

A map of New York City, specifically the area from Manhattan to Brooklyn, showing various restaurant clusters. The clusters are represented by colored circles: green circles in the upper left (Manhattan), yellow circles in the upper right (Brooklyn), red circles in the lower left (Manhattan), and white circles in the lower right (Brooklyn). The map includes street names like 9th Street, 3rd Avenue, 1st Avenue, Broadway, and FDR Drive, as well as landmarks like the Manhattan Bridge and Brooklyn Bridge. A semi-transparent white box with a blue border is centered over the map, containing the title and descriptive text.

# Restaurant Clusters

Creating clusters of venue types and visualize them based on their average rating.

The example of restaurants in New York



# Everybody Wants Venue Clusters Visualized



## Individuals

- Tourists
- Newcomers
- Locals



## Venue Operators

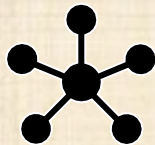
- Restaurants
- Bars
- Coffee Places



## Investors



## Media



## City Planners

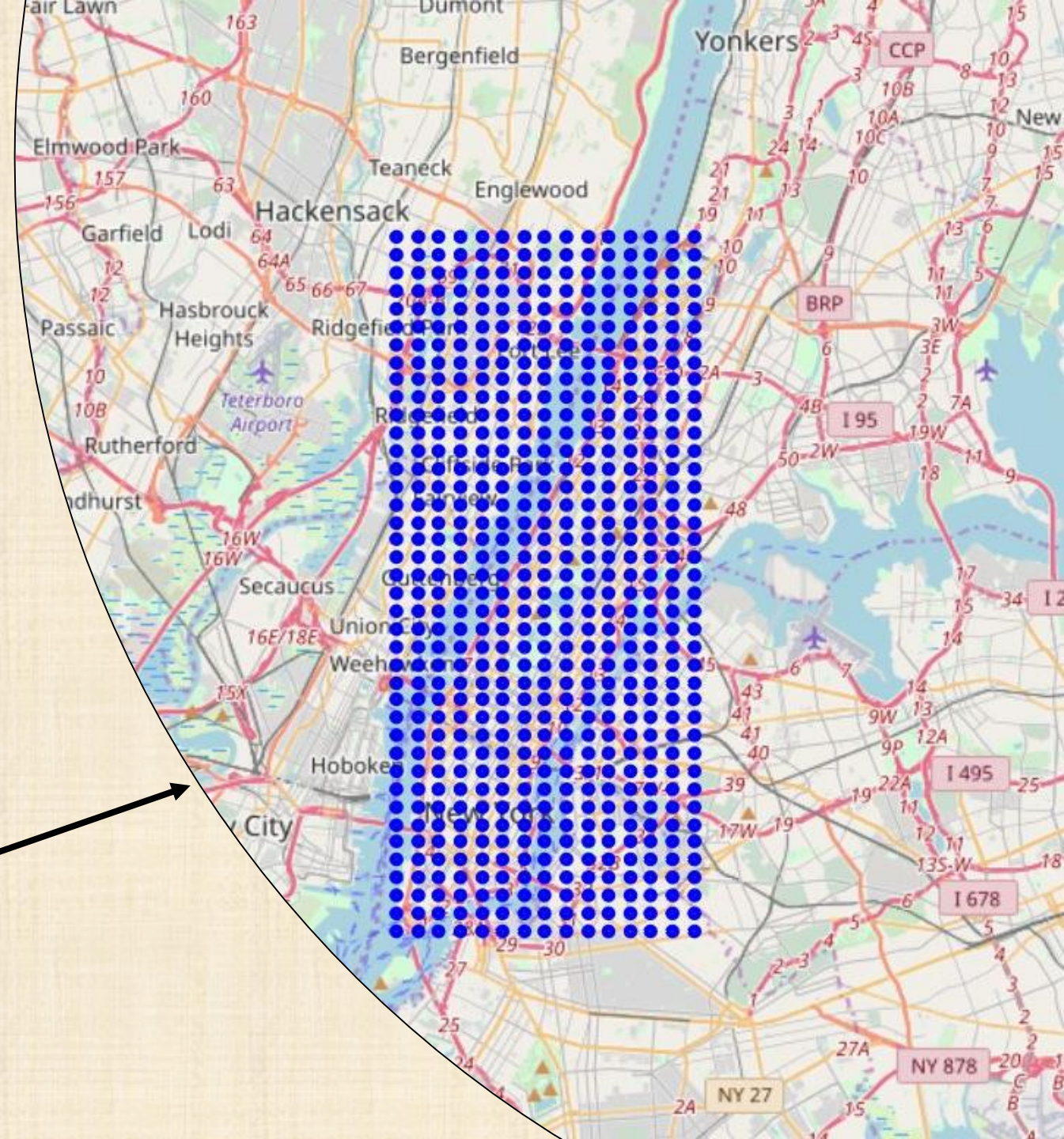


## Guide Editors

# Data

---

- Content: Restaurant location data with ratings
- Source: Foursquare API
- Problem: Limitation to 50 venues per search
- Solution: Grid with location points

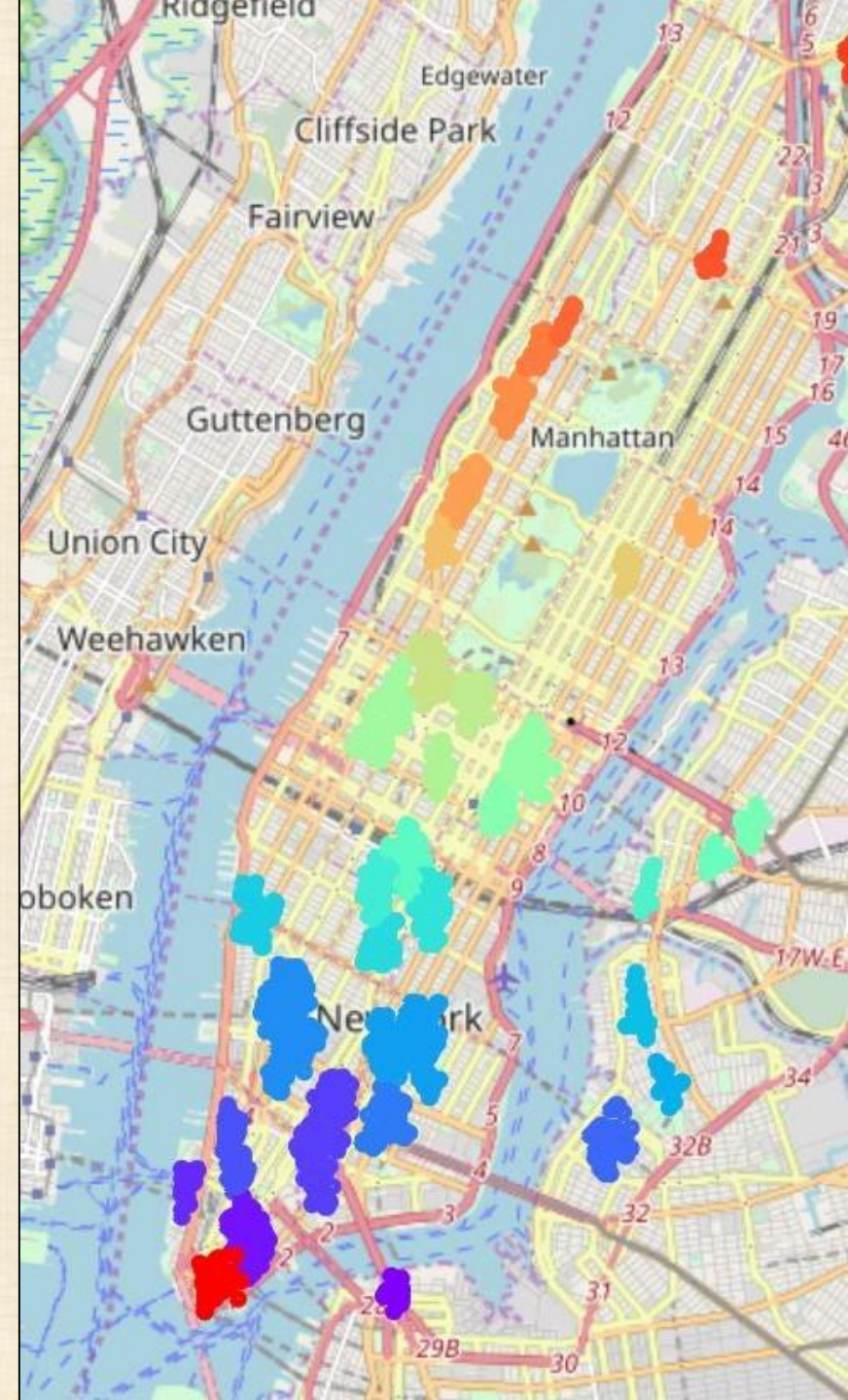




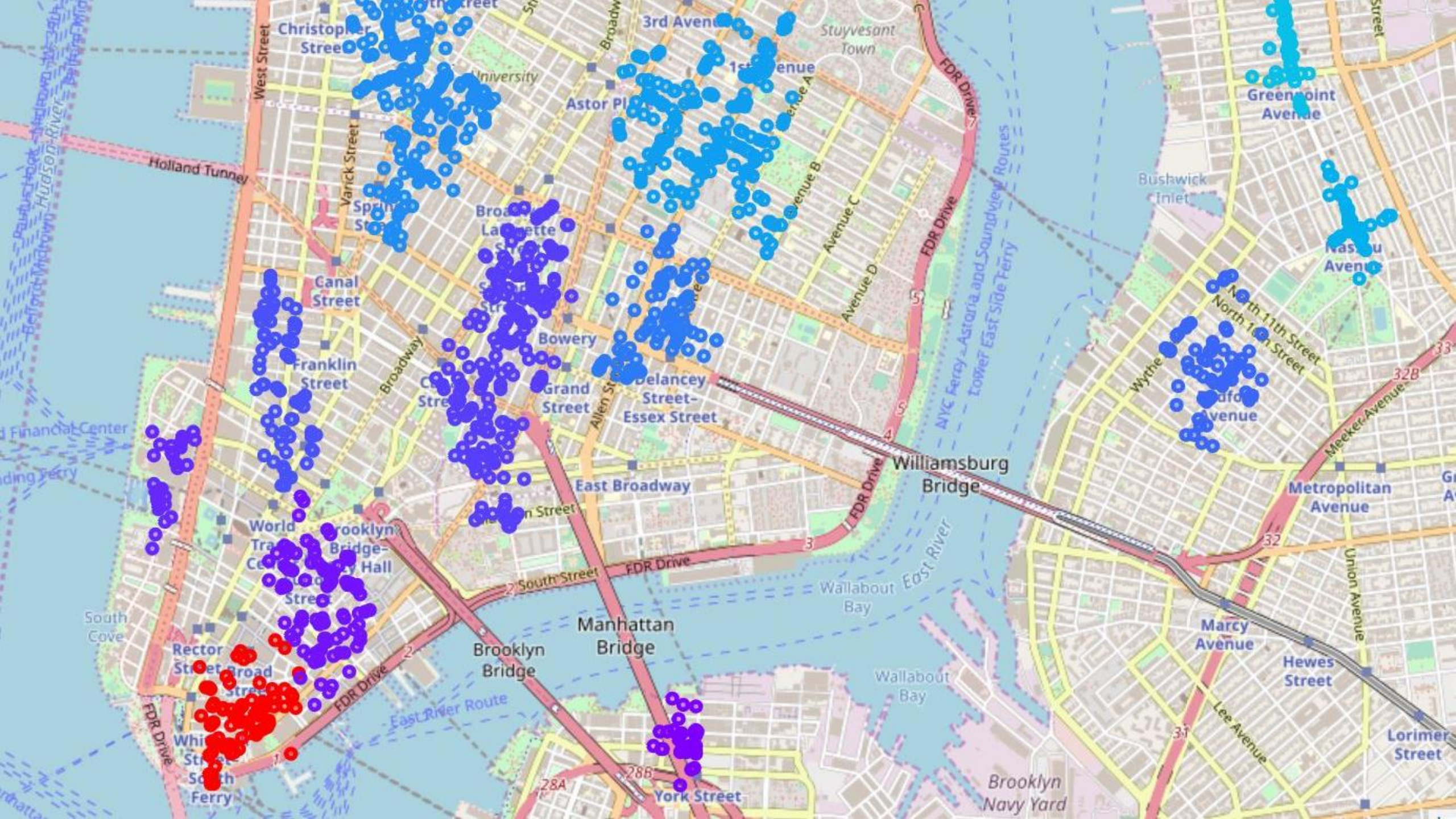
# Clustering

---

- Type: – Density-Based Clustering
- Algorithm: – DBSCAN from Scikit-learn
- Reasons: – Allows arbitrary shapes  
– Identification of Outliers  
– No specification of no. of clusters
- Parameters – Radius  $R$   
– Number within  $R$









# Rating Data

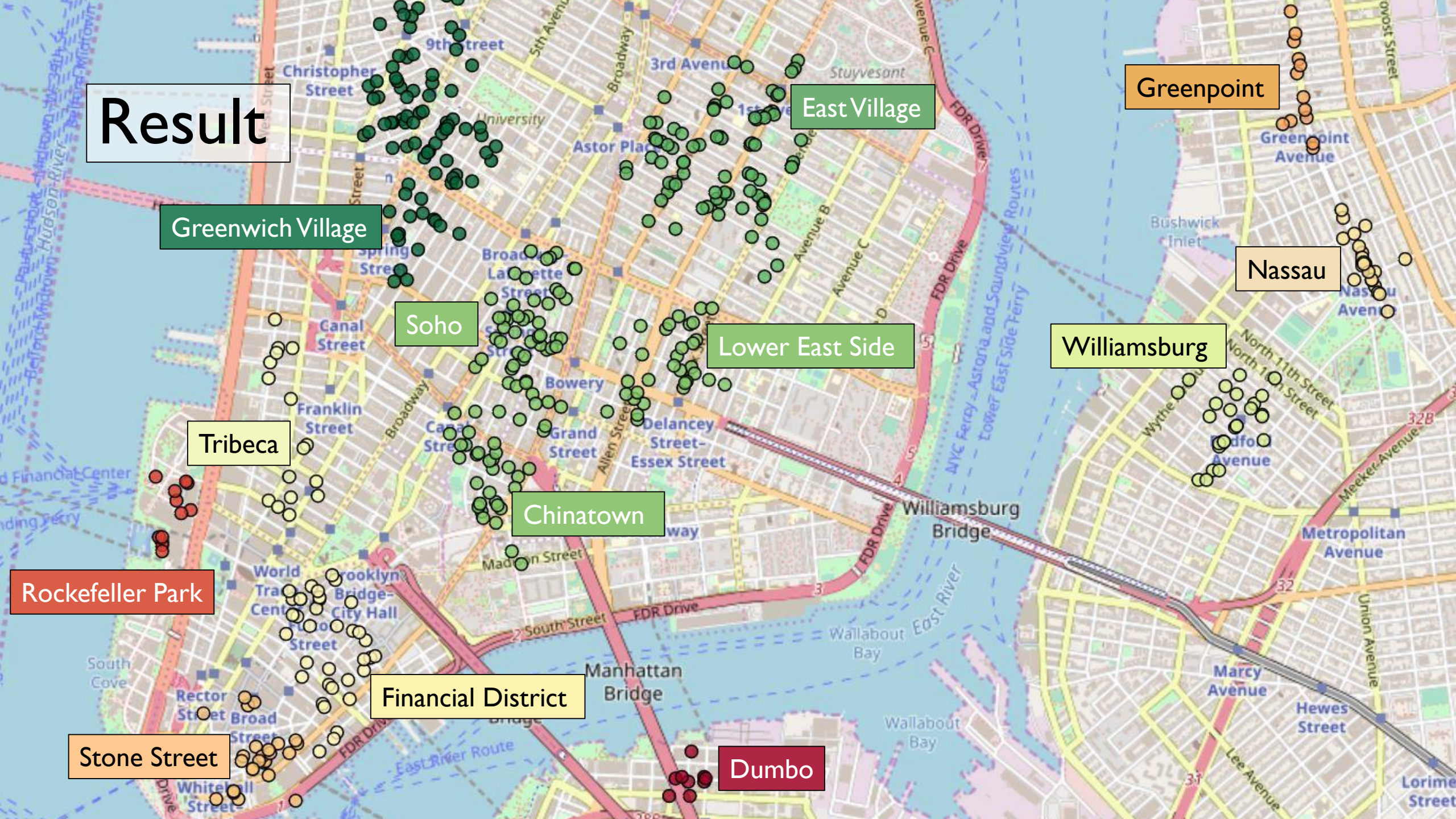
---

- Problem: Limitation to 500 queries per day
- Solution: Sample of 12 clusters for lower Manhattan and subset thereof
- Procedure:
  1. Collect rating data
  2. Group by clusters and take mean
  3. Create cluster rating column in restaurant data frame
  4. Visualize with RdYIGn color map

	Latitude	Longitude	Rating
Clus_Db			
0	40.703855	-74.011015	7.6
1	40.702890	-73.987510	7.0
2	40.708698	-74.007128	7.8
3	40.713981	-74.015500	7.2
4	40.719137	-73.996982	8.3
5	40.717426	-74.009285	7.9
6	40.717707	-73.958365	8.0
7	40.720156	-73.988669	8.3
8	40.730004	-74.001853	8.7
9	40.728620	-73.985708	8.4
10	40.724280	-73.951003	7.7
11	40.731437	-73.954612	7.5



# Result





# Discussion

---

- Value:
  - For tourists, individuals, businesses, investors, guide editors, city planners etc.
  - Easily replicable for other cities and venues
- Caveats:
  - More crowding to certain cluster
  - Adverse effect on restaurants in bad clusters
  - Exceptions not highlighted
  - Cluster parameters matter
  - Example: Chinatown and SoHo clustered together, but the restaurants in SoHo have better ratings (cf. right)
  - Food Query possibly too general (fast food)
- Hypothesis:
  - Size of cluster → better average rating

