Comparing Similarity of Cities

Samuel Chu

March 29, 2020

1. Introduction

1.1 Background
Millions of people move to different cities every year for different reasons. Some people need to move because of their job, however some people seek a change in scenery.

1.2 Problem
My project intends to help people pick a city that is either different or similar to their prior living situation.

1.3 Interest
Anyone who is moving to a new city or looking to explore what a city has to offer.

2. Data acquisition and cleaning

2.1 Overview
Used **FourSquare API** to gather information about a city's amount of entertainment, shopping, and food.

3. Methodology

3.1 Type of ML
Used K Means Unsupervised Learning Algorithm with the SkLearn, Pandas, and Numpy Packages.

3.2 Data Analysis
Based on stats for entertainment, shopping, and food it clusters the cities you want to compare to each other. If a user wants to explore LA and lives in Chicago right now, then they could compare LA and Chicago to see if they cluster together or not.

3.3 Results

Even though this can be run on any amount of cities in the world, I have found from the cities that I have tested that big cities have a lot in common. They often all have lots of entertainment and food.

3.4 Discussion

I found it interesting though that not all of the big cities had shopping. For example, there was a lot more shopping in LA than in NY. So if you like to shop LA definitely has more options for you.

4. Conclusion

Over the course of these 9 courses I have developed a deep excitement to learn more about Data Science and Machine Learning. In this project I finally got to work on a project on my own! There were a couple of times where annoying errors would not go away, but that is where StackOverFlow comes in. Overall, I am pleased with the results I have produced and data I can show about different cities all across the world with the help of the FourSquare API.