



Comparing Popular US Cities

Applied Data Capstone Project

Simran Kaur Soin



Motivation

- Many start-ups are looking to find locations in New York
- Why?
 - Target population
 - Young adults
 - Venue Popularity
 - Which types of venues do best in NYC
- ... but New York City is *expensive*



Solution?

- Find a city with cheaper property and similar population/venue popularity
- Options under investigation:
 - Boston
 - San Francisco
 - Chicago
 - Philadelphia



Data and Cleaning

- Foursquare API
 - Popular venues from each location
- Governing
 - City population + population density
- Kiplinger
 - Median Home Home Prices
- Wikipedia
 - School District Sizes



K-Means Clustering Conclusions

Number of clusters: 1

Cluster Labels		City
0	0	NYC
1	0	San Francisco
2	0	Chicago
3	0	Philadelphia
4	0	Boston

Number of clusters: 2

Cluster Labels		City
0	1	NYC
1	0	San Francisco
2	0	Chicago
3	0	Philadelphia
4	1	Boston

Number of clusters: 3

Cluster Labels		City
0	0	NYC
1	2	San Francisco
2	1	Chicago
3	0	Philadelphia
4	0	Boston

Number of clusters: 4

Cluster Labels		City
0	1	NYC
1	0	San Francisco
2	3	Chicago
3	2	Philadelphia
4	1	Boston

- Tried 4 different k values
- Boston was most similar, with Philadelphia soon after



Conclusion

- Although Boston was the most similar to NYC, it was also the most expensive with a median home price of \$421K
- Since Philadelphia had a median home price of \$220K, it was deemed the best choice for start-ups