

## Creative Cities

Contributors:

Moe Hadhrawi

Yu-Ann Wang

### I. Overview 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, and centers of economic improvement. 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.

### II. Related Work

Several pieces of work have attracted our attention:

Y combinator founder, Paul Graham, breaks down cities (NY, SF, LA, Cambridge) by its collection of people and its unique ability to generate innovative companies:

<http://paulgraham.com/cities.html>

The Economist's *A Cambrian Moment* discusses the decentralization of start-ups and how because there is widespread cloud computing power, it may be easier to start a company anywhere. This is less about public utilities, but does include emphasize the important of buildings and wi-fi.

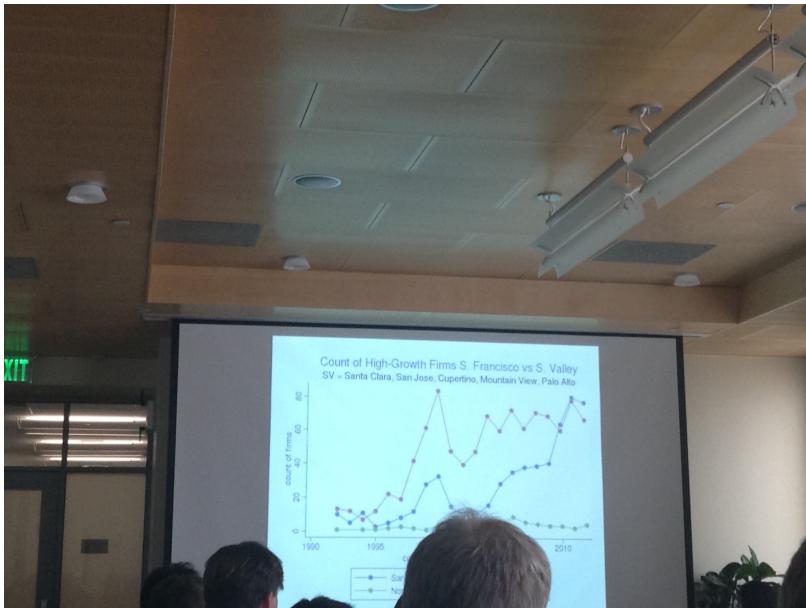
[http://www.economist.com/sites/default/files/20140118\\_tech\\_startups.pdf](http://www.economist.com/sites/default/files/20140118_tech_startups.pdf)

Tony Hsieh and the Las Vegas Downtown Project. Tony's idea is to facilitate creative collisions between people. He believes that innovative results come from unplanned interactions between people.

<http://www.wired.com/2014/01/zappos-tony-hsieh-las-vegas/>

Scott Stern and Jorge Guzman: Placecasting and Nowcasting

This paper has not been published yet, but two professors from NBER and Sloan used incorporation data to develop a predictive model on growth entrepreneurship. The professors tried to extract data and compare the shift in company growth from Route 128 to Cambridge, and Silicon Valley to San Francisco. Several images from Scott and Jorge's talk are below:



Last of all, this report from the Urban Land Institute discusses increasing a city's livability factor (i.e., parks, playgrounds) and partnerships with academic and private institutions to increase innovation.

[http://uli.org/wp-content/uploads/2012/06/Building-Innovation-Murphy.ashx\\_2.pdf](http://uli.org/wp-content/uploads/2012/06/Building-Innovation-Murphy.ashx_2.pdf)

## **Questions**

The main question we're trying to answer is in a city where there are neighborhoods for innovation, is there a correlation with smart city planning (i.e. availability of public transportation, common areas to meet and congregate, the walkability of a city, public education availability, and access to technology such as internet hotspots)

We're our mind open to new questions as we consult other teams and academics for feedback.

### **Data (source, scraping method, cleanup)**

In looking for data, we conducted a web search in addition to interviewing several people from Code for America, Code for Boston, an HBS Real Estate Professor, and two previous employees in the NYC government (for details on their inputs, please see here:

<http://bit.ly/Q7cOno>

#### **For looking at proxies of innovation, we pulled the following data**

- Paper from the US Patent Office breaking down the number of utility patents in major cities
- Total number and relative percentage of creative professions in the city (data pulled from the American Community Survey using arts, entertainment, and recreation as a proxy for a creative profession)
- Pulling city clerk information on articles of incorporation from Boston

If we have additional time or need to seek more granular data, we'd love to investigate the following:

- TechCrunch Crunchbase for number of employees, level of funding, presence of startup at a certain zipcode
- LinkedIn API for number of employees, type of industry, location of company
- Diving deeper into the US patent data since it does provide which company is the assignee of the patent (we haven't found a less manual way to look up the zipcode/neighborhood data for each patent)

#### **For looking at independent variables which could be related with innovation, we are currently surveying the following:**

- Public transportation data (supplied by the MBTA and MTA) providing average ridership in the subway by station during weekdays

- Cafes (NYC and Boston Open Data) as a proxy for places to informally congregate and discuss ideas
- Public school data (NYC and Boston Open Data) to gauge the prevalence of education access in a city
- Spatial data, including population, businesses, property statistics, and the built environment, for each grid cell in Boston

We're still in the process of trying to see if we can collect a technology centric data such as cellphone or wifi, but usually we've only been able to pull it for one city.

Most of the data available was relatively clean, but for anything which required geolocation, sometimes we had to cross reference different datasets. Our method for reconciling the data was through excel and index/match functions. In addition, some spatial data doesn't include lat/long values, but street addresses were provided. In such cases, we used Google Developer API, Geocoding service to convert street address into lat/lon values.

### **Exploratory data analysis**

We didn't focus as much on making mock visualizations as we did on outlining a set of go/no decisions based on the availability of the data. We were fairly confident that we could get innovation data on the state level and maybe the city level, but were unsure how we would go about collecting data on the more local level.

We initially hit some snags when we discovered that patent data was hard to download and didn't provide zipcodes for their authors. Another issue was working on articles of incorporation, and discovering that while Boston was generous in providing data on D/B/A (incorporation) down to the neighborhood level, other cities claimed most of their data was still in paper form.

But given that our initial cities of Boston and New York were both generous in providing resources for open data, we were able to find a lot of related variables on the neighborhood level. Hence we decided to go with a geovisualization. We're confident that given our remaining time, we can continue to find even more robust measures for innovation on the city and neighborhood level.

## Design Evolution

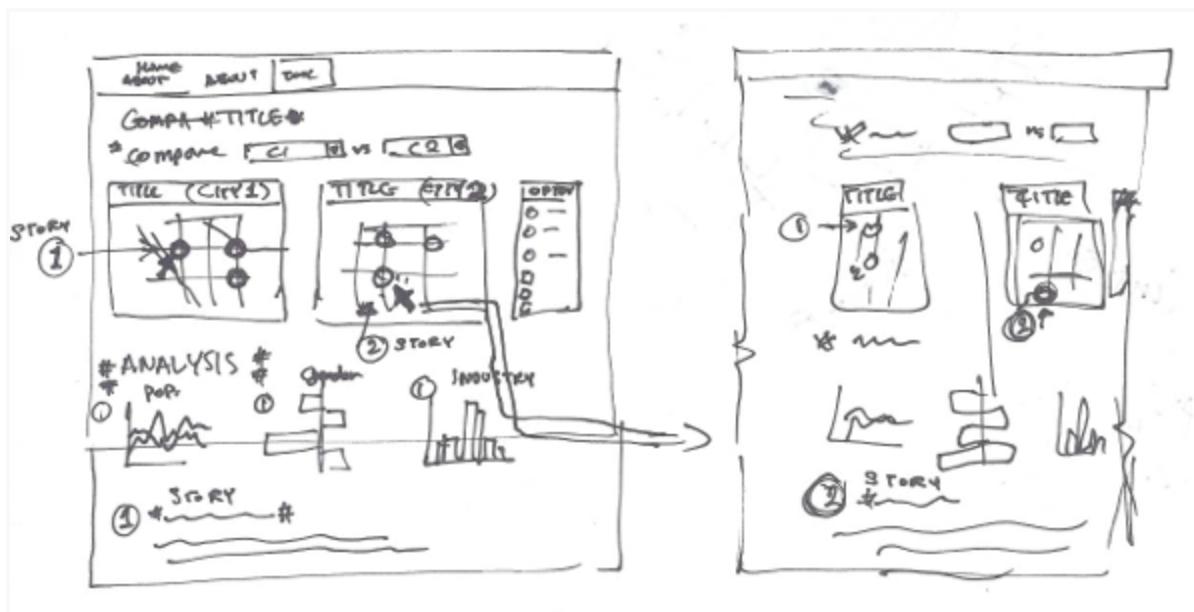
We were initially inspired by the states level data we pulled for HW4 and also the geo visualizations found in the MIT Media Lab Changing Places (<http://www.media.mit.edu/research/groups/changing-places>) and Senseable City Lab (<http://senseable.mit.edu>). We were also inspired by Chris Whong's data visualization on New York city turnstiles :

<http://chriswhong.com/open-data/visualizing-the-mtas-turnstile-data/>.

From our early discussions and design studio feedback, we realized that a geo visualization may not be helpful enough for comparing two different cities. Depending on how many layers the users examine, it's very easy to get lost in the details. We opted to also include bar charts, time series graphs on innovations and their correlated variables in the bottom of our visualization so it would be easier for our user to compare metrics in different cities.

## Implementation

We're still in the process of implementation, but we wanted to share a revised sketch of our visualization:



In this interface, the user selects two different cities from a drop down menu (or direct interaction with the maps below the drop down menes) for comparison. On the right side of the two maps, there is a panel with checkbox inputs that allow users to turn on/off

multiple spatial data layers. Below the maps, non spatial data is provided for comparison. Selected data points from the maps, or the charts could be highlighted in the story section beneath the charts analysis. An interactive prototype is provided to show the basic functionality of the visualization.

We're hoping to iterate on this after our meeting with our TF.