

Clustering U.S. Cities

Using venue distributions as feature vectors.



What exactly gives a city
its character?

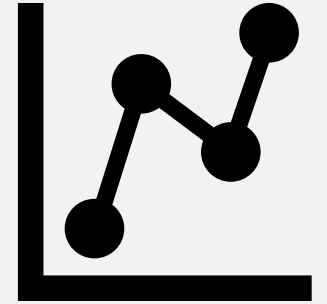
- Population
- Politics
- Health
- Jobs
- Venues*

Why would you care?

- Moving to a new city
- Opening a new business
- Travel industry

What are we trying to
do?

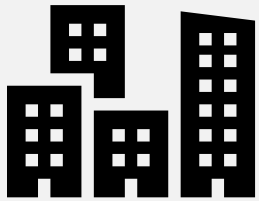
- Unsupervised learning
algorithm to cluster U.S.
cities based primarily on
the distribution of
different venues.



Introduction

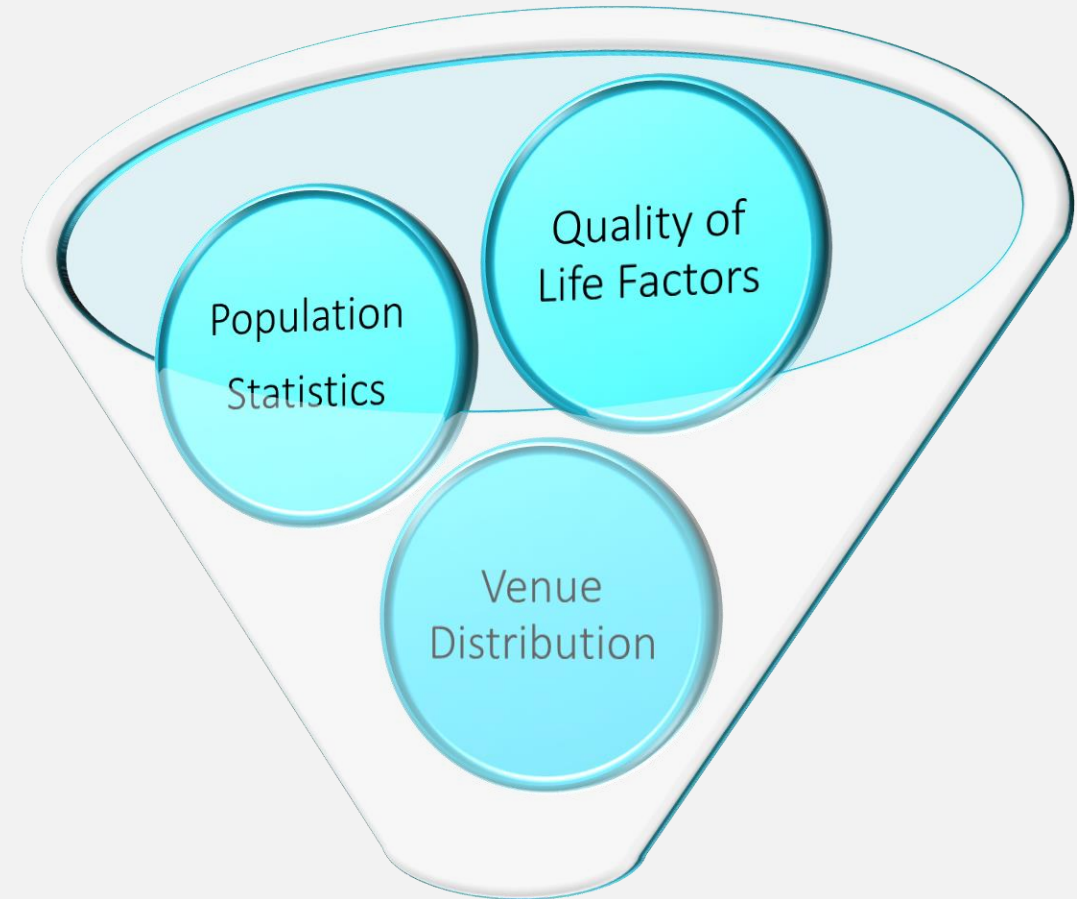
Lorem ipsum dolor sit amet,
consectetur adipiscing elit

How do we define a city's character?

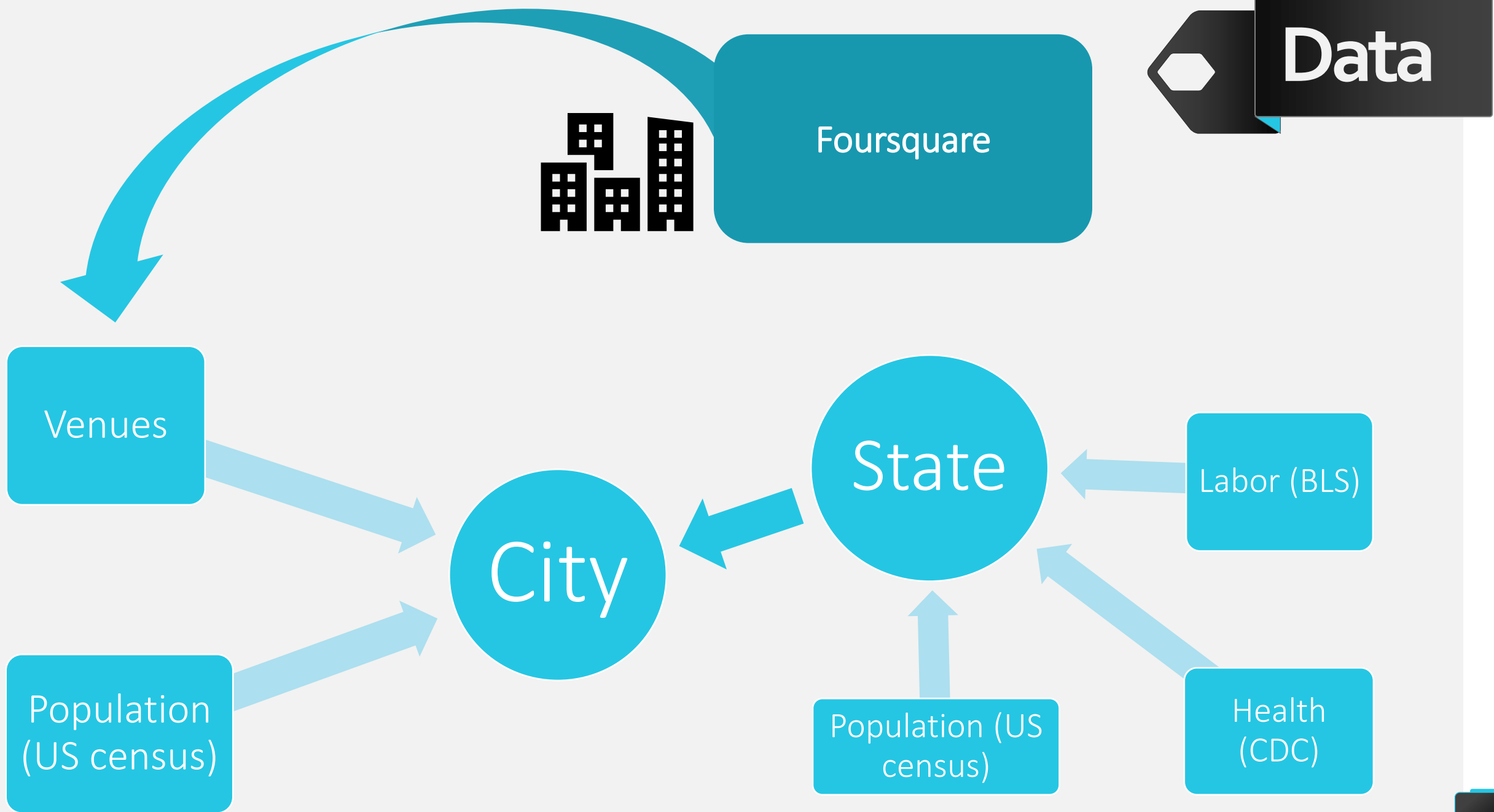


- Size?
- Entertainment?
- Politics?

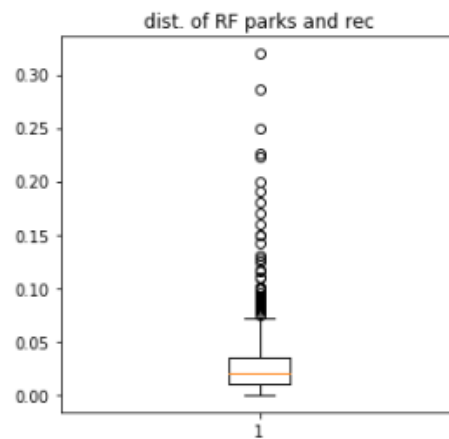
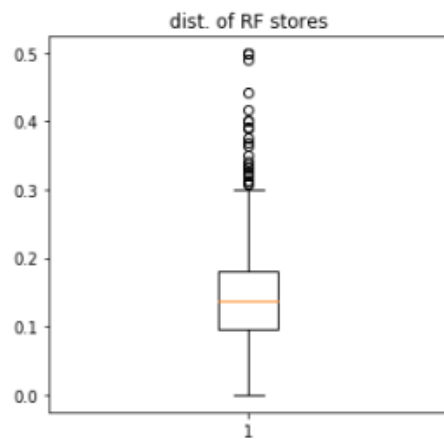
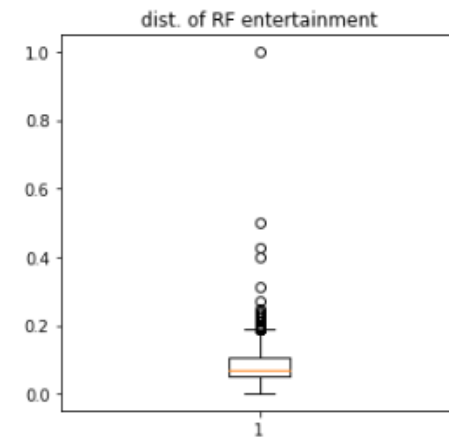
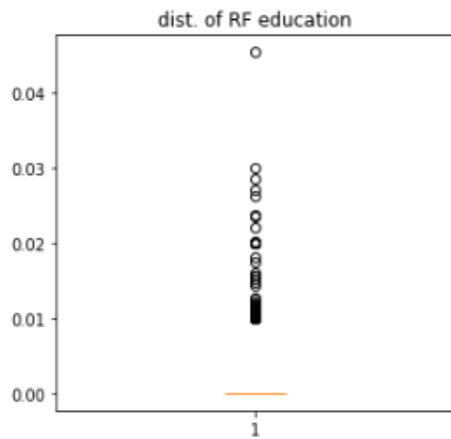
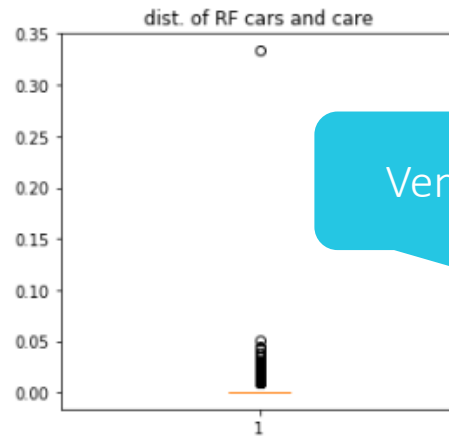
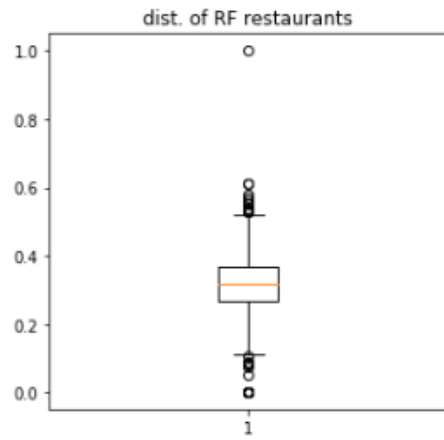
Data



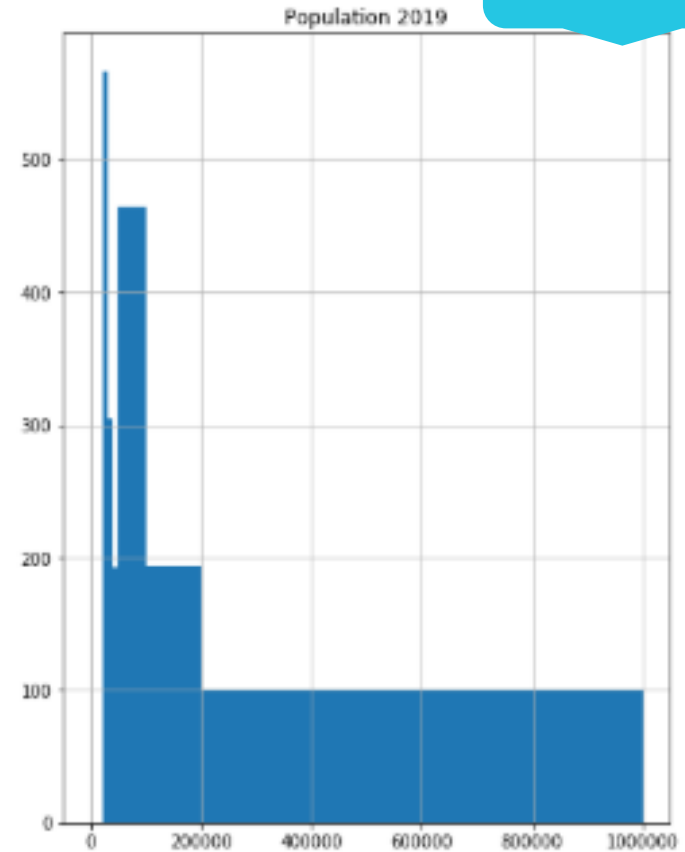
Clusters



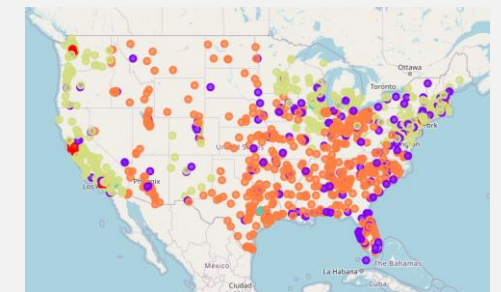
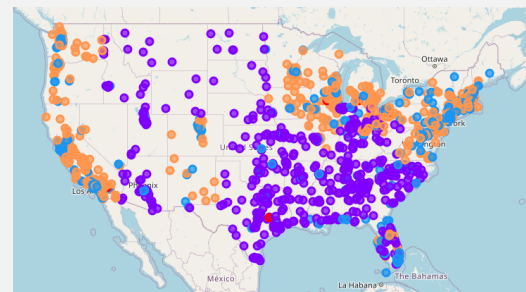
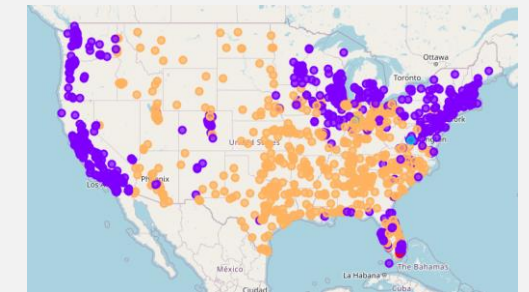
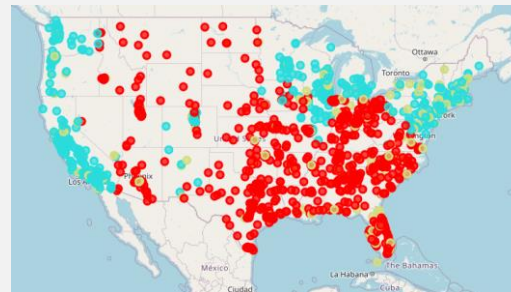
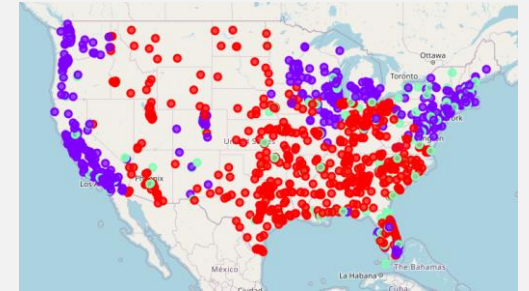
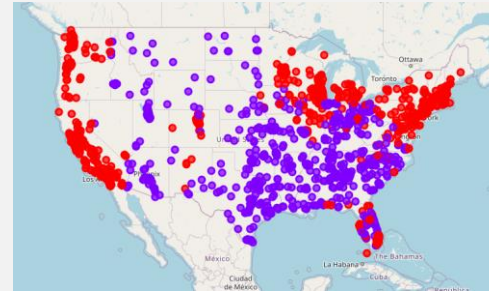
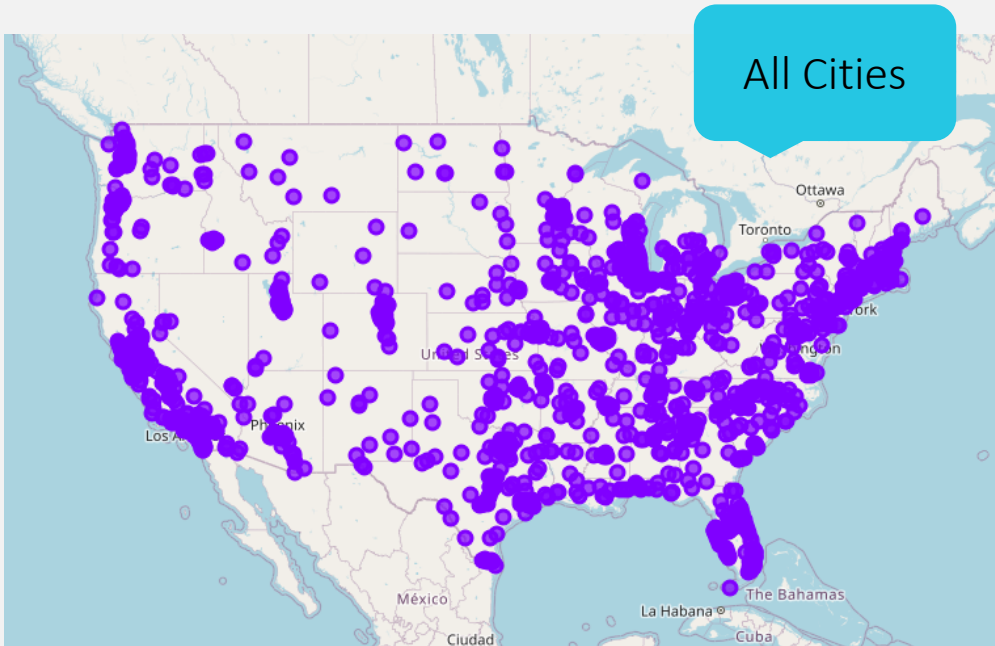
Venue Categories



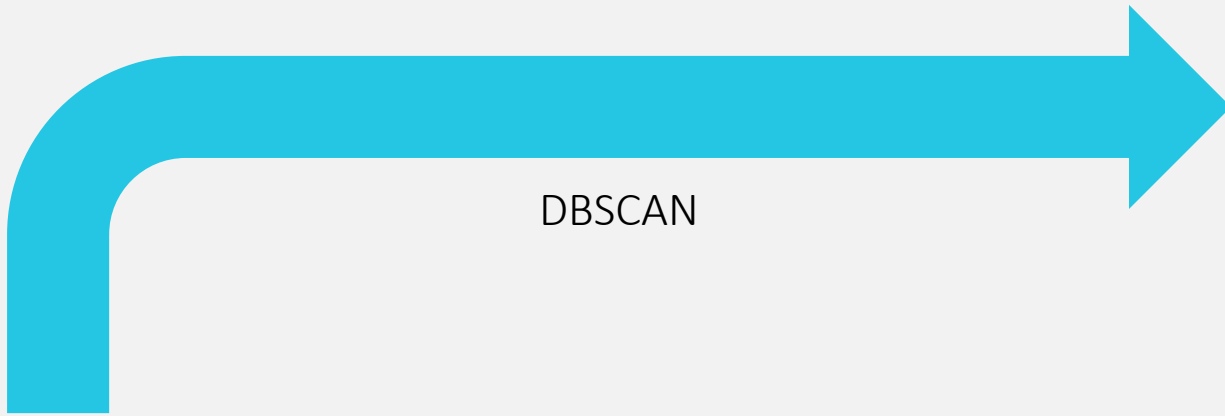
Distribution of Cities



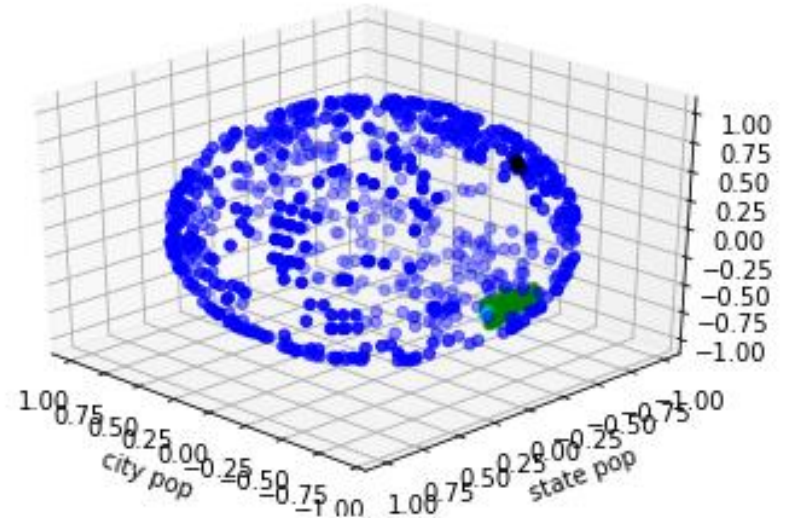
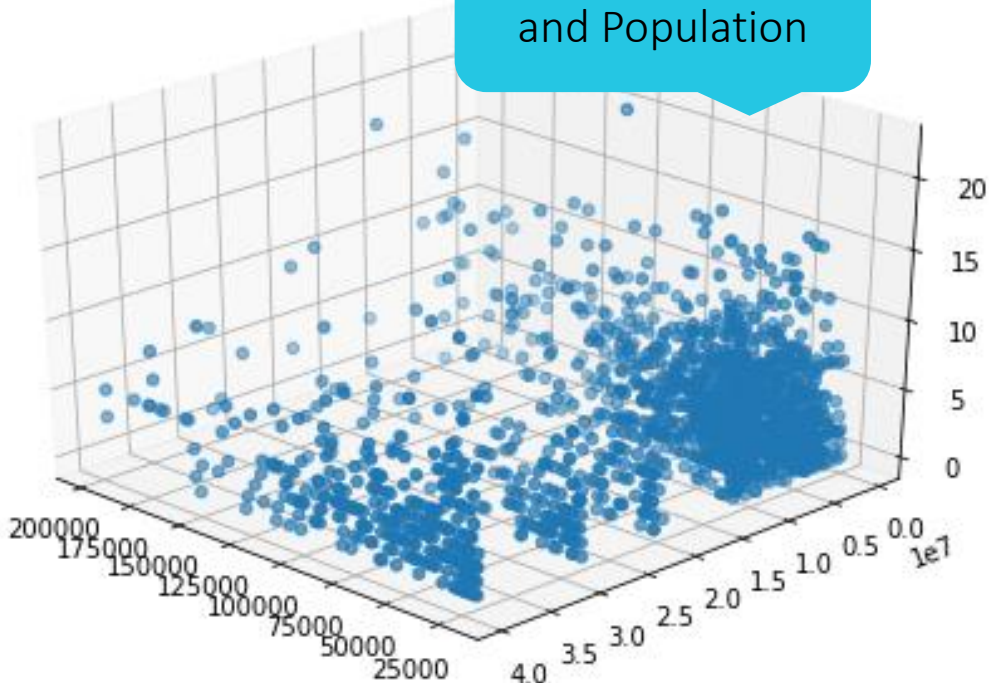
Exploratory Data Analysis



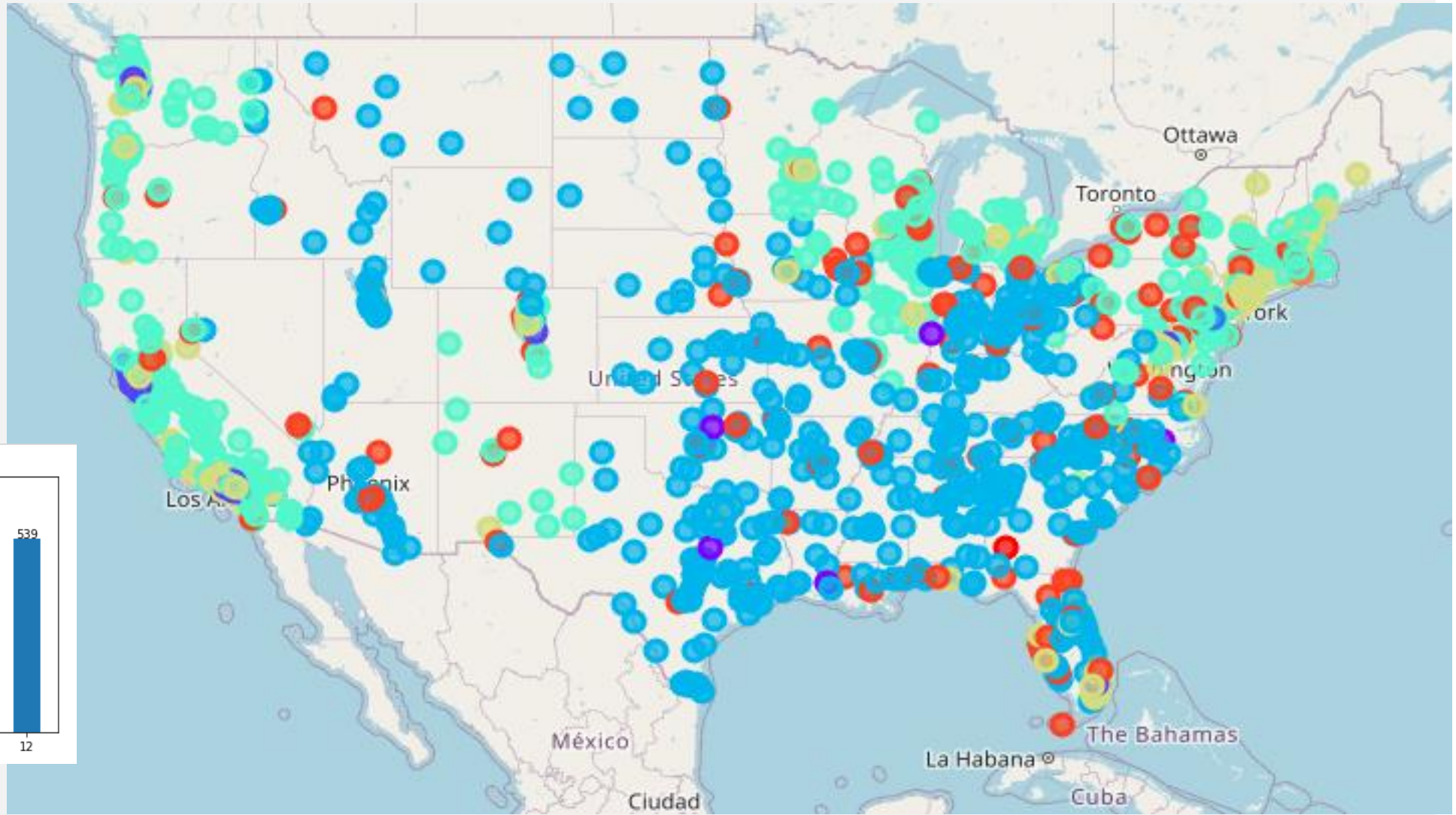
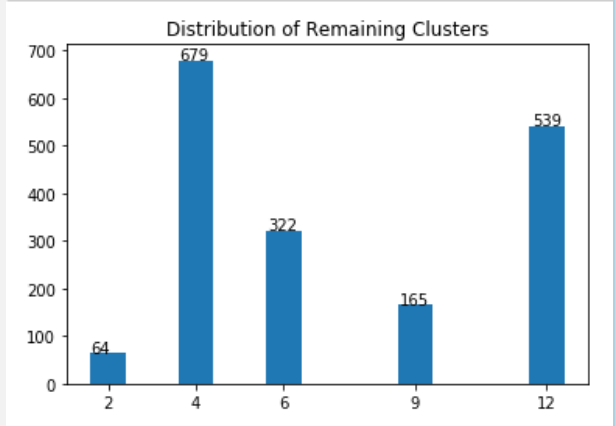
Modelling I



Alcoholic Venues
and Population

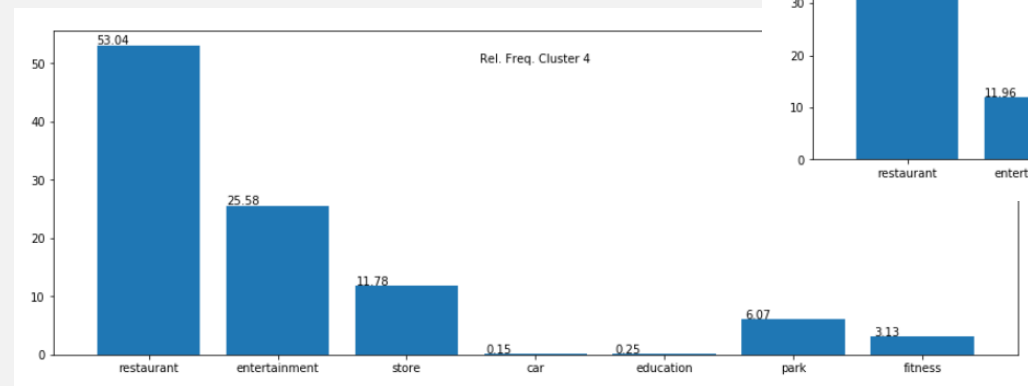
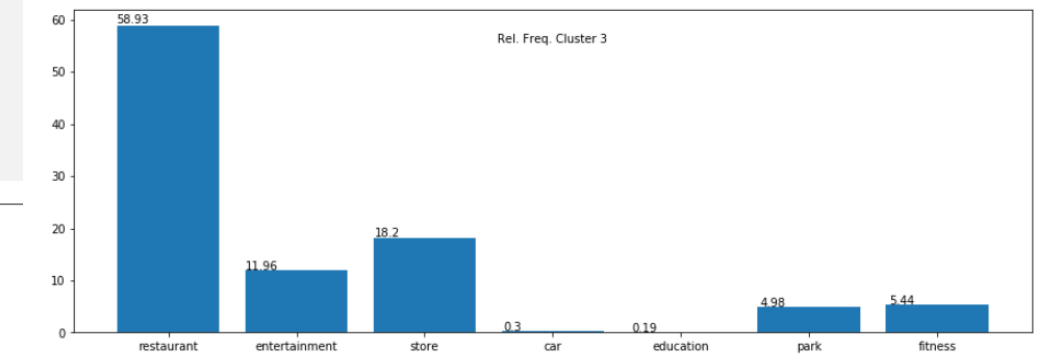
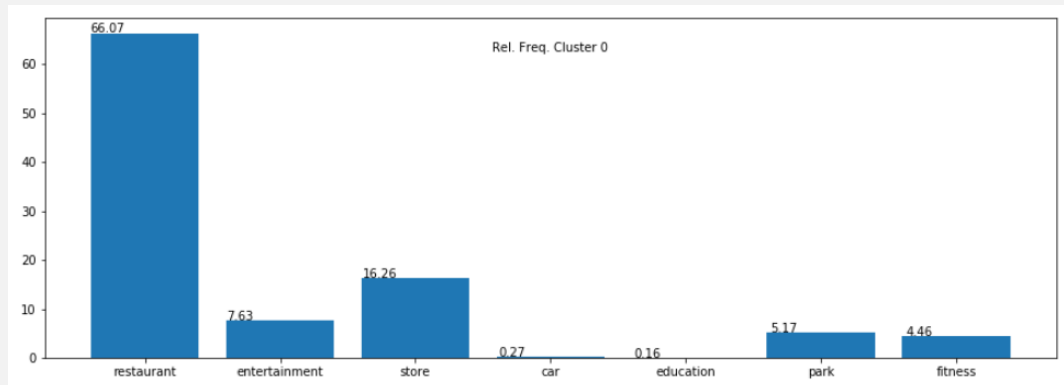
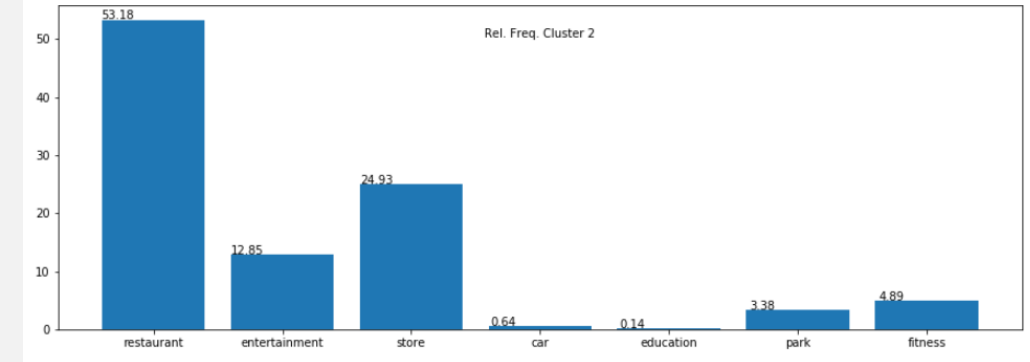
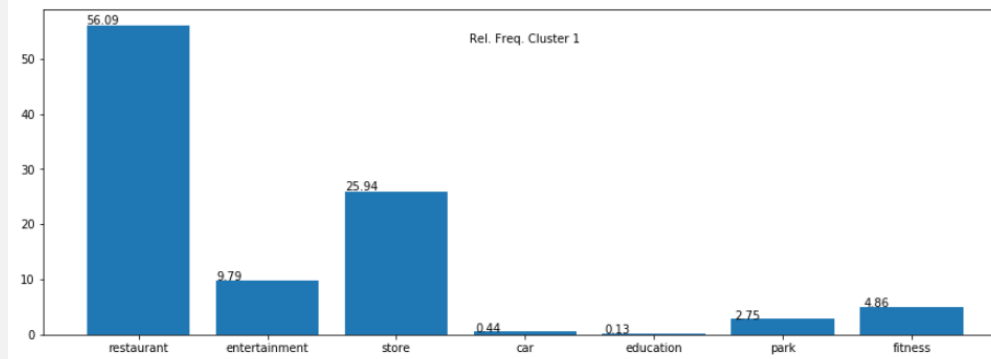


Modelling II



Results

Comparing Clusters



Results

Deeper exploration of what features are most relevant

Beginning to a larger body of work

Clusters Found

Entertainment

Shopping

Restaurant

Conclusion



Thank You

👤 Adamm Hockman

📞 +1 636-751-5946

✉ adamm.hockman@gmail.com