Livability scores of select neighborhoods world-wide

Comparing US cities and some similar cities based on their Livability score

Fadeli Mohamed Elwardi

Liveablity-based clustering is important

- You know, figuring out which cities/neighborhoods are easier or harder to live in!
- It's also crucial for predicting whether a city will be leaving its current cluster of similar cities in the near features
  - Obviously, one should leave if the current city will get harder to live in
  - Where to head?
  - Well, to the easy-living city whose probability of leaving its cluster is minimal!
- But this project takes care only of the first step: The clustering of cities.

Data acquisition and cleaning

- A modified version of the scheme shown at **This webpage** is used to estimate Livability Aspects.

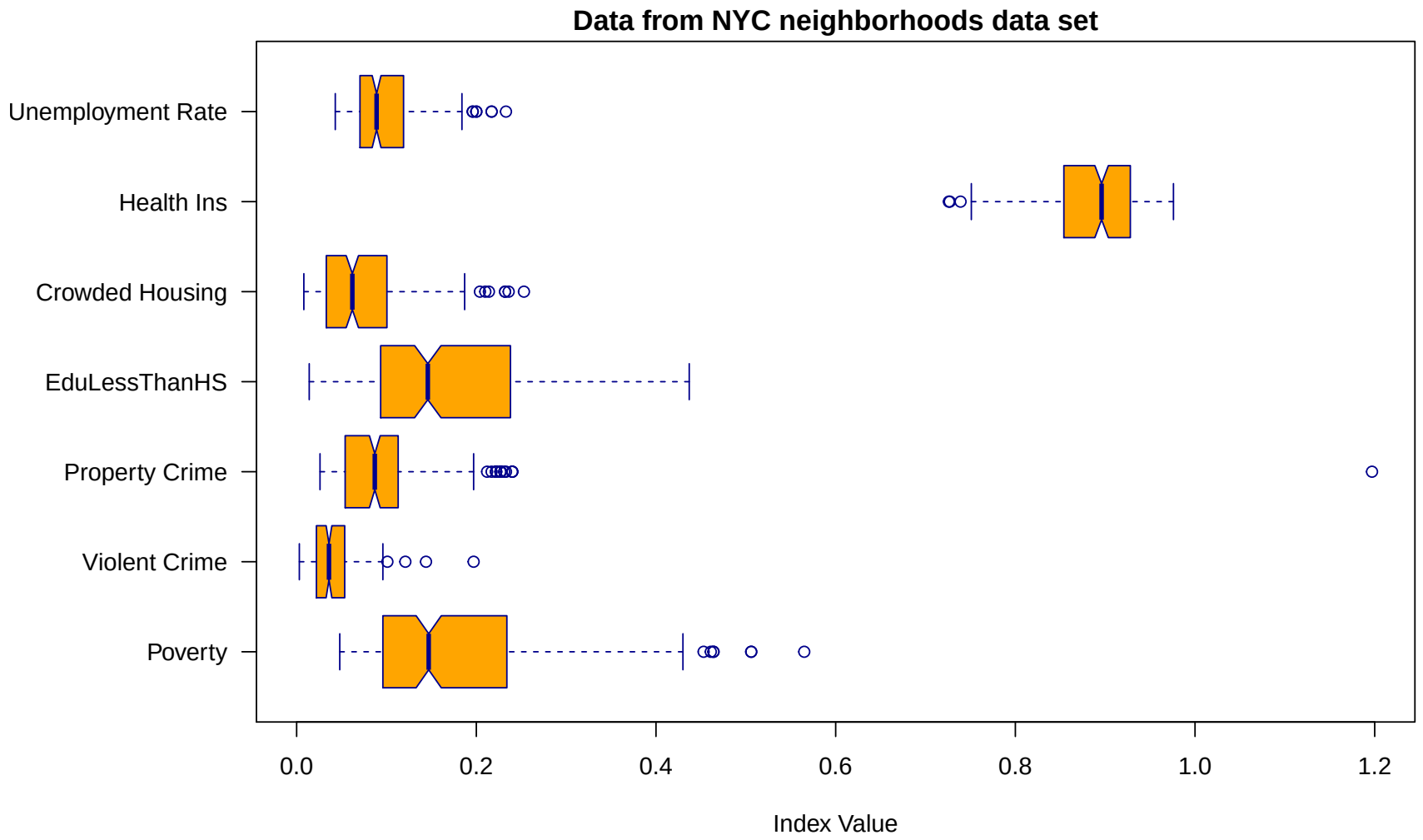
- For this project, we rely on two main sources of data:

1. **A huge data set of NYC neighborhood stats**
2. **Foursquare API**

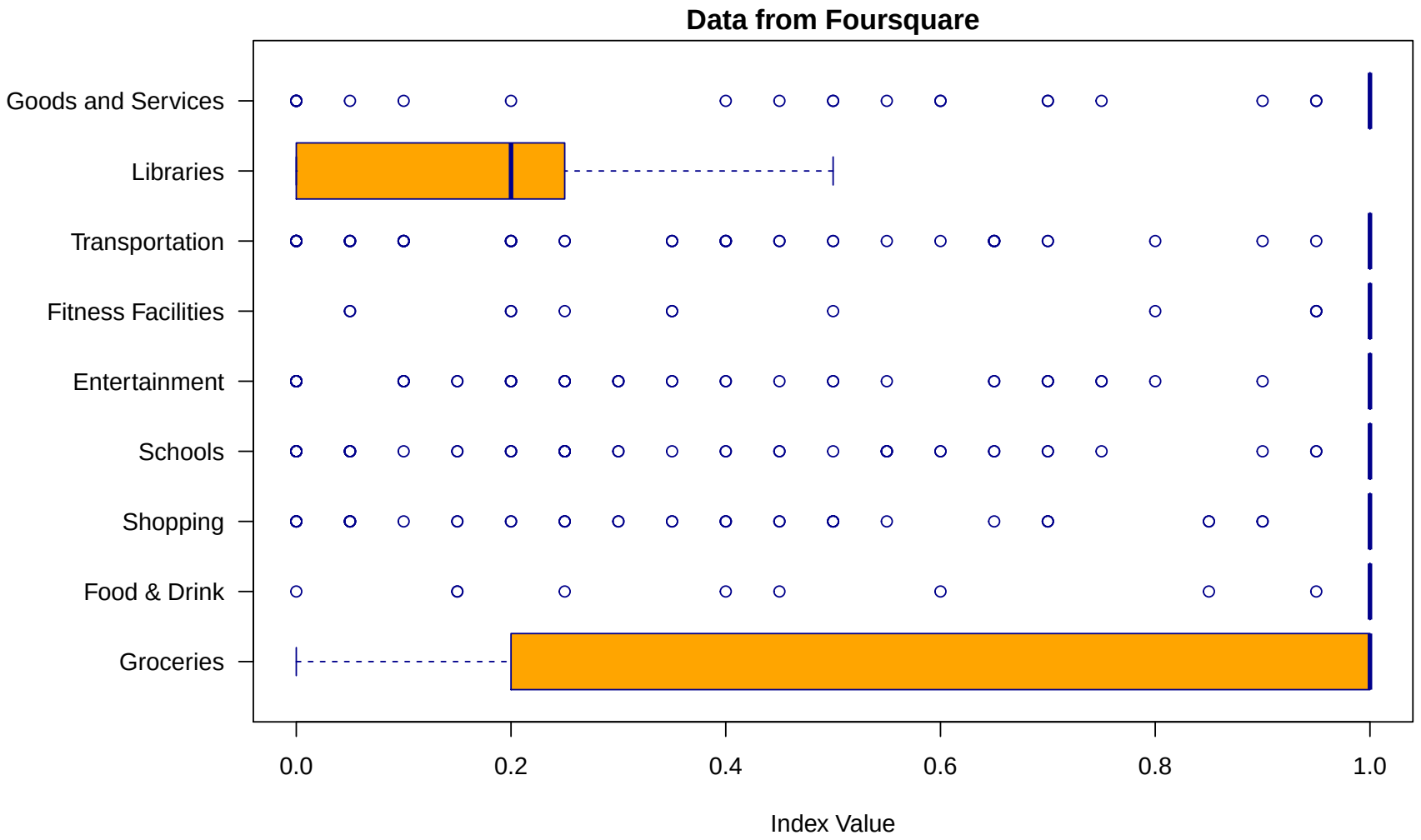
- To get arround 390 rows of data (Including some other cities worldwide that were added manually)

- Data cleanup is performed while getting the Foursquare data
  - No latitude/longitude data for each neighborhood were supplied
  - Instead, search "near" these location in Foursquare and retrieve the coordinates

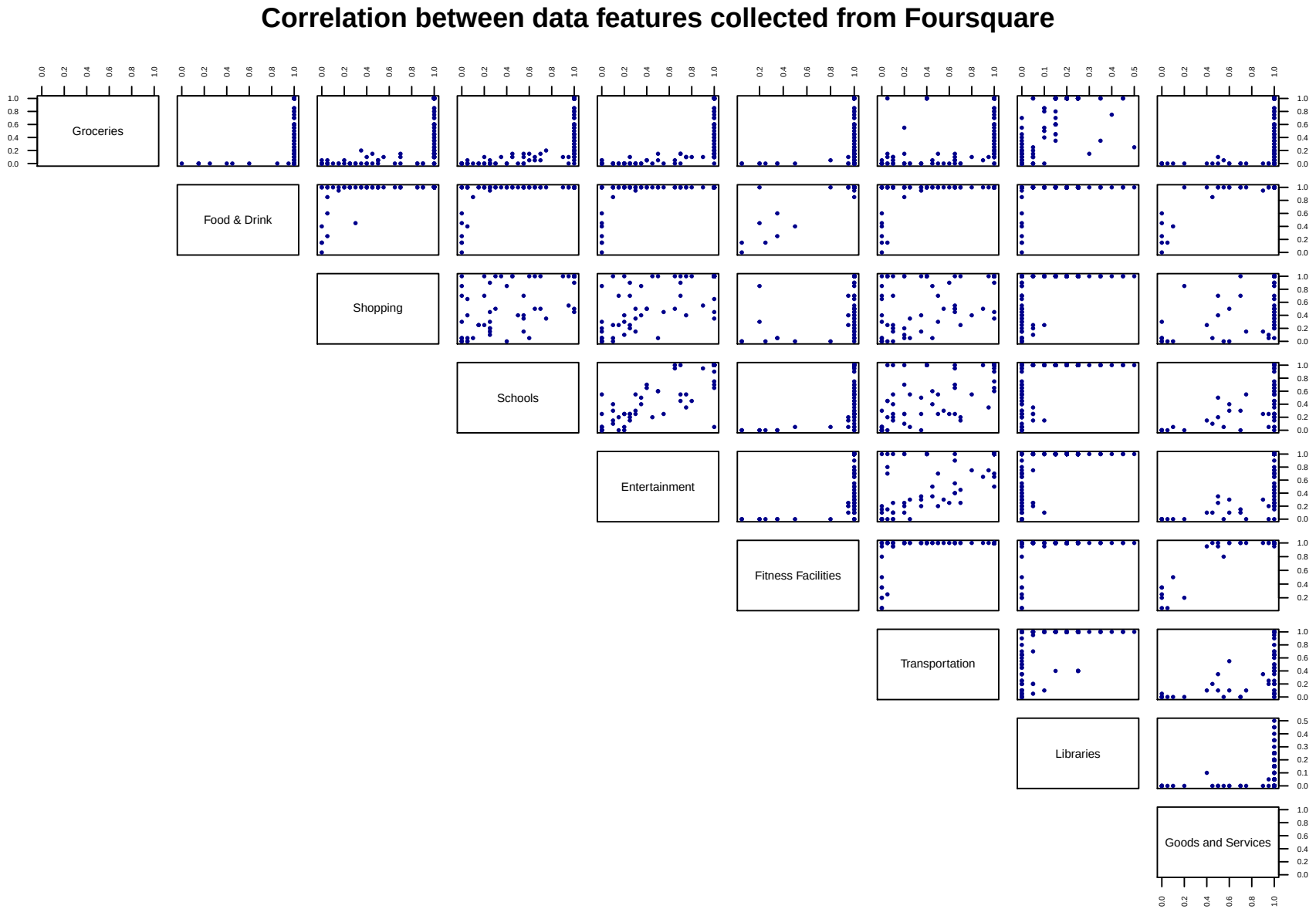
Data collected from NYC neighborhoods Database



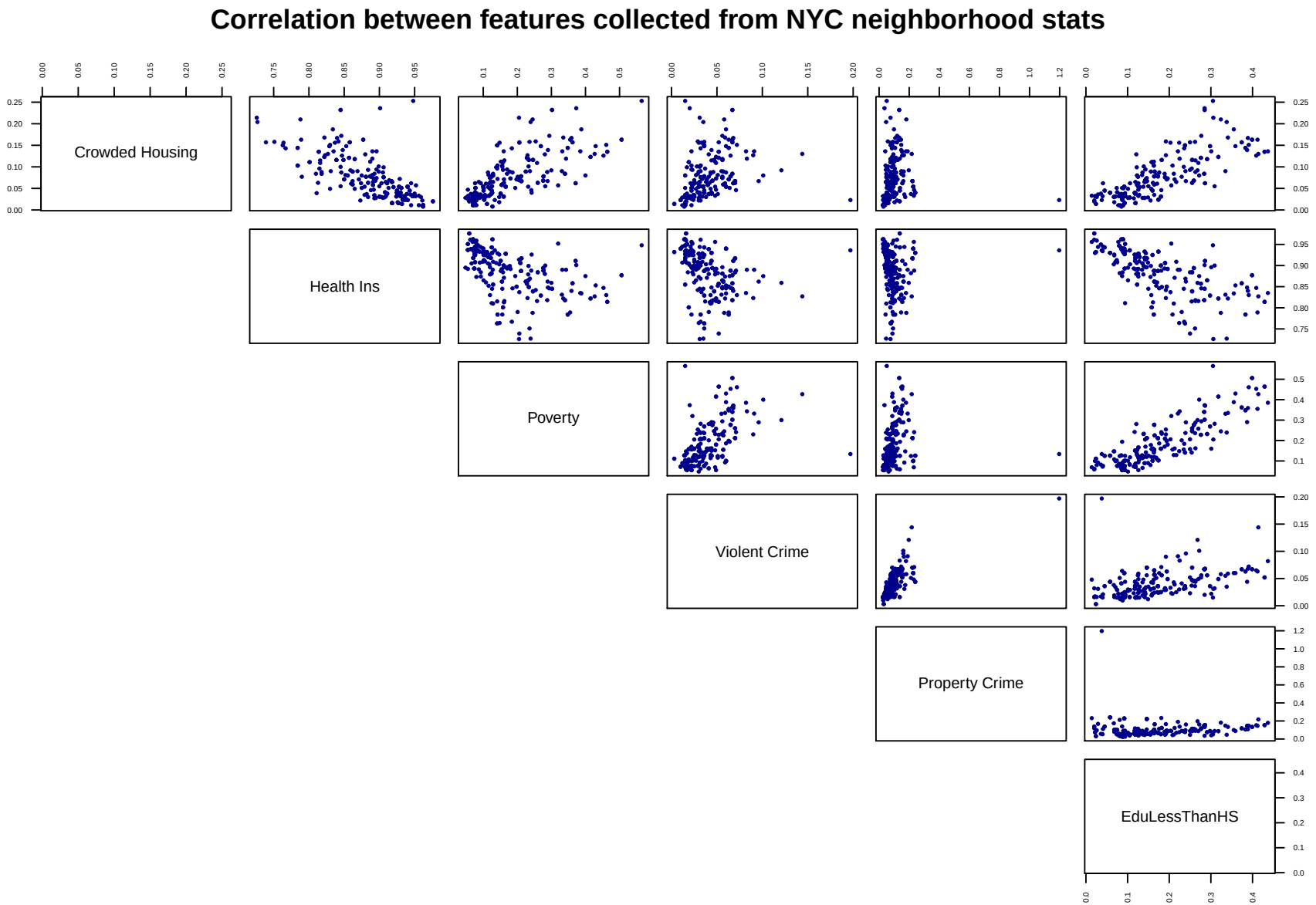
Data collected from Foursquare



How these data features correlate with each other?



How these data features correlate with each other?

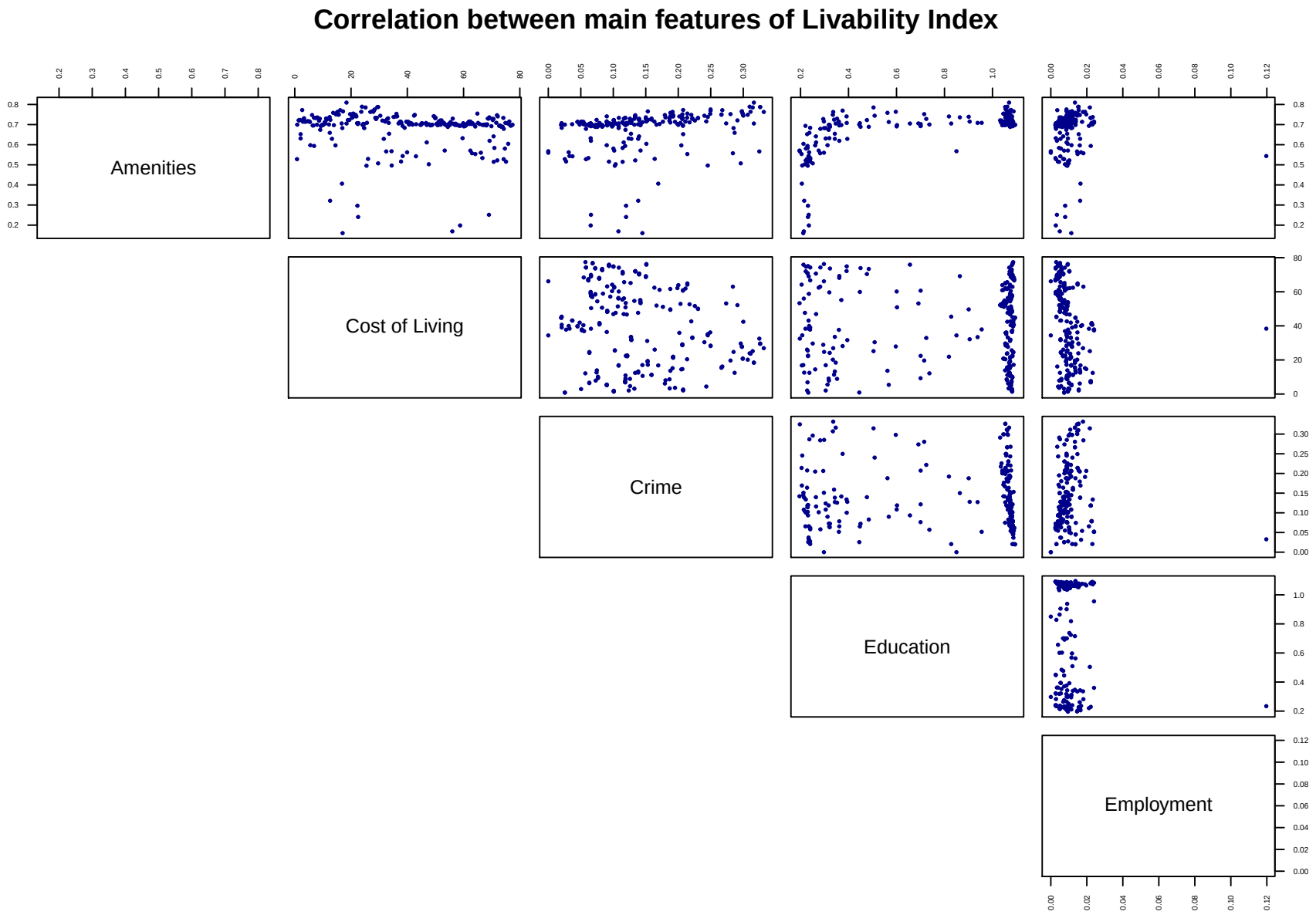


What to estimate from the data?

Each Data feature is multiplied by a contribution coefficient for each Livability Aspect:

	Amenities	Cost of Living	Crime	Education	Employment
Crowded Housing	0.00	0.25	0.00	0.00	0.0
EduLessThanHS	0.00	0.00	0.00	0.85	0.0
Entertainment	0.16	0.00	0.00	0.00	0.0
Fitness Facilities	0.06	0.00	0.00	0.00	0.0
Food & Drink	0.17	0.00	0.00	0.00	0.0
Goods and Services	0.00	0.30	0.00	0.00	0.0
Groceries	0.17	0.10	0.00	0.00	0.0
Health Ins	0.00	0.15	0.00	0.00	0.0
Libraries	0.03	0.00	0.00	0.00	0.0
Poverty	0.00	0.10	0.00	0.00	0.0
Property Crime	0.00	0.00	0.35	0.00	0.0
Schools	0.12	0.00	0.00	0.25	0.0
Shopping	0.26	0.00	0.00	0.00	0.0
Transportation	0.03	0.10	0.00	0.00	0.0
Unemployment Rate	0.00	0.00	0.00	0.00	0.1
Violent Crime	0.00	0.00	0.65	0.00	0.0

How the estimated indices correlate?



Classification - NYC neighborhoods

- NYC neighborhoods look very similar Amenities-wise

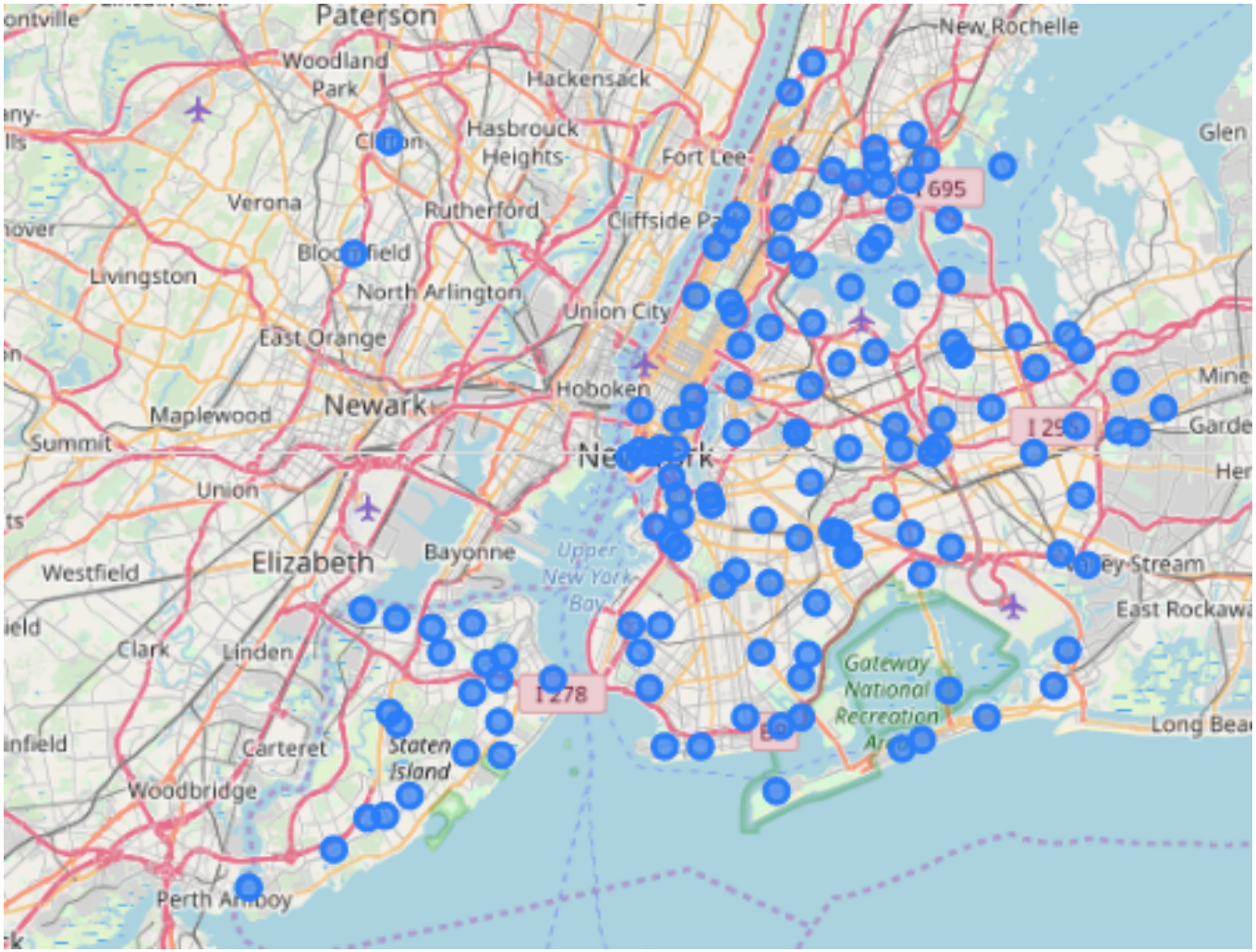


Figure 1: NYC neighborhoods clustering based on Amenities presence

Classification - NYC neighborhoods

- They can also be clustered into 2 groups (All 5 Livability Aspects):

Figure 2: NYC neighborhoods clustering based on Livability Aspects

Classification - Broader view

- The cities around New York are more diverse in terms of Livability similarity:

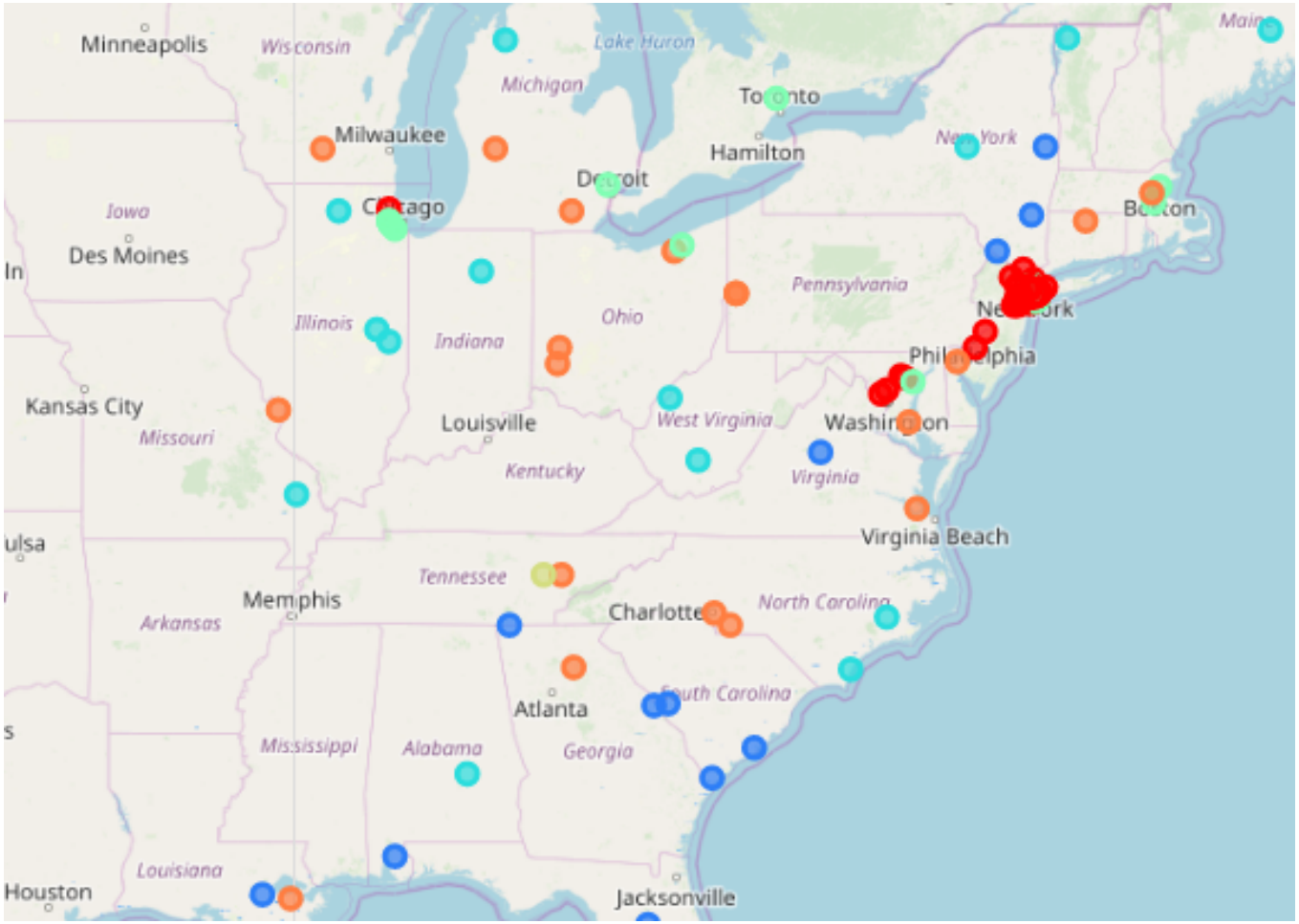


Figure 3: Clustering of some US cities based on Livability Aspects

Conclusion

- Analyzed the livability index of several important locations
- Identified which cities and neighborhoods look most similar when compared by the five main aspects of the Livability score

- Acquiring Premium Foursquare account can greatly improve the model!

- Can now predict when a city may leave its current cluster!
  - By repeatedly applying the model at different (past) time frames