How is Uber changing Taxi in New York City

Zicong Liang
Yawen Li
Haowen Ni
Diana Chenyu Zhang

Background

Uber - a location based app

- A new riding model
- Sale performance of taxi medallions drop dramatically
- 13,587 yellow cabs vs 60,000 cars from black cab and app services in NYC
- The biggest private taxi app in NYC

Data

Dataset 1: Uber pick-up data

- Over 1 million Uber pickups
- April - June 2014
- Obtained originally by FiveThirtyEight
- Downloaded from Kaggle
- Variables
  - date and time
  - pick-up location
  - uber base code

Dataset 2: NYC yellow taxi data

- Taxi data from corresponding months
- Publicly available on NYC Taxi and Limousine Commission (TLC) website
- Include more variables
  - date and time
  - pick-up/drop-off locations
  - trip distance
  - fare

Data

Additional Dataset

- NYC neighborhood boundaries Geo Coordinates

Limitations

- All UBER data are business confidential, we can't get complete dataset from 2014. Currently, we can only focus on April-June in 2014.

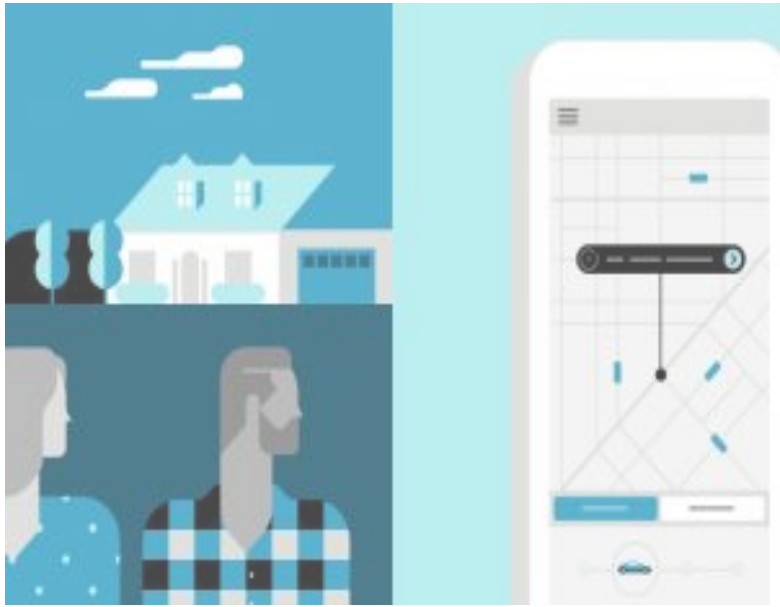
- Due to the lack of variables in Uber dataset, we can't do comparison in many aspects, such as price, numbers of passengers, distance, etc.

Use Cases

1. Uber driver wants to find out where to find more pickups at a specific time of the date.

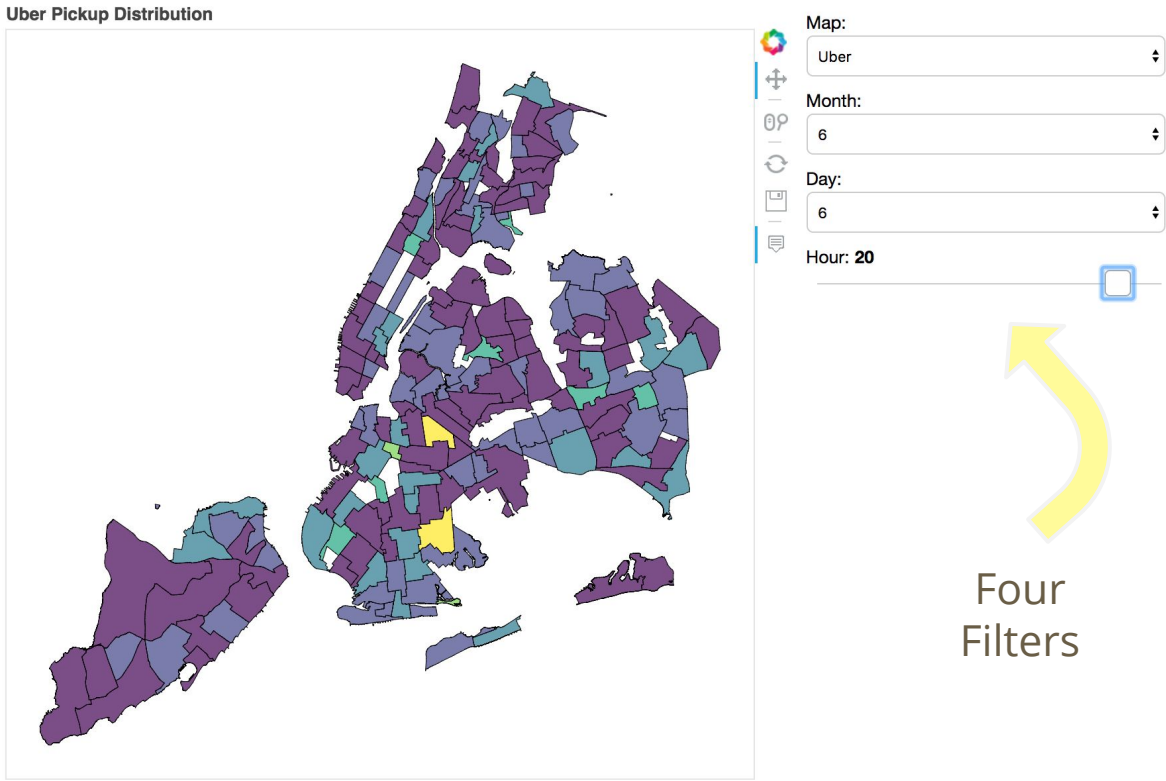
2. Passenger wants to compare the availability of Uber and taxi, and make a decision.

3. Taxi company studies their business strategies for the next year.



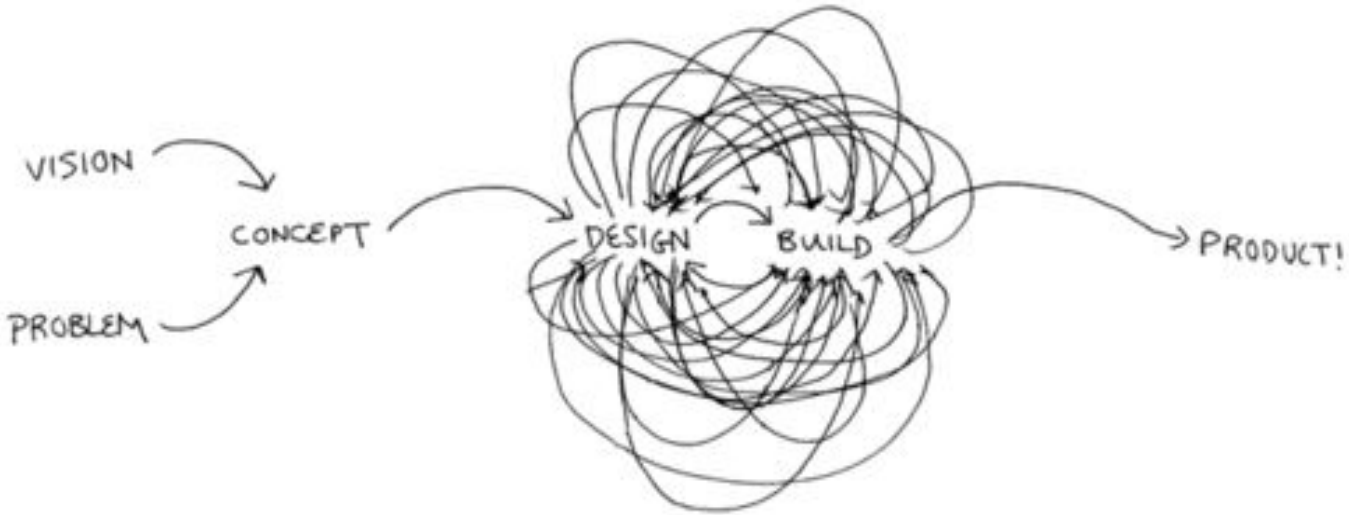
Demo

Regional Heat Map



Four Filters

Design Process



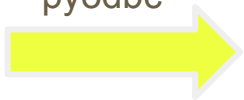
Components

- **Database**: Azure SQL Database
  - Pyodbc
- **Heatmap**: Bokeh figure, Bokeh patches
  - Hover tool
- **Filters**: Bokeh Interaction Widgets
  - Uber/Taxi
  - Month
  - Day
  - Hour

Interaction

Database

pyodbc



Heat map

Javascript Callbacks

Filter

Input:

- Month
- Day
- Hour
- Uber/Taxi Indicator

Output:

Records of aggregated pickup number returned from Azure SQL database filtered by given entry

Input:

- NYC shape data
- Data obtained from the results of SQL queries

Output:

- NYC neighborhood based heat map
- The color encodes the number of pickups.

Input:

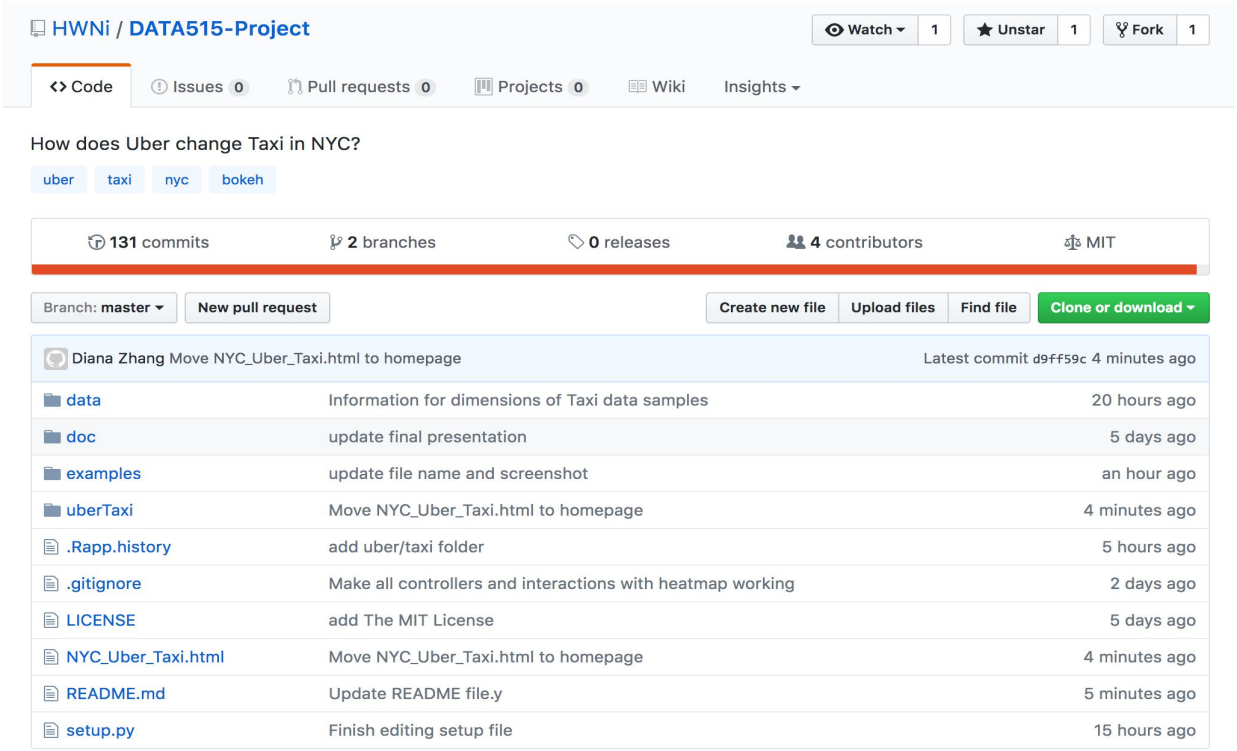
Month, day, hour, and uber/taxi selected by the slider and drop down menu

Output:

A filtered heatmap that only displays information specified by the user

Project Structure

https://github.com/HWNi/DATA515-Project/tree/master



HWNi / DATA515-Project

Watch 1 | Unstar 1 | Fork 1

Code | Issues 0 | Pull requests 0 | Projects 0 | Wiki | Insights

How does Uber change Taxi in NYC?

uber taxi nyc bokeh

131 commits | 2 branches | 0 releases | 4 contributors | MIT

Branch: master | New pull request | Create new file | Upload files | Find file | Clone or download

Diana Zhang Move NYC_Uber_Taxi.html to homepage — Latest commit d9ff59c 4 minutes ago

data	Information for dimensions of Taxi data samples	20 hours ago
doc	update final presentation	5 days ago
examples	update file name and screenshot	an hour ago
uberTaxi	Move NYC_Uber_Taxi.html to homepage	4 minutes ago
.Rapp.history	add uber/taxi folder	5 hours ago
.gitignore	Make all controllers and interactions with heatmap working	2 days ago
LICENSE	add The MIT License	5 days ago
NYC_Uber_Taxi.html	Move NYC_Uber_Taxi.html to homepage	4 minutes ago
README.md	Update README file.y	5 minutes ago
setup.py	Finish editing setup file	15 hours ago

Lesson Learned

- **Design Specification**
  - Define use case
  - Design components and interaction
- **Collaborating project with Github**
  - Use branch
  - Commit, push, and pull regularly
- **Functionality of Bokeh is limited**
  - Bokeh callback function
  - Do not use only Bokeh for creating interactive visualization
  - Combine or use other JavaScript library eg. D3.js

Future Work

- **Connect to Uber Developer API**
  - provide real-time data such as price, trip distance, and availabilities of cars nearby
  - Then we are able to do lots of interesting predictions by applying some machine learning models.
- **Use more recent data and data from more cities**
  - Such as trying to compare Uber and Yellow Cab in Seattle or even in Washington State
- **Enable narrow down to street level**

Thank you ~