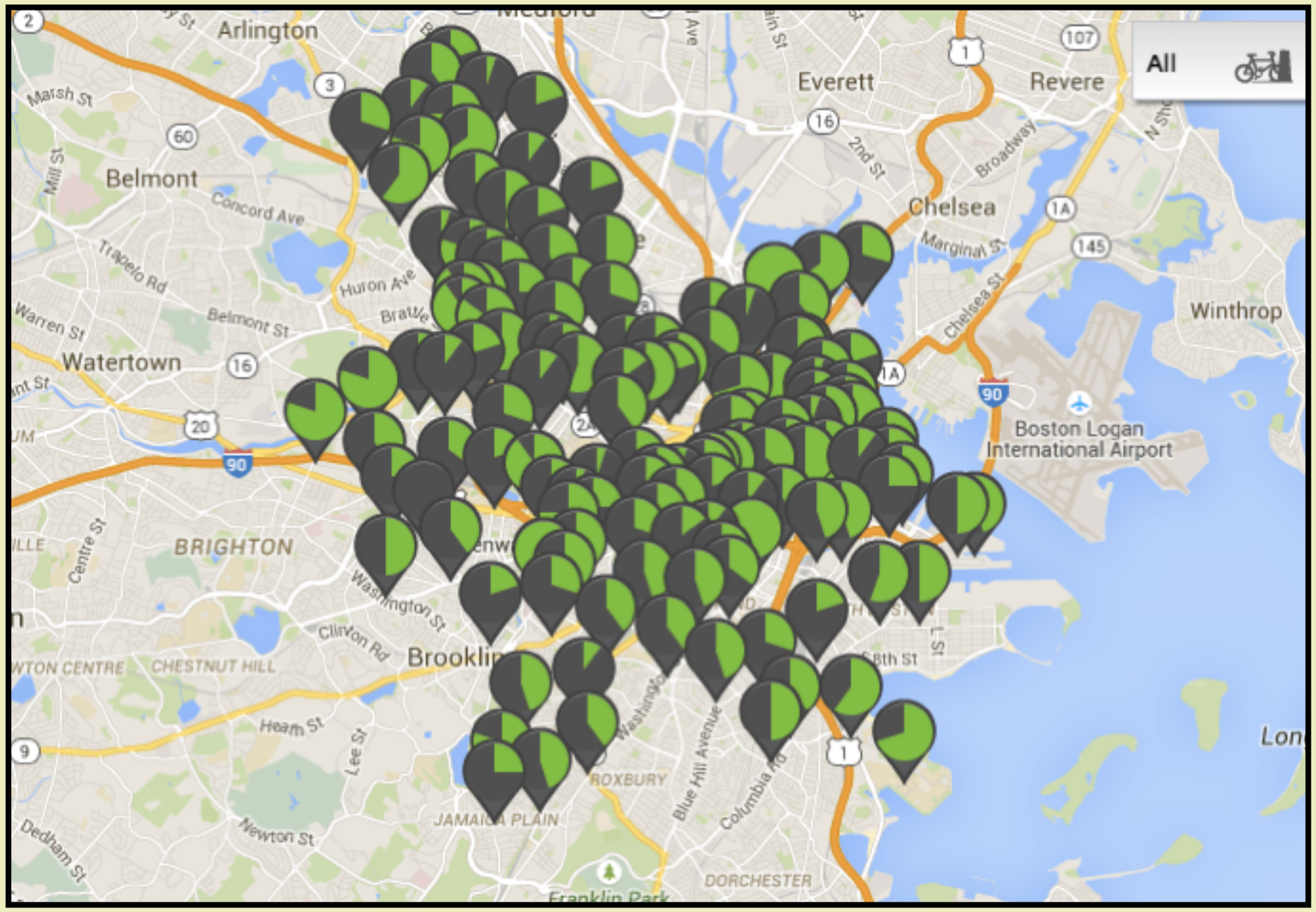
Grow Hubway: Where should Hubway expand next?

Grow Hubway: Where should Hubway expand next?



Data

Hubway Data

- Publicly available: July 2011 through November 2013
- Time and location for starting and ending station
- ~140 stations

Ancillary Data

- US Census data by zip code
  - Population size from 2010
  - Employee size from 2012
- MBTA rider data from 2013
  - PDF scraping
  - http://erikdemaine.org/maps/mbta/mbta.yaml
- Mask for Charles river from OpenStreetMap data

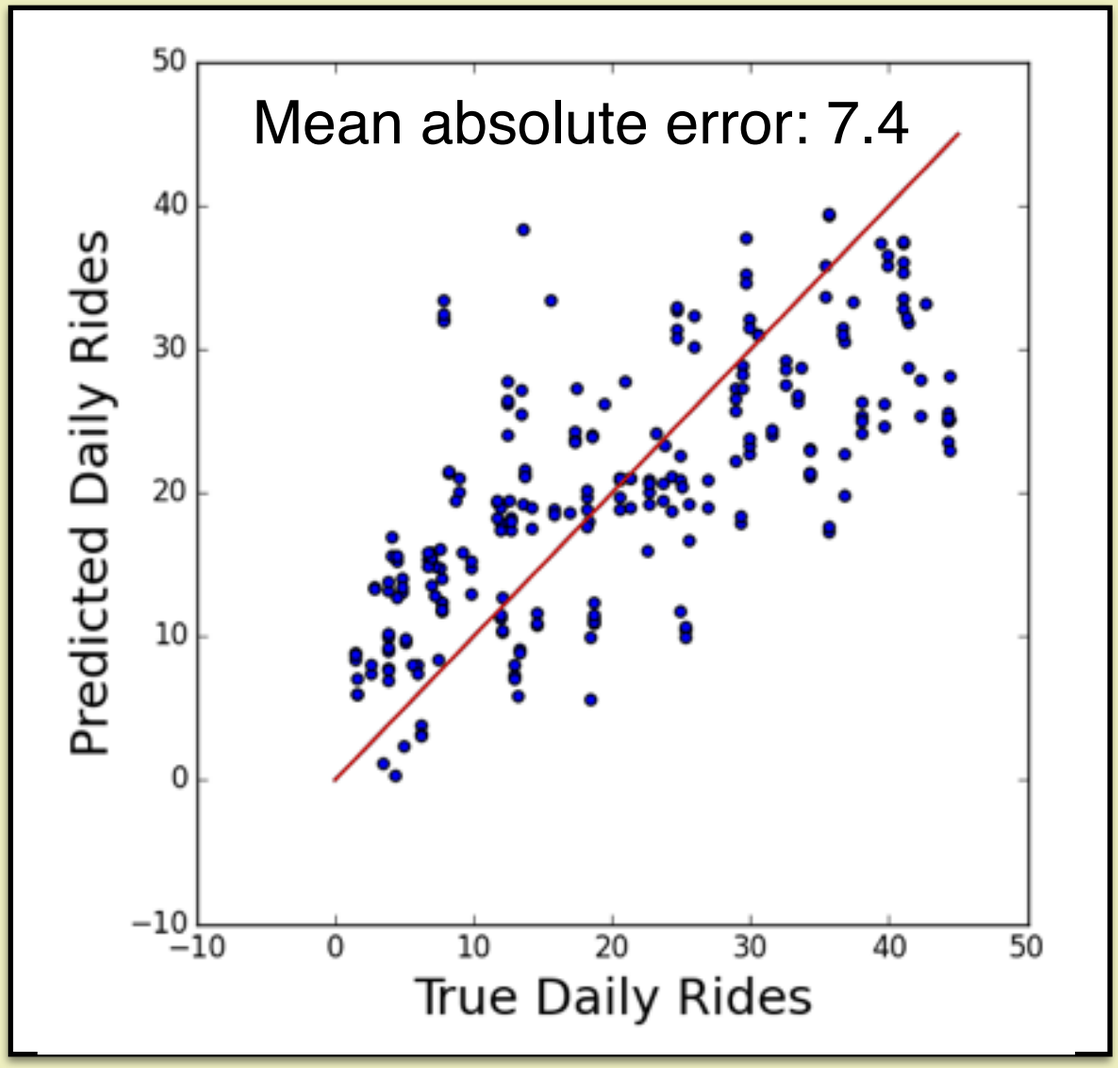
Algorithmic Approach

Features

- Population score
- Employee score
- MBTA subway score
- “origin” score and “destination” score for each of the above features

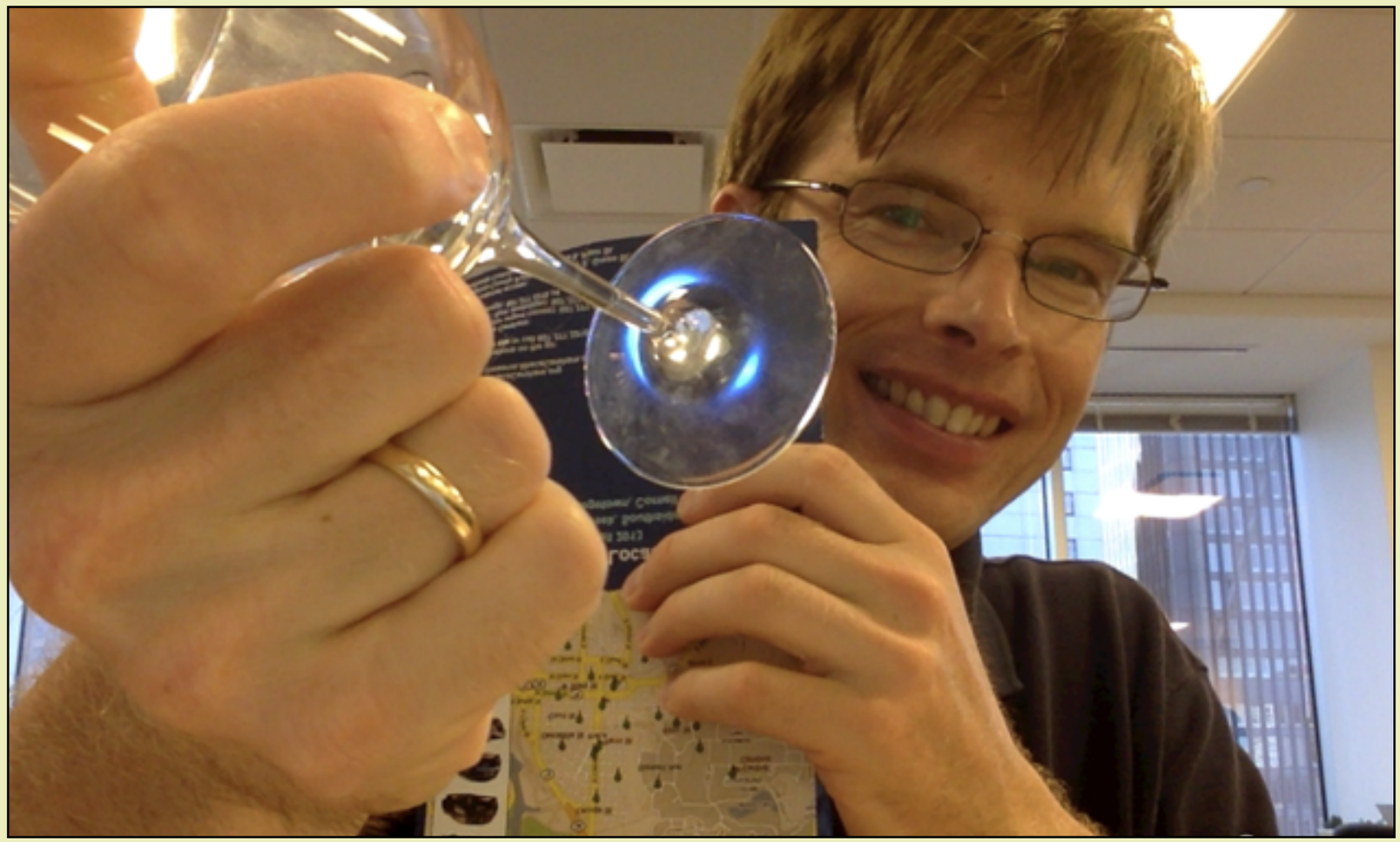
Machine Learning approach

- Linear regression: predict average daily rides
- 5-fold cross-validation
- prescription for cannibalism



About Me

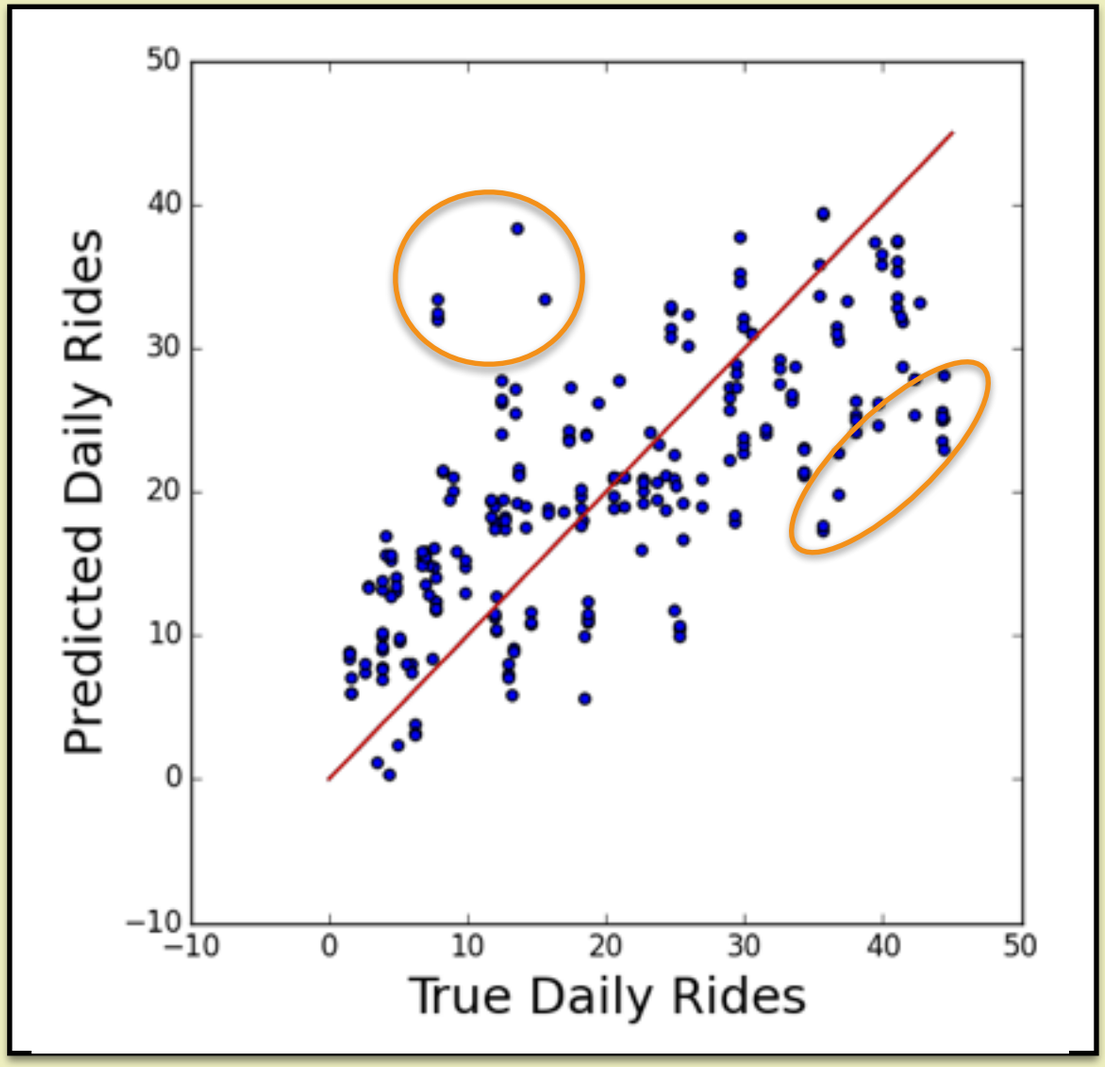
- PhD, Astronomy, University of Arizona, 2010
- Postdoc at Harvard-Smithsonian Center for Astrophysics
- Postdoc at Cornell University
- Research topic: understand how stars form in distant galaxies by studying gravitationally lensed galaxies



OLS Regression Model

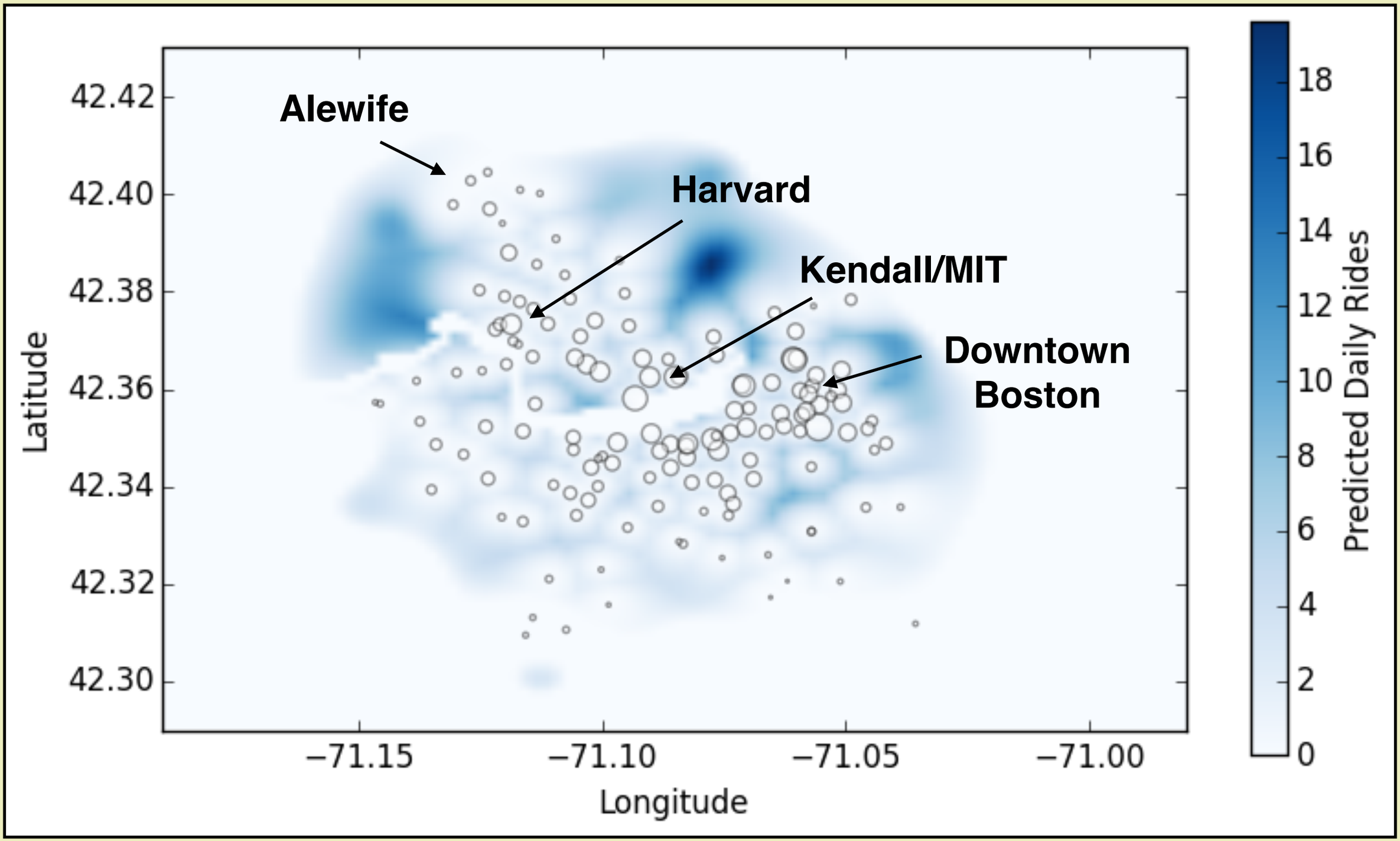
Outliers

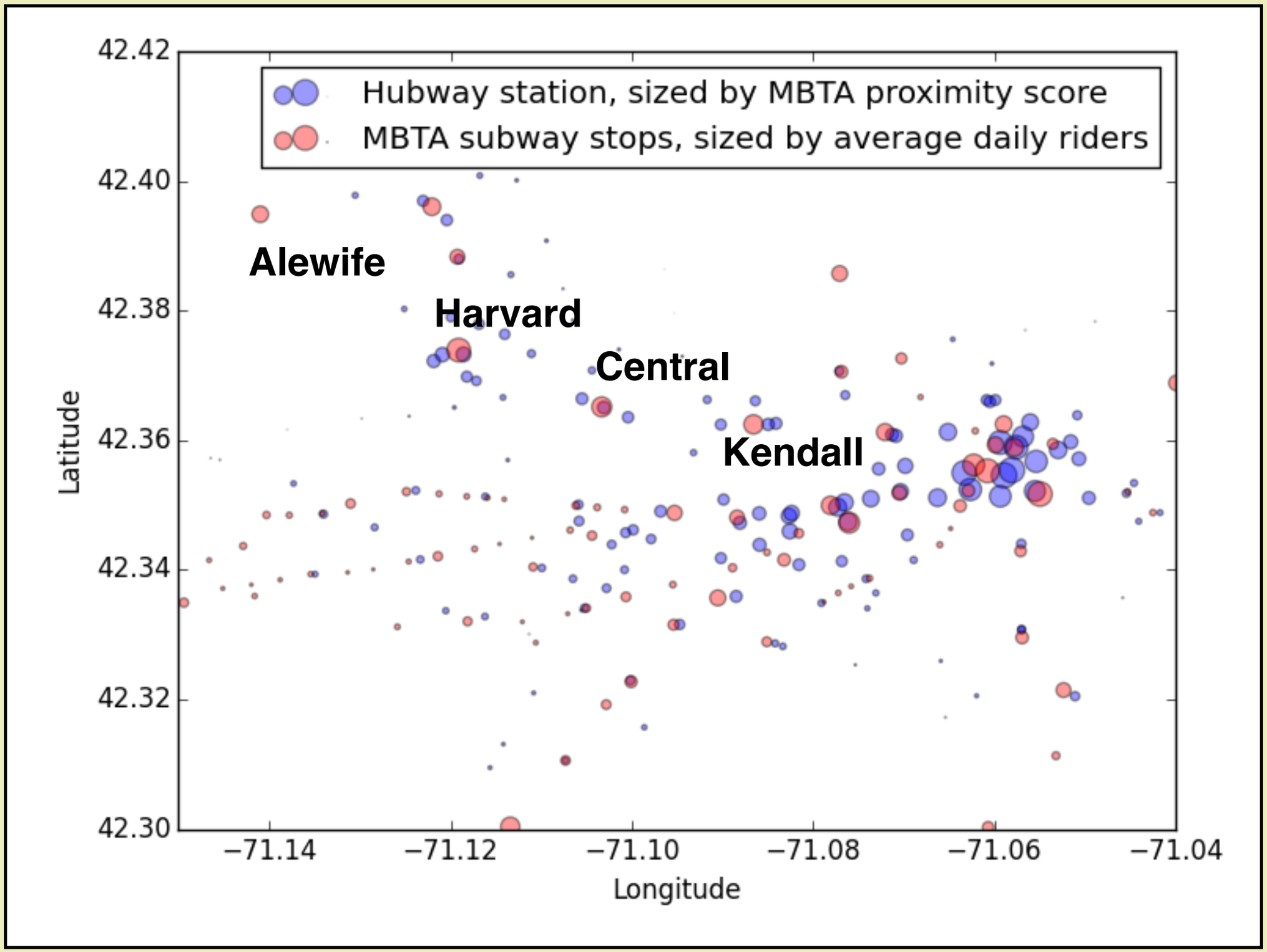
- Mean absolute error: 7.4
- Median absolute error: 5.9
- Action item: identify outliers in ($N_{predict}$, N_{true}) space and search for patterns

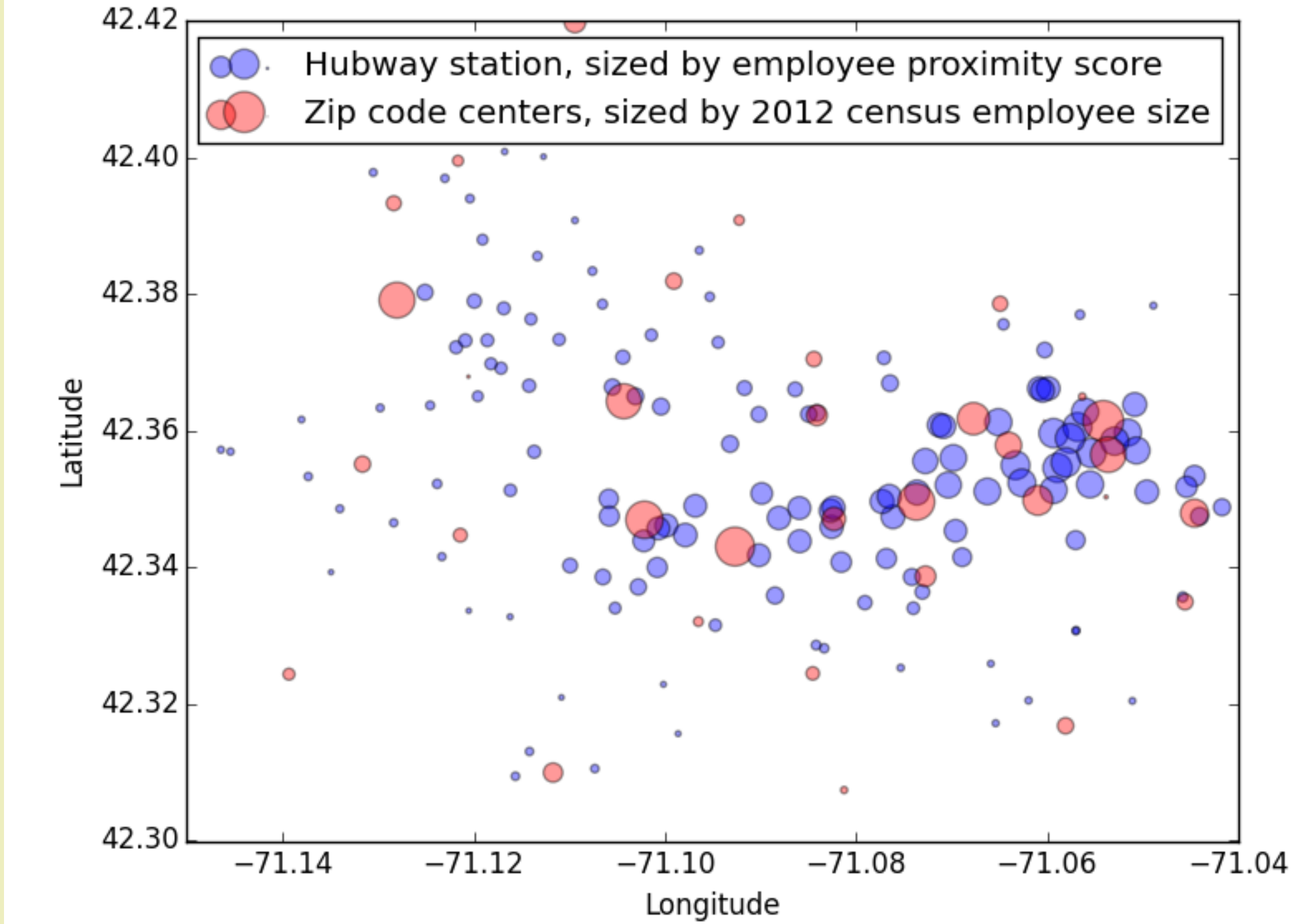


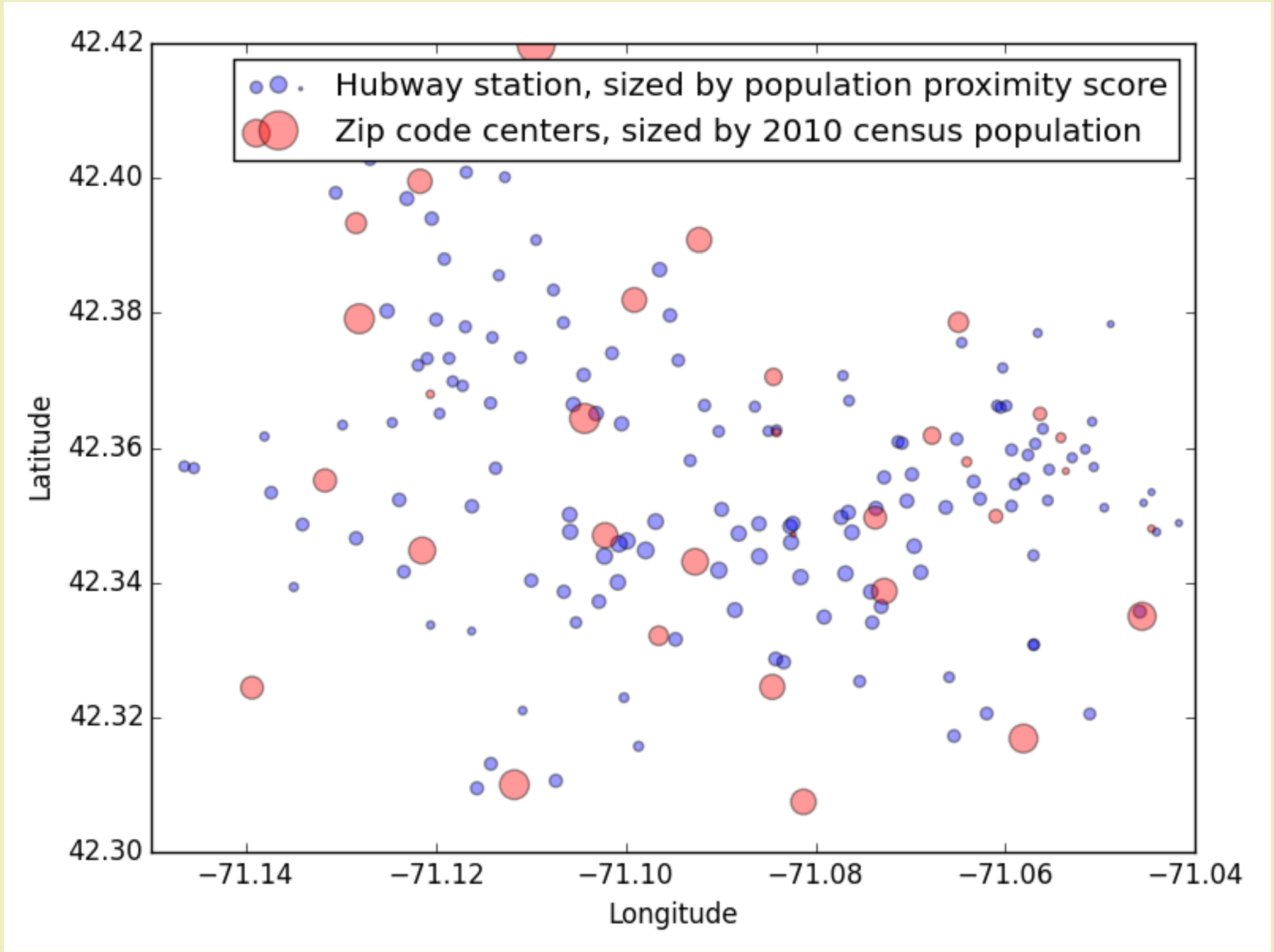
Feature Importance

	Origin	Destination
Population	-1.316	6.643
Employee	1.481	-50.954
Subway	6.643	48.217


Alewife
Harvard
Kendall/MIT
Downtown
Boston
42.42
42.40
42.38
42.36
42.34
42.32
42.30
Latitude
−71.15
−71.10
−71.05
−71.00
Longitude
18
16
14
12
10
8
6
4
2
0
Predicted Daily Rides


Hubway station, sized by MBTA proximity score
MBTA subway stops, sized by average daily riders
Alewife
Harvard
Central
Kendall
42.42
42.40
42.38
42.36
42.34
42.32
42.30
Latitude
−71.14
−71.12
−71.10
−71.08
−71.06
−71.04
Longitude


Hubway station, sized by employee proximity score
Zip code centers, sized by 2012 census employee size
42.42
42.40
42.38
42.36
42.34
42.32
42.30
Latitude
−71.14
−71.12
−71.10
−71.08
−71.06
−71.04
Longitude


Hubway station, sized by population proximity score
Zip code centers, sized by 2010 census population
42.42
42.40
42.38
42.36
42.34
42.32
42.30
Latitude
−71.14
−71.12
−71.10
−71.08
−71.06
−71.04
Longitude