INFO 5100 Project 2

Yao-chuan Chang (yc2265),
Remy Jette (rcj57),
Yeon Soo Park (yp246)

Story

Each year, Cornell students have difficulty finding the perfect housing for the upcoming academic year. Many Cornell students do not have a car as they often travel from distant locations via plane, so finding an apartment that lends itself to walking is extremely important. These convenient locations are extremely desireable, so they also typically have extremely high rent. Sifting through the available listings to find an apartment in a good, walkable location with reasonable rent is difficult, primarily due to the lack of useful tools that facilitate obtaining the necessary information to compare different options and to make the right decision. Although Cornell does offer the Off Housing Listing Service (https://sf-sas-app-002.serverfarm.cornell.edu:7793/och), it simply gives a list of all the housing depending on the filters selected and does not provide much insight unless the student already knows what specific type of housing one is looking for.

By plotting available housings on an interactive map and providing different visual cues for the rent and the walkability, we believe our visualization greatly helps students spot affordable as well as convenient housings at a glance. A student unfamiliar with Ithaca that is looking for housing can easily see how the walkability of an apartment location translates to the cost of rent in this area. In addition, an outlier that has a good Walk Score but also fairly cheap rent will be clearly identified as a green, opaque point, so someone searching for an apartment can easily find them. Lastly, the ability to hover over a point to get more information about a listing provides a good balance of information - the map doesn't cause information overload at a glance, but all of the data is easily available if someone is interested in a particular listing.

Data

First, we obtained the off-campus housing data from Cornell University's Off Campus Housing Listing Services (https://sf-sas-app-002.serverfarm.cornell.edu:7793/och). We wrote a jQuery script to scrape the data and save it as a json file. This first set of data included:

- **Address**: the address of the housing
- **Date available**: when the housing is available
- **Neighborhood**: region which the housing belongs to
- **Rent**: rent per month
- **Lease length**: minimum length of the lease

- **Bedrooms**: number of rooms
- **Distance**: distance from Cornell University

To allow users to compare rent more easily across all housings regardless of the bedroom counts, we created an extra variable: rent per person. We divided the rent by the number of bedrooms, assuming that each bedroom (including the studio) will hold only one person

- **Rent per person**: arithmetic average of rent per bedroom

Second, to quantify the walkability of each housing, we collected Walk Scores from WalkScore.com (https://www.walkscore.com). Then, to later plot housings on a map, we converted each address into latitude and longitude using LatLong.net (http://www.latlong.net). So our second set of data included:

- **Address**: the address of the housing
- **Walk Score**: Walk Score of the housing
- **Latitude**: latitude coordinate of the address
- **Longitude**: longitude coordinate of the address

Next, we needed to clean and merge our data. We noticed some unreasonable values for the rent: for example, the rent for a housing with 4 bedrooms located in Collegetown was $600/month, which clearly seemed like it was the rent for one bedroom in a 4 bedroom housing rather than the rent for all 4 bedrooms. So, taking the number of bedrooms and the location into consideration, we re-examined the rent per person for each listing (paying close attention to housings with less than $400 rent per person). Finally, our JavaScript code that powers the visualization merges these two data sets as its first step, using the address as the key, so that all of the data can be used easily.

Visualization

We wanted to visualize our data on a single map so that the users can easily determine whether the rent is reasonable based on the location and whether that place convenient for those without a car. We used Google Maps API to provide zooming, local businesses and services information, and street views and used the JavaScript D3 library to plot each listing on the map as a circle.

We used color to represent different levels of rent per bedroom. As it is common that green is used for indicating underbudget and red is used for indicating overbudget, we set the color from green to red as a linear scale for the domain of 100-1800. We hardcoded this scale for our dataset (and set our slider to match) rather than simply using the minimum and maximum rent values because otherwise a single outlier could completely skew the results. This would still be valid if this dataset were to be updated in the future as rent in Ithaca does not change

that drastically from year-to-year, but may need to be changed if this were to be reused with another location (such as San Francisco, for example).

```
var colorScale = d3.scale.linear().domain([100,900,1800]).range(["green","yellow","red"]);
```

We used opacity to represent different levels of Walk Score. In order to help our users to more easily identify housings that are convenient for living, we set higher opacity for those with high Walk Scores (so that they are more visible) and lower opacity for those with lower Walk Scores. Note that a WalkScore of zero does not make the point fully transparent, as we want those locations to still be visible on the map - thus, the opacity ranges from 0.2 to 1. In addition, regardless of the Walk Score the border of the point is fully opaque, to make sure the points can be found easily.

```
var opacityScale = d3.scale.linear().domain([0,100]).range([0.2,1]);
```

In order to make our data circles more noticeable, we transformed the map which originally had many colors into grayscale, so that the color scale we used to indicate the cost of rent was more noticeable.

By using D3 and jQuery, we created three filters which dynamically updates the points on the map based on the criteria of the user, so that they can find all the possible housing within their budget and desired walkability and ignore listings that aren't suitable (such as those with many bedrooms when they intend on living by themselves):

- **Room type selector**: Number of bedrooms in the apartment (or Studio)
- **Rent selector**: Selects the range of rent per bedroom - when a new Room Type is selected, this will change to match the minimum and maximum rent per person of that room type.
- **Walk Score selecto**r: Walk Score from 0 to 100

Finally, we added a hover-over action to each data circle to display a popup information window that displays more detailed information for that housing such as neighborhood, distance to Cornell, date available, lease length, etc., to make that data available in case the user is interested in a particular listing without incurring information overload on the main map.

Citations

Source of dataset:
- Cornell University's Off Campus Housing Listing Service: https://sf-sas-app-002.serverfarm.cornell.edu:7793/och
- Walk Score: https://www.walkscore.com
- Latitude & Longitude: http://www.latlong.net

Reference:
- Google Maps APIs: https://developers.google.com/maps
- Google Maps + D3: http://bl.ocks.org/mbostock/899711