

Applied Data Science Capstone Week4 Assignment

Peer-Graded Assignment: Capstone Project – The Battle of Neighborhoods Week 1

Introducing the Opportunity:

Upon retirement, I'm entertaining the opportunity of opening a small coffee / music house in one of two college cities where I lived while a student, giving back to the origin communities. Those cities are Harrisonburg VA and Boulder CO. Now, there are significant differences between the two, beginning with geographically and weather patterns that for this exercise will be excluded. However, there are some commonalities which we will use to generalize through data analysis as to where is a better, safer opportunity upon retirement. For example, we'll use data sources to help investigate the following questions:

- What areas within both cities contain the highest levels of crime?
- Does population density correlate to criminal behaviors?
- Using venue data from Foursquare, what venues and locations are most common in both cities?
- Where is there a greater, and more prosperous opportunity to establish a college level coffee shop / music house?

Harrisonburg VA Tourism Image:



Boulder CO Tourism image:



Data Sources:

To understand and explore the opportunity with greater thoroughness, we will need the following city open data and other access points, as indicated below:

- Boulder CO Open Data Site: <https://bouldercolorado.gov/open-data>
- Boulder CO Neighborhood Data Site: <https://www.bouldercoloradousa.com/about-boulder/boulder-neighborhoods/>
- Harrisonburg VA Open Data Site: <https://www.harrisonburgva.gov/GIS>
- Harrisonburg VA Neighborhood Data Site: <https://www.harrisonburghomes.com/neighborhoods.php>
- Foursquare Venue Data Site: <https://foursquare.com/>

Next Steps:

Using this data will allow exploration and examination to better address the initial questions posed in the introduction. The city data sources will enable us to properly group crime by neighborhood. The census data will enable us to compare population density to examine if areas of highest crime also associate with population density. Locations of interest will then allow us to cluster and quantitatively understand the venues most common to those locations, providing a better understanding as to the opportunity that may exist, and finally summarizing and making comparison between the two cities will lend us to determine a general need and potential sustainability for the business.