Creative Cities

Contributors:
Moe Hadhrawi
Yu-Ann Wang

Sharing with: Oscar Mejia

Overall Summary:

We did our feedback session remotely with Oscar because we had class conflicts on Tuesdays.

When providing written feedback for our study, Oscar brought up several interesting points:

(1) The "layers" drawn on top of the city map showing public transportation, schooling, etc. will be a very compelling part of the presentation, however, users will not be able to derive as much from this information. Instead, they may only be able to track how "busy" a city is based on this data. The charts on the bottom, even though less striking, will help the readers compare direct data sets and provide a cleaner interface.
(2) Aggregating or simplify all the data sets we are using to a score
(3) Consider contrasting NYC or Boston to a less developed city and see what kind of insights could be derived.

Our response to Oscar's feedback was:
(1) It's true that it may be harder to understand relationships from a map. We've been very excited about exploring neighborhood data, but at the end of the day, we should focus on the bottom charts to measure and compare proxies for innovation, and correlated variables.
(2) The aggregated score idea is interesting. We're not sure if we would have enough knowledge to create such a score, but maybe we could follow up this conversation with a cluster analysis expert student from one of our classes.
(3) This again could achieve very interesting insights, our main concern is that NYC and Boston happen to be very supportive of the open data movement. A less developed city may not possess the capability to track and organize city data to the same level.

// Here is a copy of materials we provided Oscar and response back to us.

Background and Motivation

It's now the 21st century and new cities are being built and old cities are being renovated to accommodate global population growth and urbanization. By 2009, more people in the world lived in cities than rural settings. This number will double by 2050 to 3 Bn (UN World Settlement Program).

Historically, people have long felt ambivalence towards cities. Cities have been the wellspring

of ideas, the cornerstone of the industrial revolution. It's also been the source of socioeconomic equality, congestion, and crime. Inspired by several pieces of literature, we wanted to explore cities in a positive light and see what elements in a city could be related with higher levels of innovation.

Both our backgrounds tie into cities and creativity:

Moe is working in the Changing Places group, MIT Media Lab. He's investigating what makes cities creative/innovative using a tangible interactive decision support system. Moe is interested to connect tangible with graphical data representation to reduce cognitive load in understanding complex systems.

Yu-Ann has lived in multiple cities for work, including NYC, Los Angeles, Hong Kong, and Jakarta. In each place, she was struck by how urban planning has facilitated or disrupted people's ability to get together and work effectively. She is currently taking two courses on digital innovation and sustainable cities. She hopes to tie together these two courses and study why certain cities are effective in breeding entrepreneurship and innovation.

Project Objectives

Through our project, we're hoping to study different innovations using proxies (i.e., concentration of creative professions, patents issued in a city, venture capital funding, etc.) and understand community and city based correlations (i.e., university presence, coffee shops, public transportation.

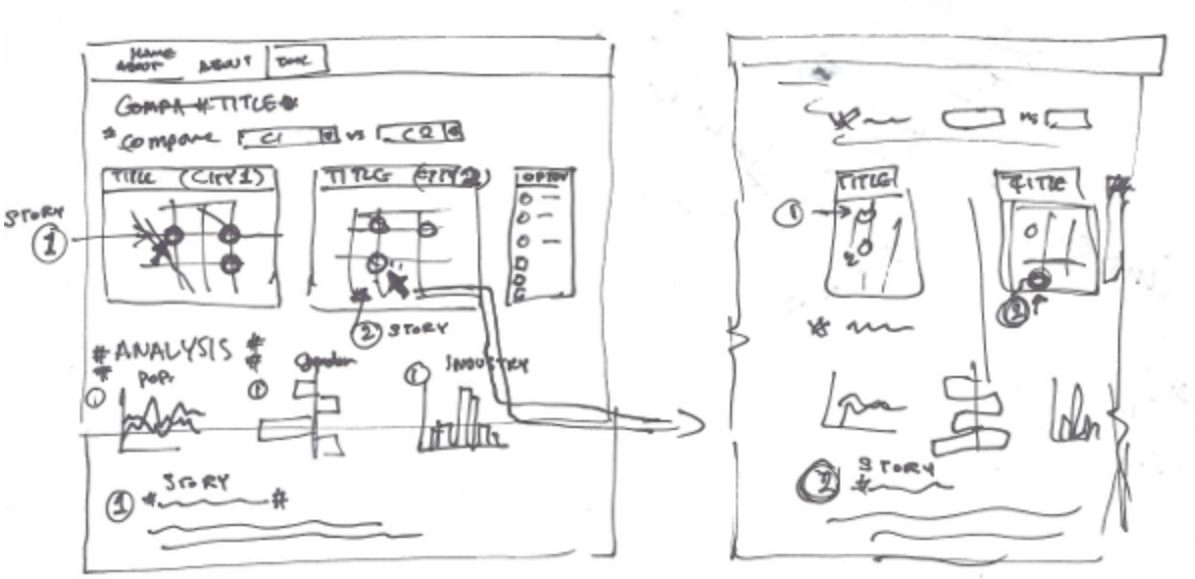
Data set

Currently:
- creative professions are pulled through the US Census (ACS). We relied on the paper *Metropolitan Migration Flows of the Creative Class by Occupation* to code professions as "creative"
- Boston open data.gov and New York open data.gov provides us access with the quantity of cafes through time (a proxy for common areas which help generate innovation)
- MTA and MBTA data is pulled from their respective websites
- patents over time is pulled from the US patent office website

Data Processing

We're fortunate that most of the data has been cleaned (Boston and New York has a pretty good open data track record). The main component we struggle with is the patent data, which is inflexible in search and nearly impossible to download.

Visualization



Here is a sketch of our visualization, we will explain the features in detail below.

Must have features

- comparison between two cities (New York and Boston)
- geomap with zoom-in function to cities (from a general map of the United States)
- layering effect including several variables we believe are correlated with innovation (i.e. public transportation, sidewalk cafes, etc.)
- graphs about innovation (i.e., creative careers in time series, patents, etc.)

Optional features

- Include more cities besides New York and Boston
- Include neighborhood level data (this would involve using articles of incorporation data, pulling from TechCrunch's Crunchbase to determine zip code level data, maybe using LinkedIn API to pull smaller companies with info on number of employees and years in operation)

Schedule

Our initial schedule was too aggressive, so we are revising it going forward:

April 14 - 18: revise feedback from TF, revise prototype based on feedback from TF, start collecting other city data and neighborhood level data
April 21 - 25: ideally move into optional features, consult professors on any blind spots
April 28 - May 2: final revision

Oscar's response:

Hi Yu-Ann, Mohammad,

Please see my comments below. I love the project! I look forward to try your visualization once it's done.

I like the concept of the visualization. It's going to be very interesting to explore and interact with it once it is completed.

I think this visualization is innovative because it tries to find hidden relationships that are assumed. Most people would agree that variables like public transportation is "good" for the city, but it may not be very clear for most people what impact it may have in terms of fueling innovation.

this visualization will not scale to a large number of cities, it would be difficult to compare more than 2 or maybe at the same time, but I think that is a compromise that has to be made. It shows a lot of data and different encodings that are difficult to combine.

The most prominent part of the visualization are the maps and the layers of data they will render. It's going to be challenging to present those layer in way that is easy to understand for the audience. The maps may be a little difficult for comparing both cites, I think people will spend more time comparing the charts at the bottom. The maps will only give an indications of how "busy" each city is.

It may also be interesting to come up with some kind of score. Some way to aggregate all the variables, compute them and see if they show the result people would expect.

Also, it would be interesting to see how these cities compare with a less developed city. I know that getting the data is what may take more time in the project, but I think that having 2 cities, and allow users to select between them would be a nice feature and may show the relationships in a more natural way.