Background

Wake County is located in central North Carolina. Its largest city is Raleigh, the state capital. Along with eleven other municipalities that surround Raleigh, Wake County has a population of 1.1M, making it the most populated county in North Carolina. It is also home to the largest school district in the state.

Within a half-day's drive from both the mountains and the ocean, and midway along the Atlantic seaboard, Wake County's location is ideal to access many points of interest. Ranking 3rd for the Best Counties to Live in America, 14th for Best Counties for Families in America, and combined with low unemployment, Wake County has experienced phenomenal population growth.

Recently, a friend has decided to move to Wake County and start a pizzeria. In order to help, an analysis will be created to pinpoint the top-5 locations in the county to open the shop.

Problem Statement

In order to find the best potential locations, one of the factors will be to find dense population centers which lack pizzerias. This approach should maximize clientele while minimizing competition.

This analysis will uncover a metric to calculate and locate where the top-5 potential locations should be, through analytics of data from multiple sources.

With that said, other factors should come into play, too, in order to find the best potential locations. However, these are out of scope for this analysis but are discussed briefly later in the document.

The ideal audience for this analysis would be any entrepreneur looking to open a pizzeria in Wake County. The approach, though, could then be easily modified to accommodate any cuisine or location.

Data Sources

Two data sources will be used.

1) Wake County Open Data (http://data-wake.opendata.arcgis.com/)

   This resource contains a wealth of data about Wake County, including properties, parks, education, public safety, and zoning data. Of interest, it contains demographic information. Specifically, one table includes 2010 census data by census tract. This table includes the area of the tract and population. This is ideal to calculate segmented population density throughout the county.
   (http://data-wake.opendata.arcgis.com/datasets/census-tracts-2010)

2) Foursquare (https://developer.foursquare.com/)

A comprehensive resource of crowd-sourced data, Foursquare will be leveraged to find all the pizzeria locations in Wake County via its developer portal.

Approach

a) Using the **Wake County Census Tract Data**, obtain population and area (km^2), along with longitude and latitude of each tract. This is ideal for calculating relative population density throughout the county.

   Census Tract is an excellent a proxy for neighborhood data, especially considering customers will be coming various neighborhoods. Wake County does not have boroughs.

b) Using **Foursquare**, pull all the restaurants in Wake County that can be keyed on Italian cuisine or pizza, along with its location.

   This list will be further refined, as the entrepreneur believes his pizza product and its unique recipe stands alone. Therefore, he does NOT consider national chains to be direct competitors. The Foursquare data will be scrubbed to exclude the following: Papa John's, Domino's, Little Caesar's, and Pizza Hut.

c) Merge the two data sets on longitude and latitude. For the location of each Census Tract, count the number of pizzerias and Italian restaurants within a 5-kilometer radius. This is roughly the extent of current delivery radii and translates into a 10-minute drive.

d) Calculate the population per square kilometer and divide it by the number of competitor locations. This provides the maximum number of potential customers within close driving proximity to the location, relative to competing restaurants.

e) Sort this metric in descending order and document the location of the top-5 census tracts. A higher metric translates into higher potential customer-base with minimal competition.