**Selecting Optimal Location for Pizzeria Restaurant**

**in the Sacramento Region**

by Michael K. Christoff

January 8, 2020

**1. Introduction**

My clients were moving from the east coast to Sacramento, CA in order to be closer to their extended family. They currently ran a successful pizzeria in New Jersey which they are planning to sell, and with the proceeds open a new pizzeria in the Sacramento region. They asked me to help determine, through data science research, which regions would be an ideal spot for their new restaurant.

The cost of retail space was not an important consideration for my clients, but they insisted the new location be a high traffic, growing region that is near event venues, corporate centers and recreational attractions. They also wanted to minimize competition by avoiding areas already inundated with pizzerias.

They made it clear that they really love downtown Sacramento, but they were aware the region surrounding Sacramento had many great opportunities. As long as they are just a few hours away from Sacramento to be able to visit family regularly, they want to explore every opportunity before making their decision on the best location.

1. **Data**

I knew they prized the downtown area, so the first part of my research would focus on the city of Sacramento and neighborhoods. The second part of my research would focus on the cities within a two hour drive of Sacramento.

**2.1 Data for city of Sacramento search**

For my search within the Sacramento city limits, I will use a tourism website planetware, which highlighted top-rated attractions within the city of Sacramento. Based on each attraction, I will explore their immediate surrounding area to identify popular venues and then make comparisons. For this, I will use the Foursquare website. Foursquare, the world’s leader in location-based data, uses crowdsourcing to collect specific information about locations and venues all over the world. I will use Nominatim geolocator location data to covert location name or address to map markers onto a Folium map. I will use K-means clustering to determine cluster labels to add to the Folium map and my results.

**2.2 Data for cities outside Sacramento search**

For the extensive region surrounding Sacramento, I will concentrate on the largest populated cities and those with the highest median family income. I will use U.S. Census Bureau to determine the largest cities in the Sacramento region, as well as specific population and median family income for these cities. I will explore cities surrounding Sacramento based on venue information from Foursquare. I will use Nominatim geolocator location data to map markers onto a Folium map. I will use K-means clustering to determine cluster labels to add to the Folium map and my results.