

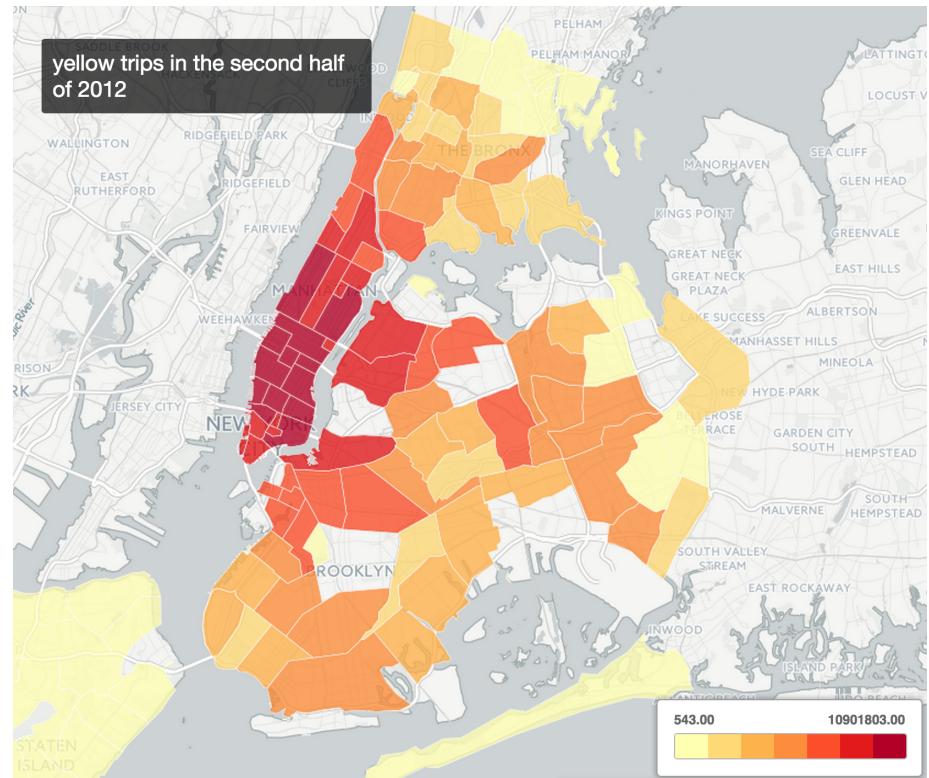
# **Effects of Boro Taxis on the New York City Taxi Market**

Israel Malkin, Alex Pine, Tian Wang  
Big Data, Spring 2015

# Motivation

- Yellow “medallion” taxis do not serve outer boroughs well.
- 2013: NYC government creates new green “Boro” taxi for outer boroughs
- Only permitted to pick up riders in Brooklyn, the Bronx, Queens, and northern Manhattan.

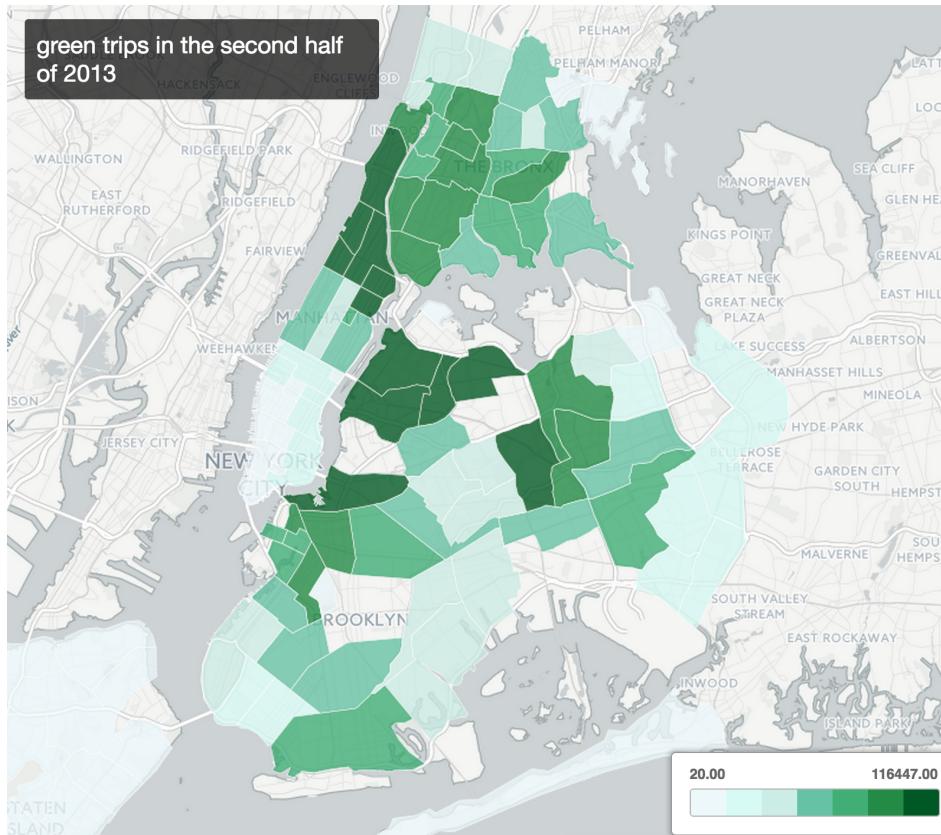
# Motivation



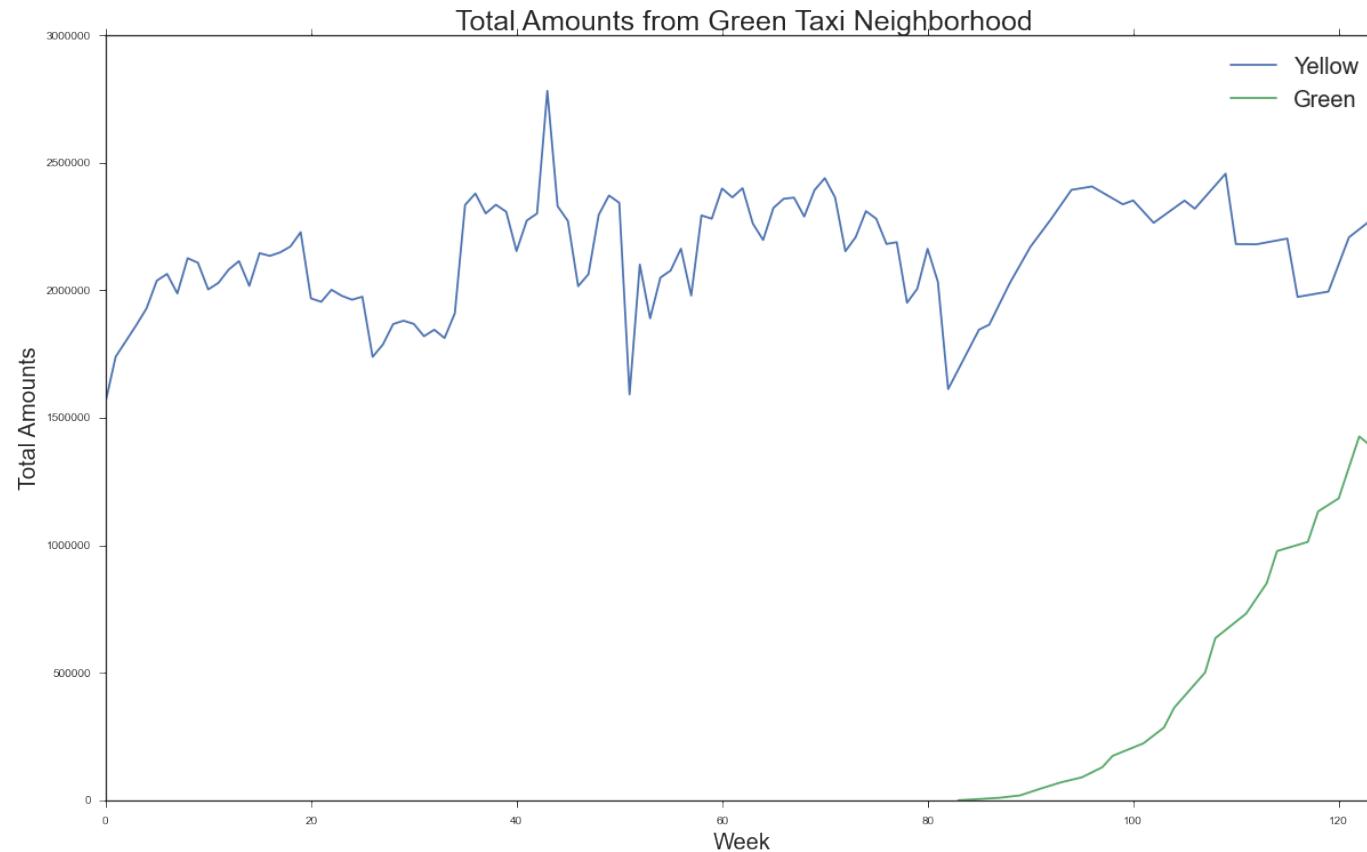
# Motivation

- Yellow taxi drivers protest and file a lawsuit in response.
- Did the introduction of Boro taxis negatively impact the business of the medallion taxis?
- Or did the Boro taxis simply fulfill pent-up demand?

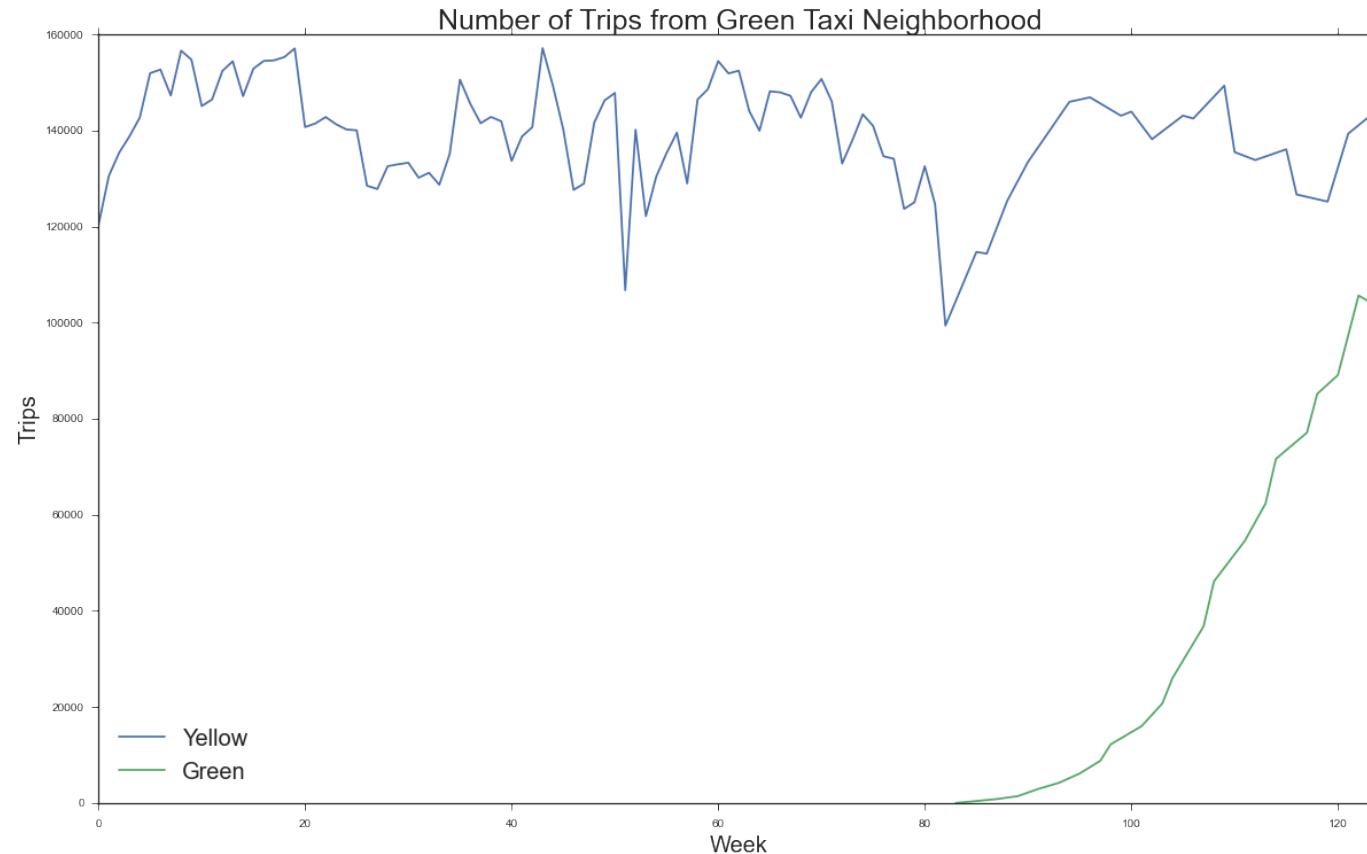
# Results



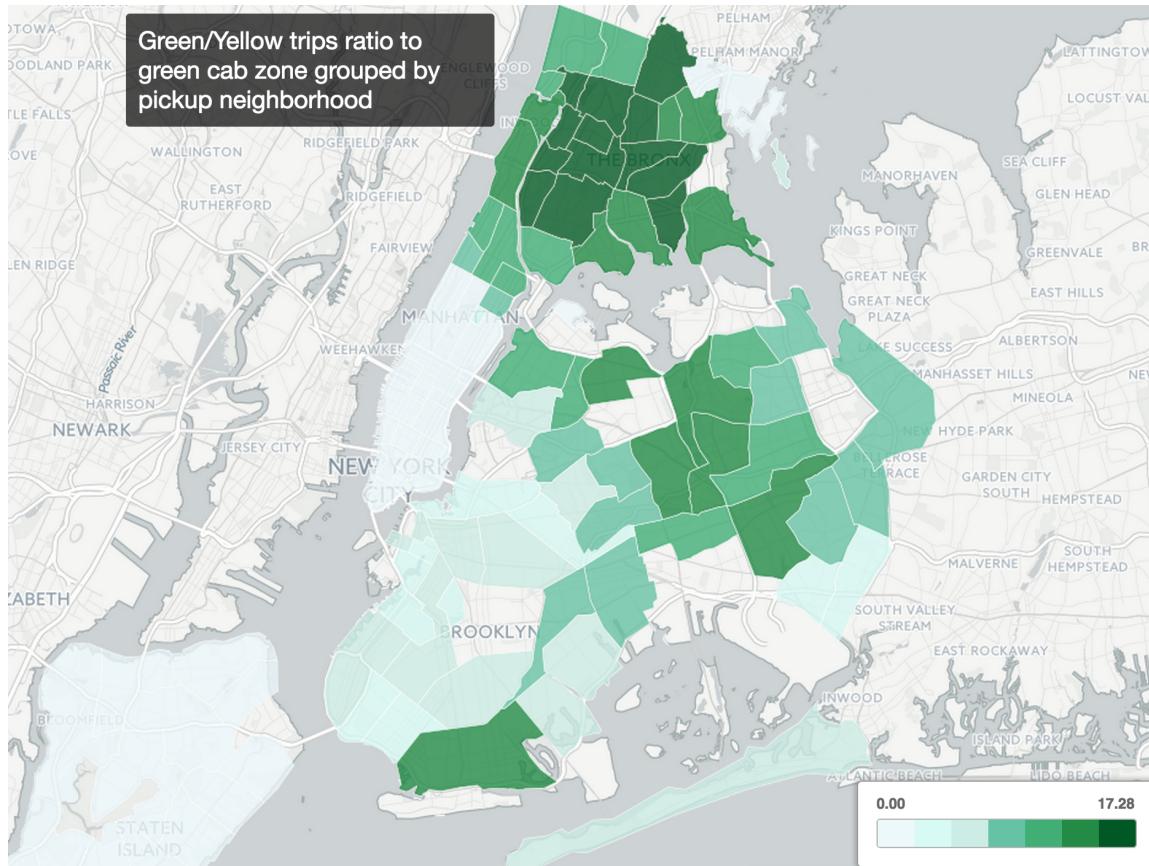
# Results



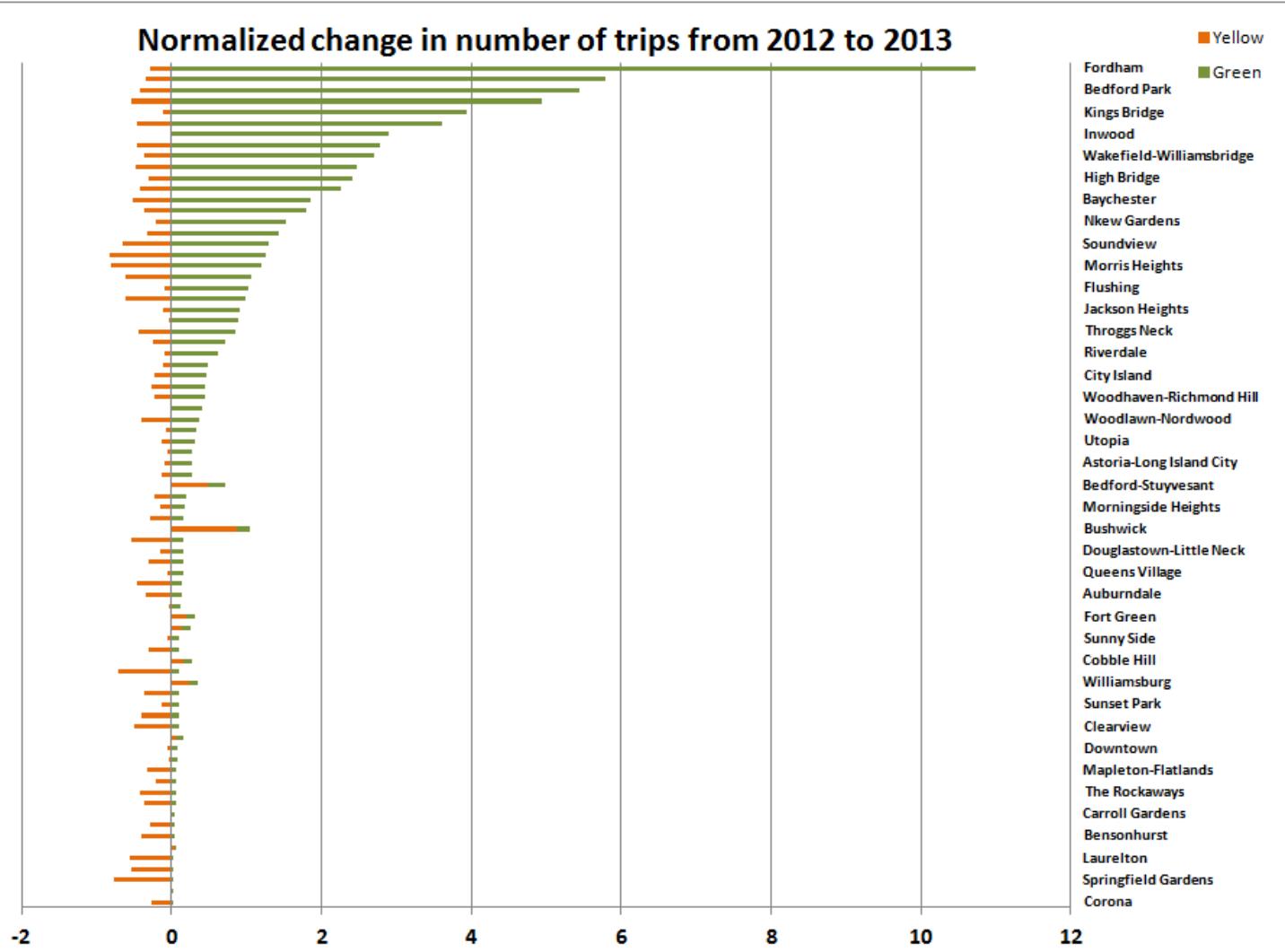
# Results



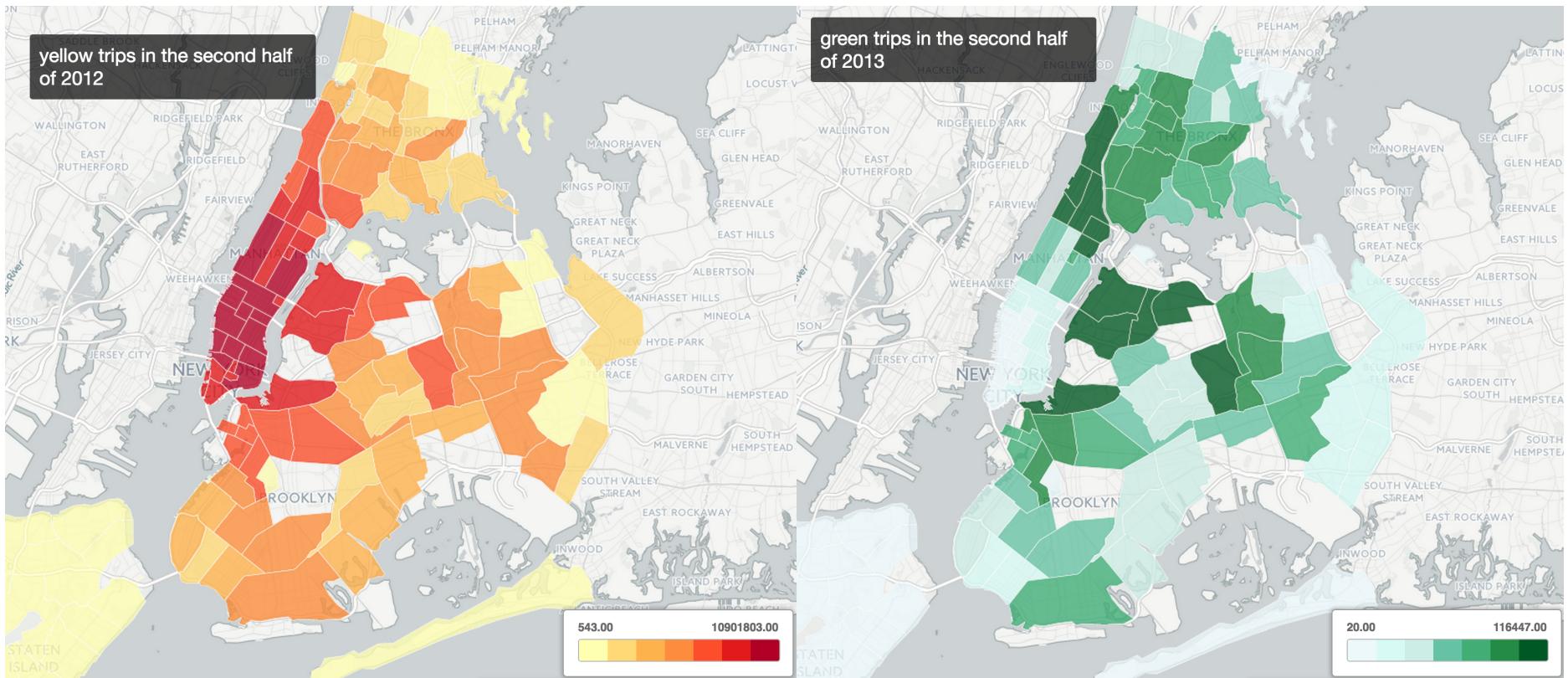
# Results



## Normalized change in number of trips from 2012 to 2013



# Results



# Conclusions

- Boro Taxi succeeded in their stated goal
- Yellow Cabs didn't suffer a loss in business
- By the end of 2013, Boro Taxi accounted for almost 40% of pickups in green zones
- Boro Taxi dominated pickups in the far outer neighborhoods
- In the most extreme case: Fordham experienced a 10-fold increase in pickups

# Methods

Processing:

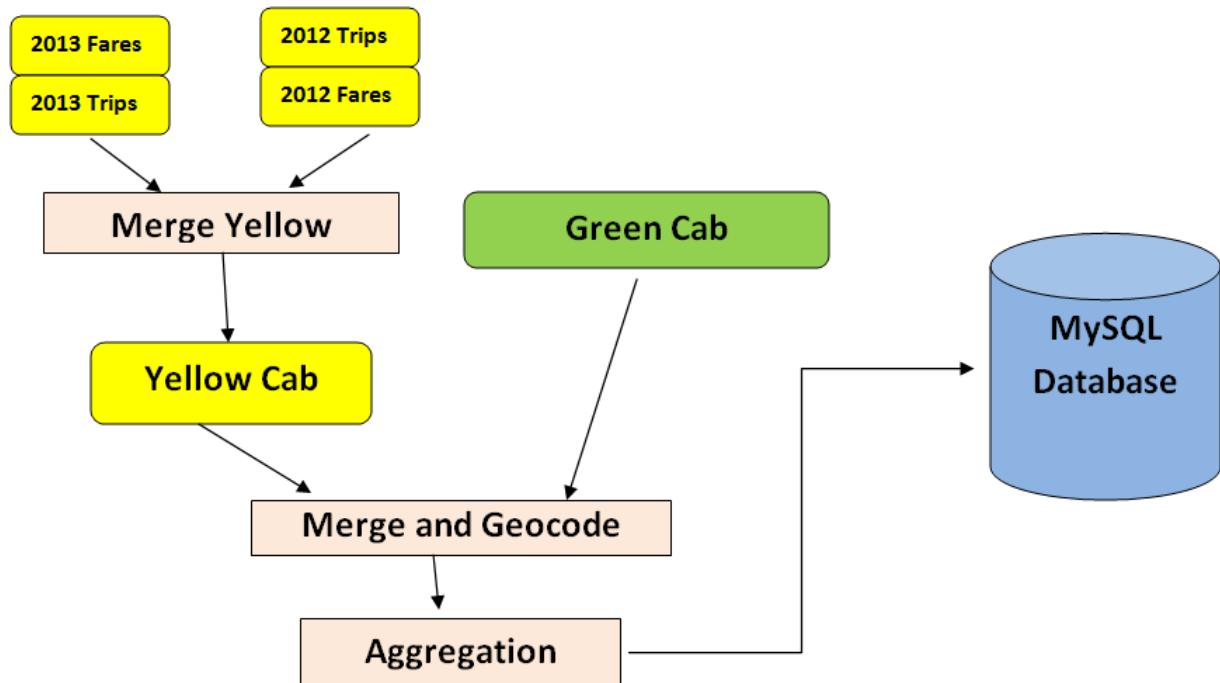
AWS, Python

Analysis:

MySQL

Plotting:

CartoDB, Python



# Methods

- Lots of different big-data analysis tools to choose from
  - Hive, Pig, Spark, etc.
- Realized we didn't need per-trip granularity.
- Aggregated data by date, hour, pick up and drop-off neighborhood.
- Reduced data size from 30 GB to 1.7 GB.
- Allowed us to use MySQL.

# Issues

- 2014 yellow taxi data not yet available.
- missing neighborhoods in shapefile.
- long pipeline forced us to restart the entire process.
- AWS, S3 user interfaces are lacking.