

Coursera Capstone Project Presentation

Finding Similar City Analysis

Isaac Trussell

September 21, 2019

Problem Introduction

I currently live in Boise, ID. I have recently decided that I would like to make a change and move to another part of the US for both career and personal reasons. I don't really have a destination and am pretty open to a lot of different options. The thought of leaving Boise and starting brand new at a brand new place really scares me and seems like a big leap especially since I don't really have any problems with Boise itself other than that the job market for my current career field of mechanical engineering isn't exactly booming. This lead me perfectly to my idea for this project. I should cluster cities that have similar attributes to that of Boise. So basically my goal is to find a city that is super similar to mine and I can look for jobs in that area and it won't be that big of a change for me.

Objective

Find cities that have similar qualities to Boise, ID

Project Methodology

The basic structure of the project should look as follows:

- Find data about population and location
- Find data about temperature and climate
- Gather information using the Foursquare API about popular venues in the area
- Cluster the cities based on similarities

Future steps for this analysis will include:

- Find job posting information for mechanical engineering jobs in the area
- Cost of living comparision

Data Sources

After a quick search for city data, it was pretty easy to find general information about different cities online. The first table that I came across was a wikipedia article that had the following information that I plan to use in my analysis:

- City
- 2018 Population Estimate
- 2010 Census Population
- 2016 Land Area
- 2016 Population Density
- Latitude & Longitude

My next data source was found here. It has some interesting climate data for lots of different cities. Unfortunately it is smaller than that of the wikipedia data, so we will lose some cities that we can compare with Boise. The data I will be using from this website are the following:

- Average Winter Temp
- Average Spring Temp
- Average Summer Temp
- Average Fall Temp
- Average Yearly Precipitation
- Average Yearly Snowfall

Lastly, I plan to use the Foursquare API to pull popular venues in the area. I am going to categorize the venues based on their parent "category" as alot of the categories go into great depth. For example, I want to try and add up all the venues that are in the area that are simply resturants, rather than having 1 american food resturant and 1 chinese food resturant. For this high level analysis I don't think I need that in depth of information quite yet.