

# Optimizing Capacity Utilization & Location Planning of the Hudson Bike Share System

Albert Alaluf, Alex Caruso, Hongyi Chen, Chong Fu

Instructor: Edward Stohr



**STEVENS**  
INSTITUTE of TECHNOLOGY  
THE INNOVATION UNIVERSITY®

Business Intelligence & Analytics

## Motivation

- Hoboken deployed a bike sharing system in late 2015 called by Hudson Bike Share.
- Subscribed users can rent bikes at any of the 32 stations and return them at any stations within 45 minutes.
- As more users are signing up, Hudson Bike Share is looking to expand and create more stations.
- Some existing stations get little usage, while others are overwhelmed and are short on capacity.

## Technology

- Python for cleansing the raw data, which has pieces of incomplete or unusable data;
- Gephi for generating visualizations of usage for each existing stations;
- Tableau to layer existing stations onto population density map
- K-means Clustering Algorithm to minimize the sum of square rooted distances between existing stations in cluster and new location

## Current & Future Work

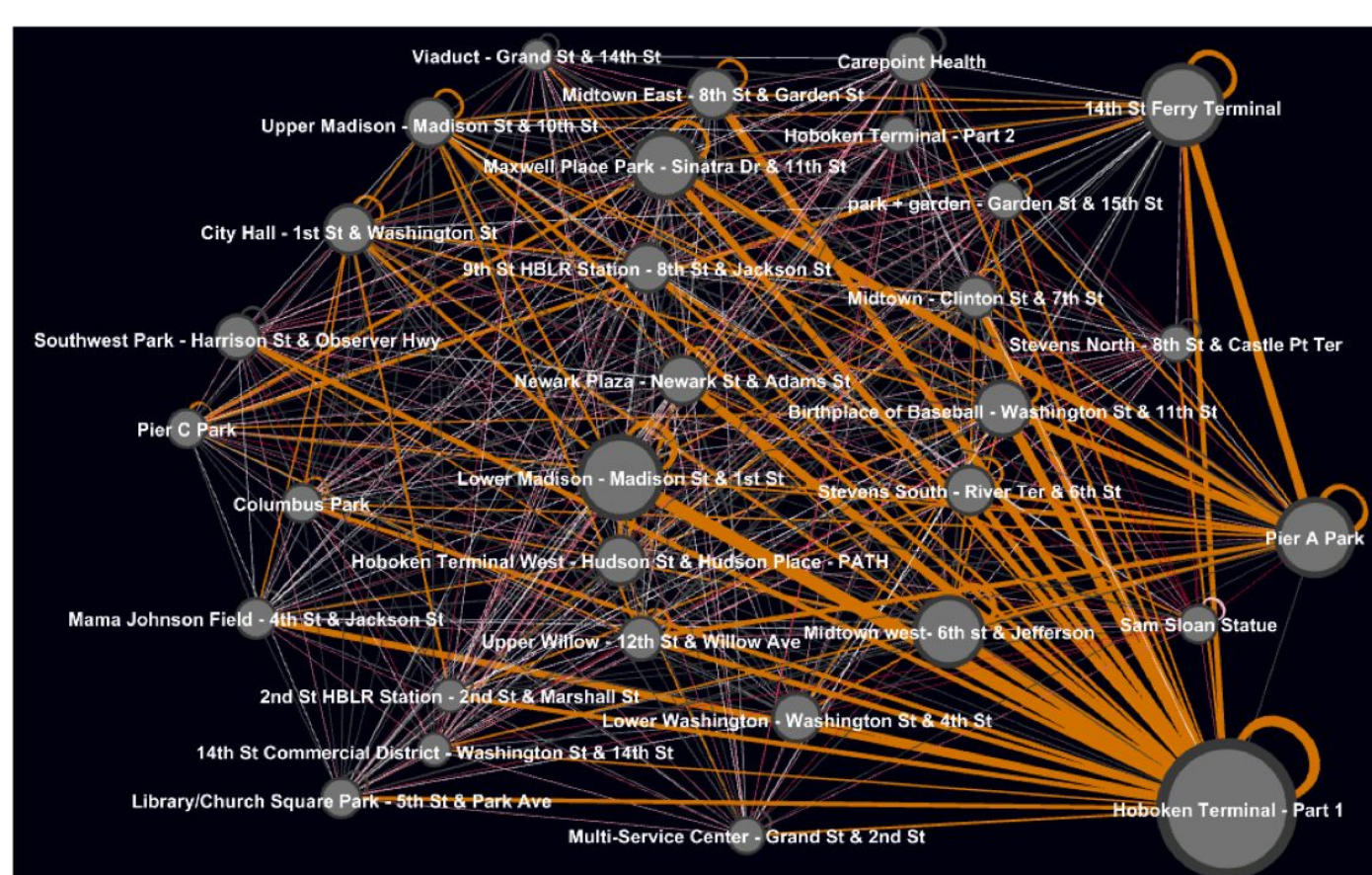
- Cleanse the raw data of trips based on the starting & ending point and time of each trips using Python;
- Visualize the number of trips using Network Statistic Analysis on Gephi;
- Analyze the current total monthly usage of bikes and forecasted the trend of growth using multiple linear regression;
- Implement Gravity Location Model based on population density map of Hoboken;
- Cluster the existing stations based on the heatmap using K-Means Clustering Algorithm;
- Suggest Optimal locations for new stations and/or relocations for existing stations based on the Gravity Location Model.

## Data & Scope

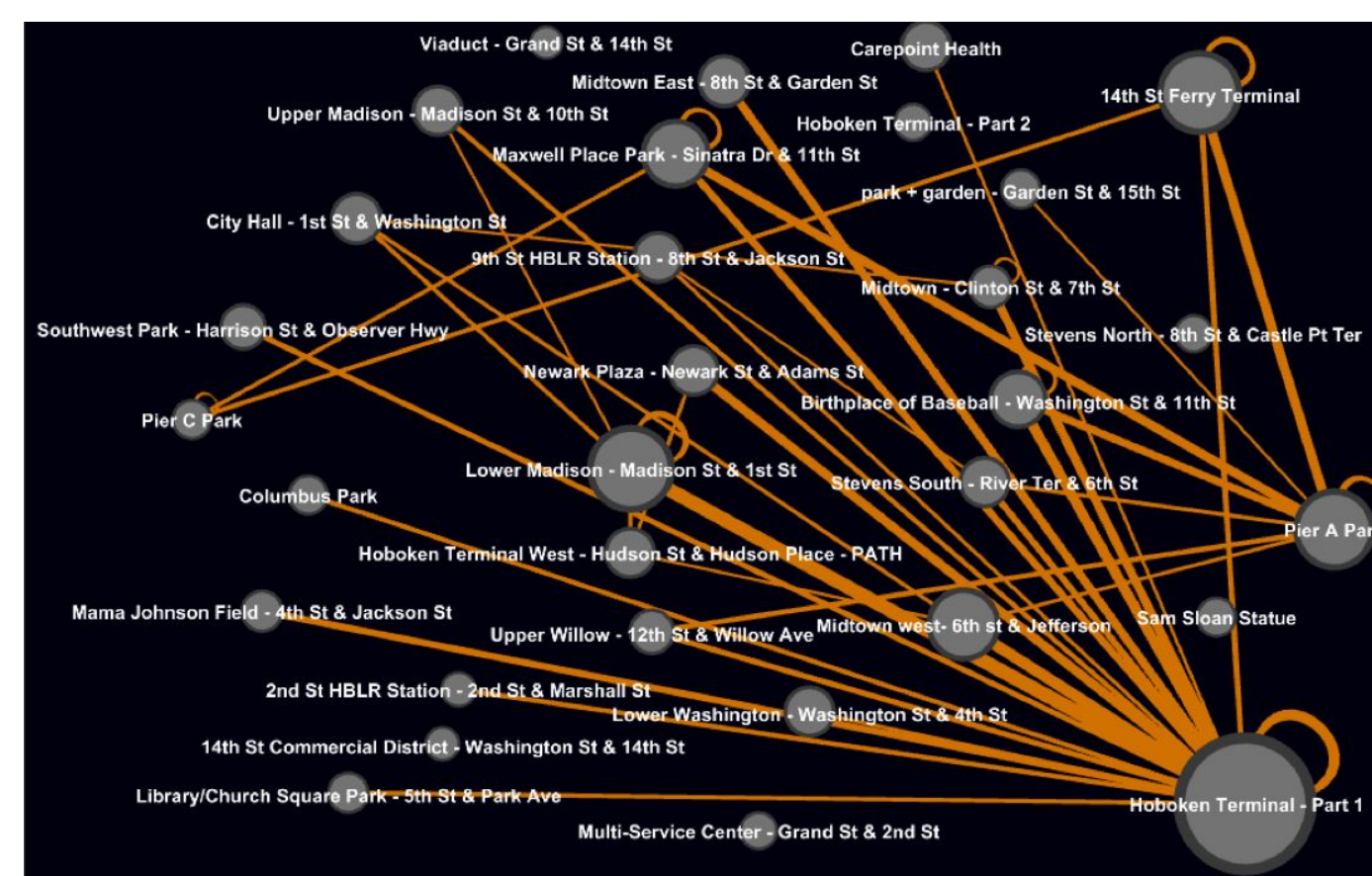
1919	Maxwell Place Park - Sinatra Dr & 11th St	1919	Maxwell Place Park - Sinatra Dr & 11th St
1901	Hoboken Terminal West - Hudson St & Hudson Pl - PATH	1901	Hoboken Terminal West - Hudson St & Hudson Pl - PATH
1900	Hoboken Terminal - Part 1	1920	Birthplace of Baseball - Washington St & 11th St
1902	Pier A Park	1902	Pier A Park
1926	Lower Madison - Madison St & 1st St	1901	Hoboken Terminal West - Hudson St & Hudson Pl - PATH
1907	Stevens South - River Ter & 6th St	1907	Stevens South - River Ter & 6th St
1921	Upper Willow - 12th St & Willow Ave	1921	Upper Willow - 12th St & Willow Ave
1921	Upper Willow - 12th St & Willow Ave	1921	Upper Willow - 12th St & Willow Ave
1921	Upper Willow - 12th St & Willow Ave	1907	Stevens South - River Ter & 6th St
1900	Hoboken Terminal - Part 1	1900	Hoboken Terminal - Part 1
1916	City Hall - 1st St & Washington St	1916	City Hall - 1st St & Washington St
1900	Hoboken Terminal - Part 1	1926	Lower Madison - Madison St & 1st St
1902	Pier A Park	1902	Pier A Park
1921	Upper Willow - 12th St & Willow Ave	1921	Upper Willow - 12th St & Willow Ave
1900	Hoboken Terminal - Part 1	1900	Hoboken Terminal - Part 1

```
1 ##CLEAN DATA TO CREATE SIMPLIFIED GEPHI NETWORK WITH ONLY ACTIVE STATIONS & PHYSICAL TO PHYSICAL CONNECTIONS
2 import csv
3 with open("raw_data.csv", "r") as f:
4     reader=csv.reader(f, delimiter=',')
5     rows=[(row[5], row[6], row[7], row[8]) for row in reader]
6
7 physical_trips = []
8 for row in rows[1:]:
9     if row[0] != '' and row[2] != '':
10         physical_trips.append(row)
11
12 active_station_ids = range(1900,1914) + range(1916,1931) + range(1949,1952)
13 active_station_trips = []
14 for row in physical_trips:
15     if int(row[0]) in active_station_ids and int(row[2]) in active_station_ids:
16         active_station_trips.append(row)
17
18 with open("active_station_trips.csv", "w") as f:
19     writer=csv.writer(f,delimiter=',')
20     writer.writerows(active_station_trips)
```

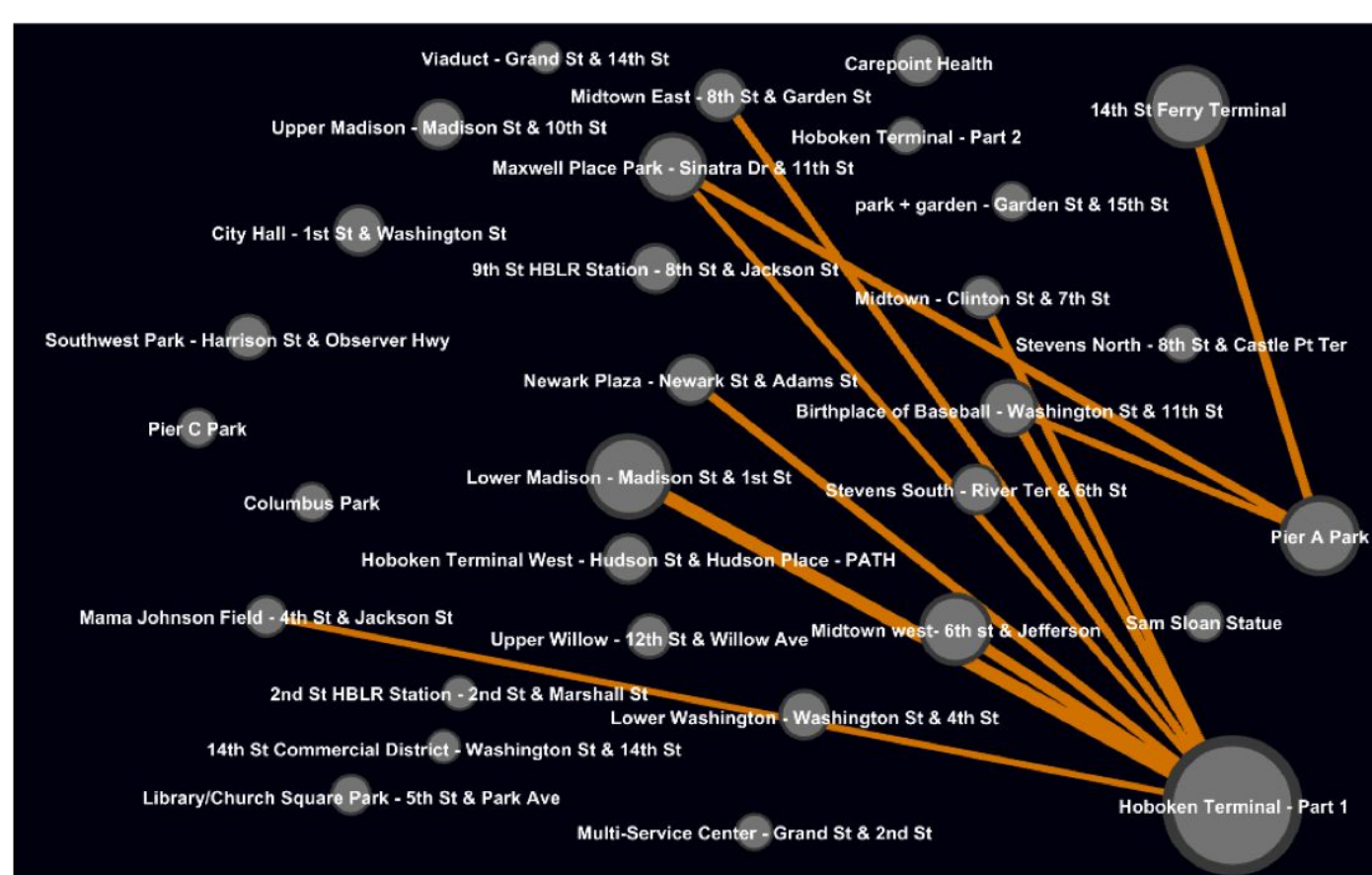
## Usage Visualizing - Network Statistic Analysis



Original Station Usage Network



Filtered Station Usage Network (1)

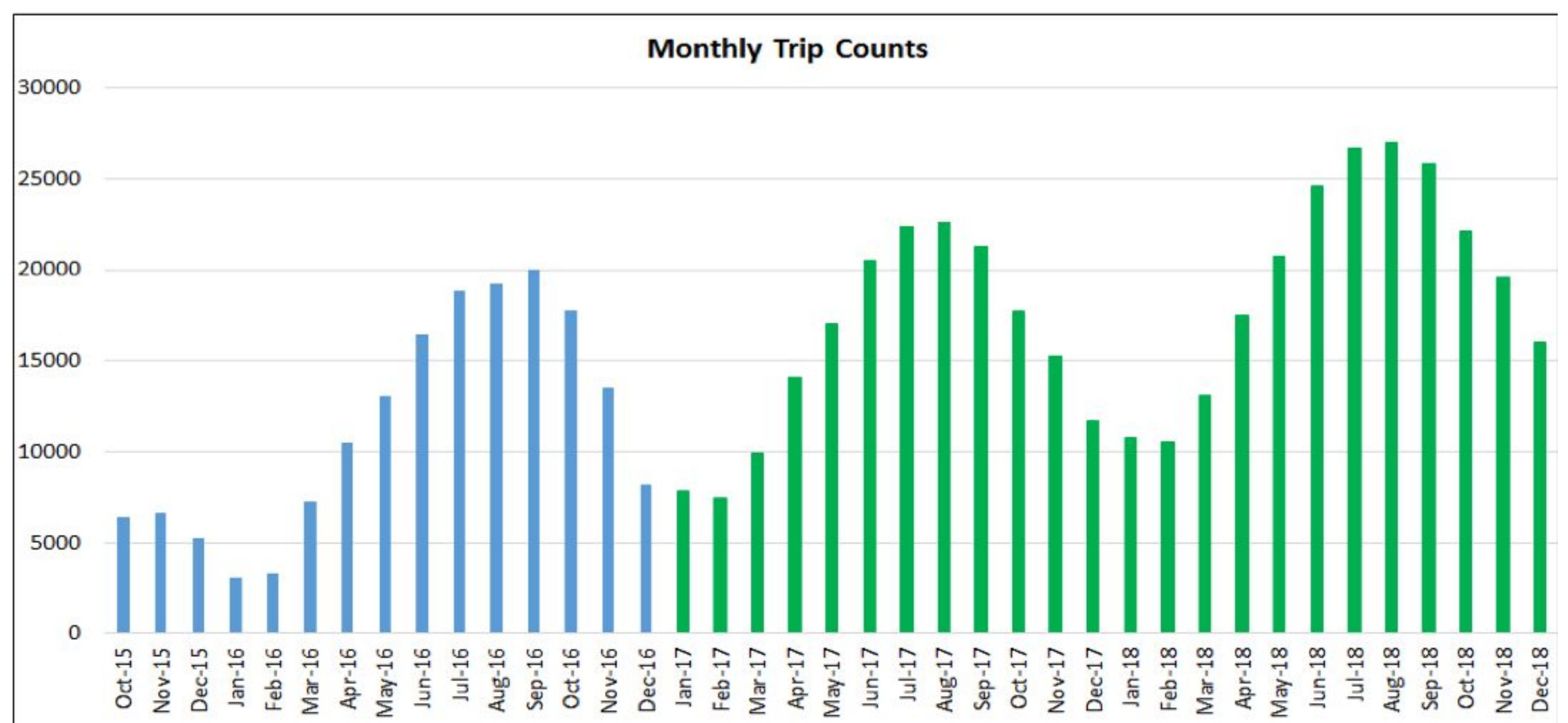


Filtered Station Usage Network (2)

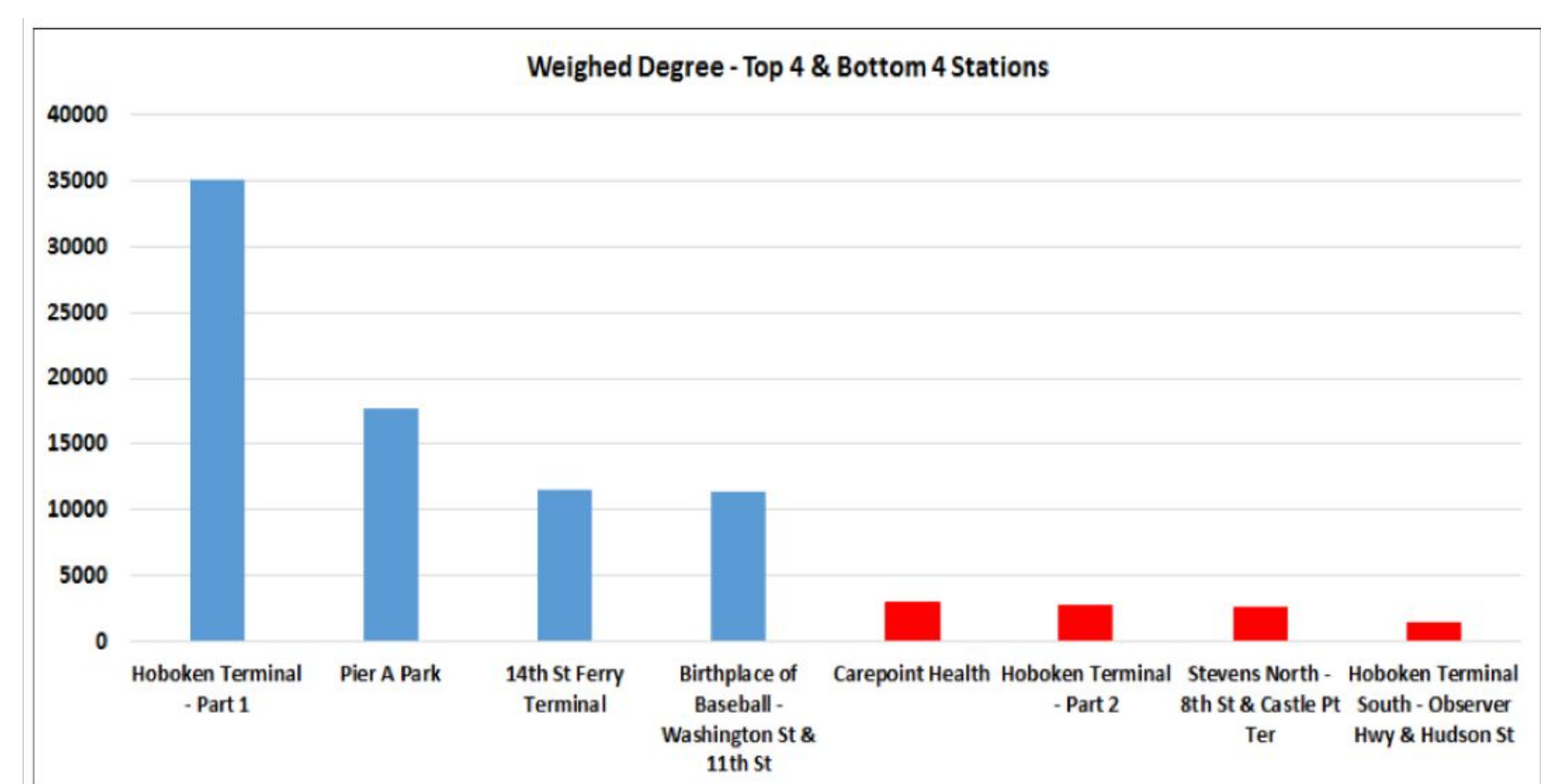


Timed Station Usage

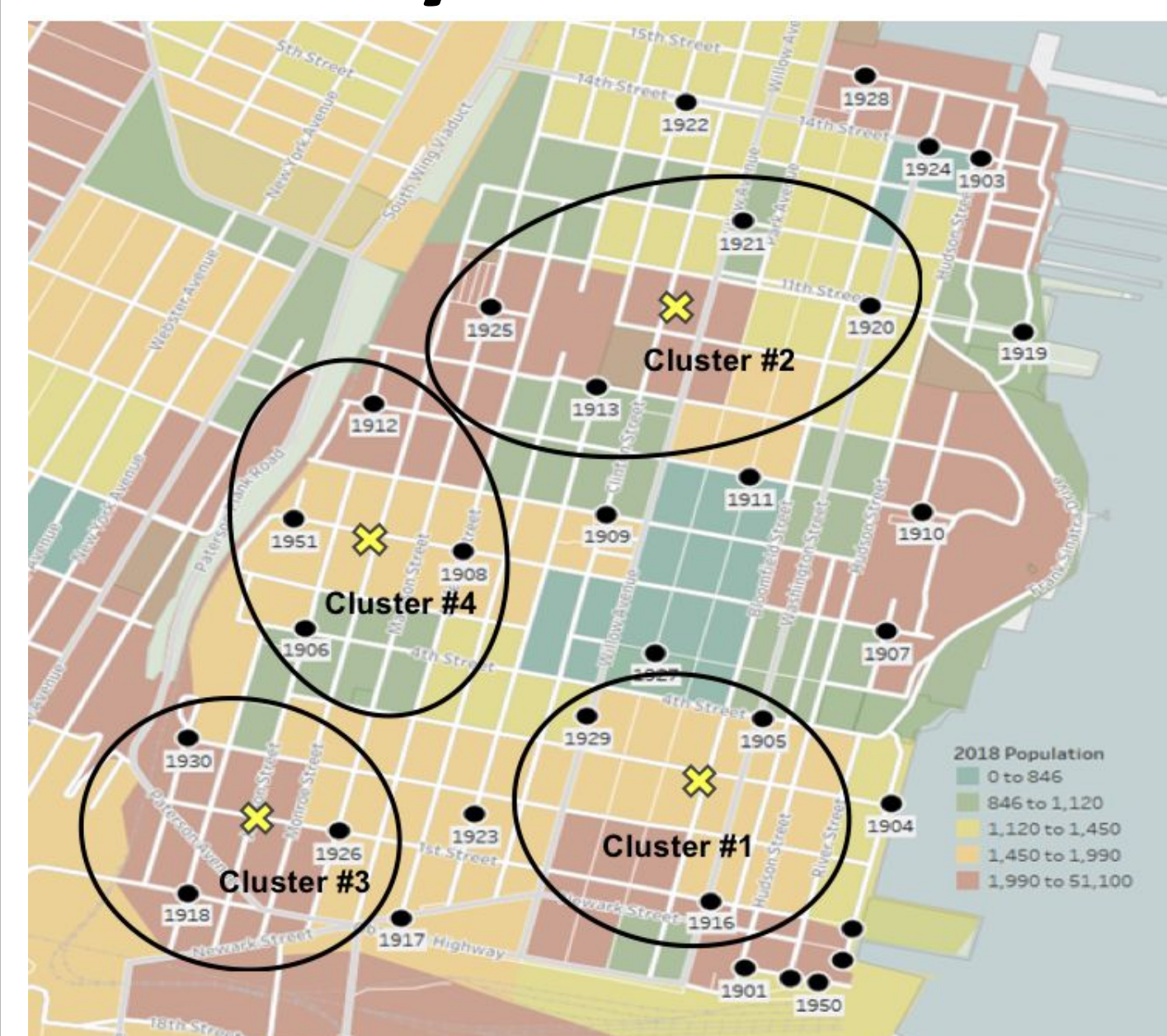
## Usage Forecasting - Multiple Linear Regression



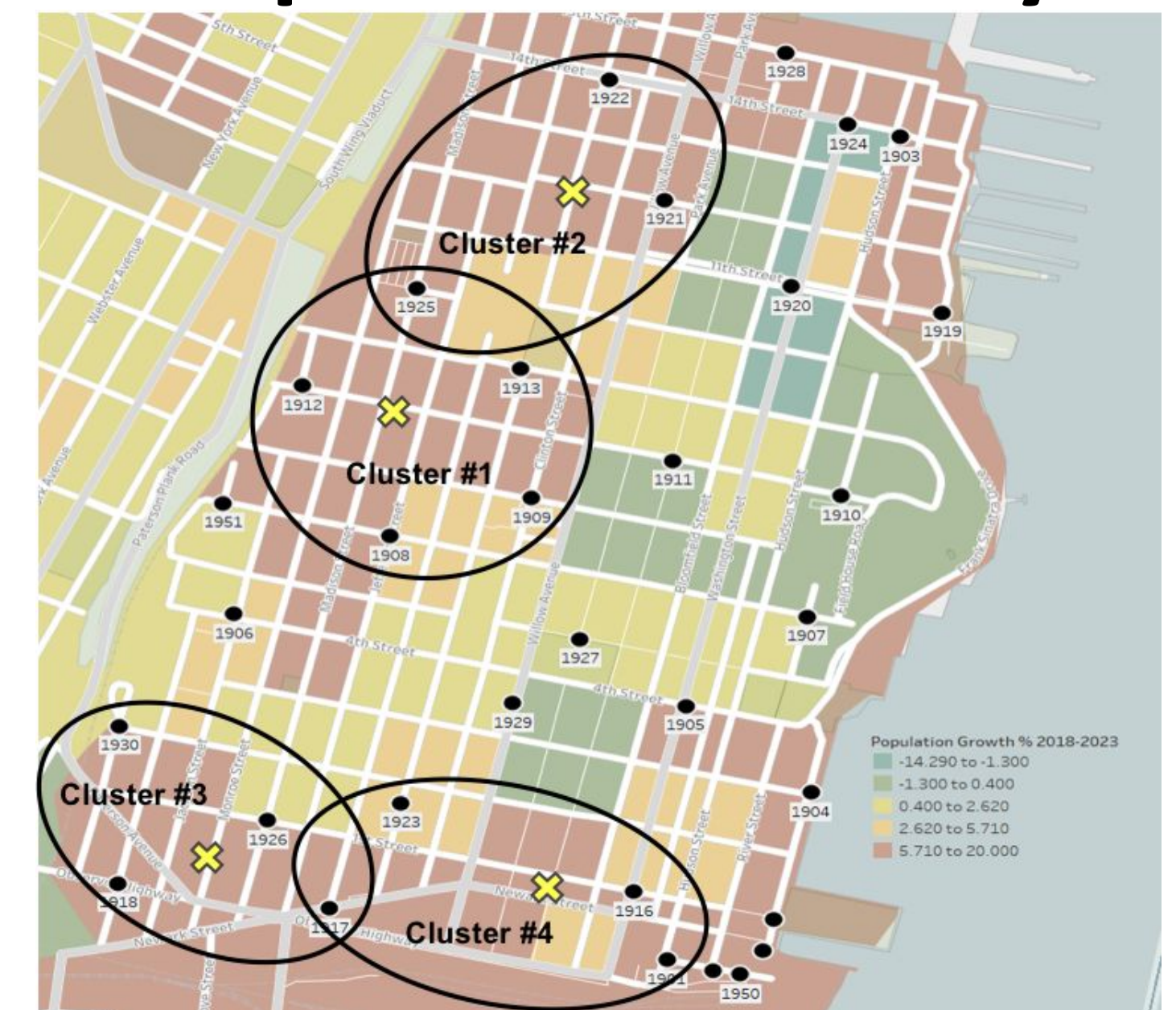
## Weighted Degree – Top 4 & Bottom 4 Stations



## Gravity Location Model - Population Density



2018 population density



2018 - 2023 expected % growth