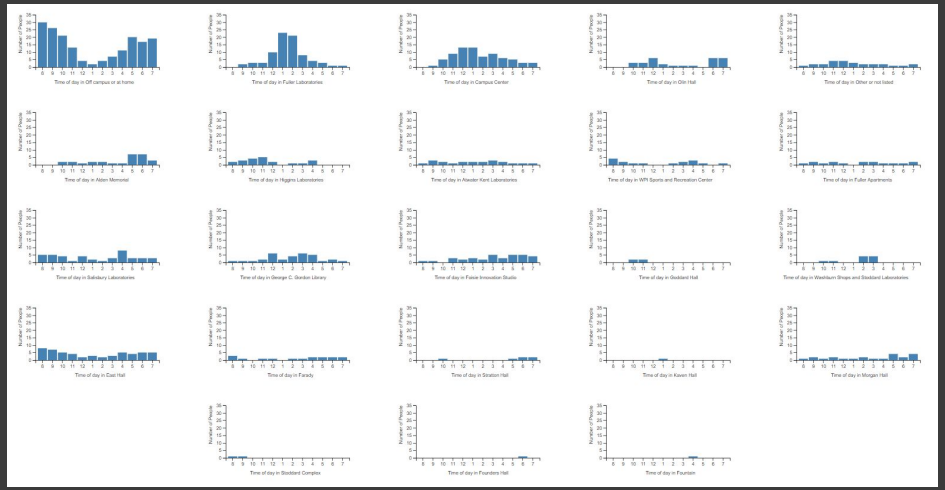

We decided to make 2 different kinds of heatmaps for our data. The first visualization on the map does not let you easily compare the popularity of a location over time. Although the data sets are the same we have to different takeaways from looking at our first 2 graphics. This one makes it easier to see different times for the same building side by side.

# Bar Graph of all locations

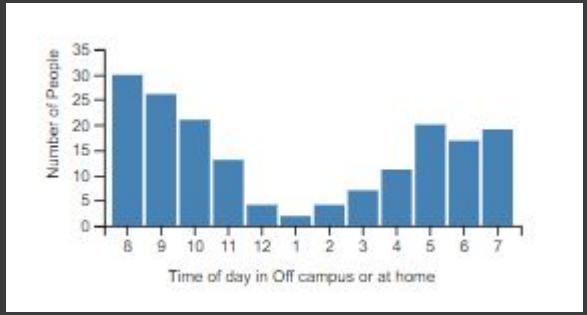
We also wanted to have a way that you can more easily compare the popularity of buildings at certain times. So we created a bar graph that displays all the the number of people at each location at a specific hour of the day. Users have the ability to change what hour that is and the bar graph would update.



# Bar graph of single locations



Comparing a buildings popularity over time was the focus on the last bar graph. We decided to create a bar graph for each location so that you can easily tell its popularity over the course of the day and compare the numbers.



# Evaluation

➔ **Did it work?**

The visualization was a success and we had multiple different channels that allowed people to easily gather different conclusion.

➔ **Improving**

The best way for us to improve our visualization would be to simply gather more data. We only had 60 responses so many places were not represented and the data was skewed towards Fuller Labs because our lecture were the primary respondents.