

# Capstone Project - The Battle of Neighborhoods

A Study of Neighborhoods in Overland Park

Huiping Qin

May 11, 2019

## 1. Introduction

Why Overland Park shows up frequently on the top list of best places to live in US? There are various info on the Internet, however deep neighborhoods analysis is lacking. Through data analysis and machine learning technology, we may find evidences to the answer, which may in turn help local business on recruiting marketing and attract more talents to Overland Park. It may also help people who want to find a better place to retire or to raise a family to make informed decision. I have lived in Overland Park for over 20 years and searching an answer on the question may help me to know my community better and to love it even more. I love my local community, with easy access to good education, libraries, museums, shopping, entertainment, park and trails, and so forth. But what are other neighborhoods in Overland Park like, what venues do they have to contribute to the listing of Overland Park on top places to live in the US?

Equipped with knowledge and skills just acquired from the Coursera Data Sciences classes, I hope to find the answer by using a combination of location data and machine learning to explore neighborhoods in the city of Overland Park. In this study, I am going to use the Foursquare API to explore neighborhoods in Overland Park. I will use the **explore** function to get the most common venue categories in each neighborhood, and then use this feature to group the neighborhoods into clusters. I will use the *k*-means clustering algorithm to complete this task. Finally, I will use the Folium library to visualize the neighborhoods in Overland Park and their emerging clusters.