Neighborhood Venues Analysis for Café & Bar Crawls

IBM Data Science Capstone Project

Anthony Canterbury

September 29th, 2019

1 Introduction

Within the city of Cincinnati, OH a nonprofit organization wants to put on a series of outreach events which take donating participants to multiple neighborhood cafés and bars to socialize. The goal of this "crawl" is not only to take donations but to make the charity or issue as visible as possible in these neighborhoods.

To make these "night & day crawls" successful the venues in these locations must meet several criteria:

- First the location must have at least two cafés and two bars in proximity. Seven venues max per cluster. Can seven be considered within walking distance (less than 0.7 miles apart)?
- The venues must be open sometime on Saturdays between 3pm and 10pm.
- The venues must be popular (well-reviewed).
- The venues must not be expensive.

The problem to answer is which clusters of venues in these neighborhoods are most ideal to recommend for partnering with for these events based on available data. Also, which café to meet at and which route is optimal to travel will be answered within the analysis.

2 Data

The data required for this analysis:

- List of Neighborhoods within Cincinnati, OH. A list of the 50 neighborhoods with approximate geographical boundaries comes from https://data-cagisportal.opendata.arcgis.com/datasets/cincinnati-sna-boundary
- List of neighborhood venues with reviews and approximate prices. This data will be supplied by the Foursquare API.

While bars are easy to filter for, a café is a little more subjective. For this analysis we will consider anything labeled a café, coffee shop, teahouse, or dessert shop (e.g. donut shop, ice cream shop, pastry shop) a café. Beyond basic cleaning and formatting for all datasets some of the data from the

Foursquare API will most certainly need to be predicted due to missing data for prices and review ratings. Missing data is expected since new venues appear regularly in many of the neighborhoods.