**Finding the Best Pizza in Town**

Kevin Miller

October 9, 2019

1. **Introduction**
   1. **Background**

I started a new job in Rochester, MN at the beginning of this year and recently moved to the area. My wife and I love to try new restaurants and foods, but there’s one food that never seems to get old for us, and many others for that matter: Pizza. In every area I’ve ever lived people seem to be on a search for the best pizza in town, however everyone has a different opinion about what pizza they like best. Since I’m living in a new area and haven’t tried many pizza places here yet, I thought it would be interesting to see if I could use the skills I learned in this course to determine objectively what the best pizza in Rochester, MN is.

* 1. **Problem**

The fact that everyone’s opinion differs from one another makes finding the ‘best’ pizza a challenge. In fact, this makes finding the ‘best’ anything a challenge. In reality, there isn’t an objectively ‘best’ anything as someone is bound to have a different opinion, so the challenge is then to find out what most people prefer. This information can easily be used for marketing purposes, and certainly recommendation services like apps or internet search engines. There are a myriad of ways one could go about solving this problem, but I’m choosing to use user ratings and number of likes for pizza restaurants from the Foursquare API to try and determine where the coveted ‘best pizza in town’ in Rochester, MN.

1. **Data**

**2.1 Data Source**

For this project I chose to use data from the Foursquare API exclusively. The Foursquare API has a multitude of location related data, but I’ll be primarily using the following:

* Latitude and Longitude data for Rochester, MN and venues
* Venue Name
* Venue Category
* Venue Rating
* Venue Likes

Given the lack of available daily pulls allowed for my Foursquare Developer account, I’m limiting my analysis to include the data listed above. If I had more time and a better account tier, I’d consider looking into the number of dislikes and possibly tips provided by users to paint a better picture of what users think of the different pizza places. I plan to use the venue rating as my first determinant of the best pizza and town, and number of likes as a tiebreaker if necessary.

**2.2 Data Preparation & Cleaning**

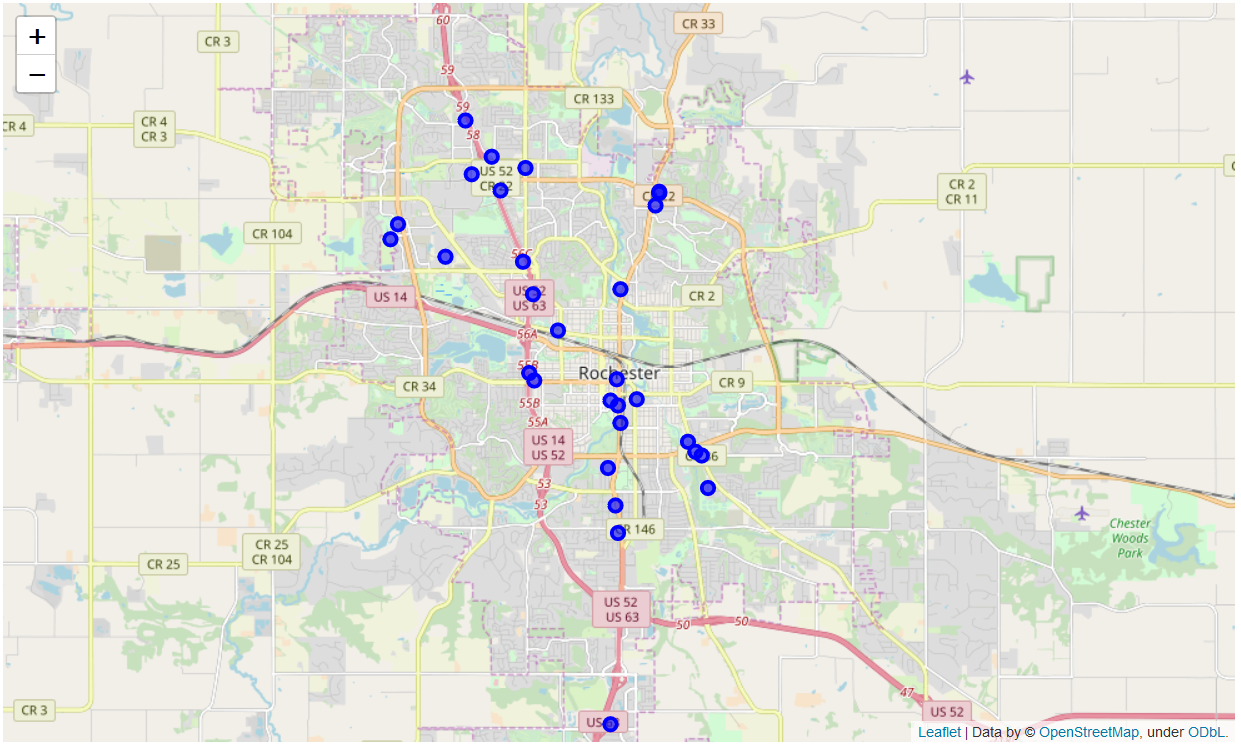
When working through this project, I had the Sandbox Tier of the Foursquare Developer account, which allowed me to make a limited amount of calls to the Foursquare API per day. I simplified my analysis for this reason, however there were still a few steps to retrieve the data I needed and manipulate it into the format necessary for analysis.

I started by using the geopy library for Python to pull the latitude and longitude values for Rochester, MN since this was my starting point from which to find pizza places. I then made a call to the Foursquare API to search for venues with ‘Pizz’ in the name within Rochester city limits. I searched for ‘Pizz’ rather than ‘Pizza’ or ‘Pizzeria’ because when searching for one or the other I would lose relevant restaurants. When I made the call to the API, the output was in a JSON file, so I used the json\_normalize method within the pandas library for Python to convert the results in to a useable pandas data frame.

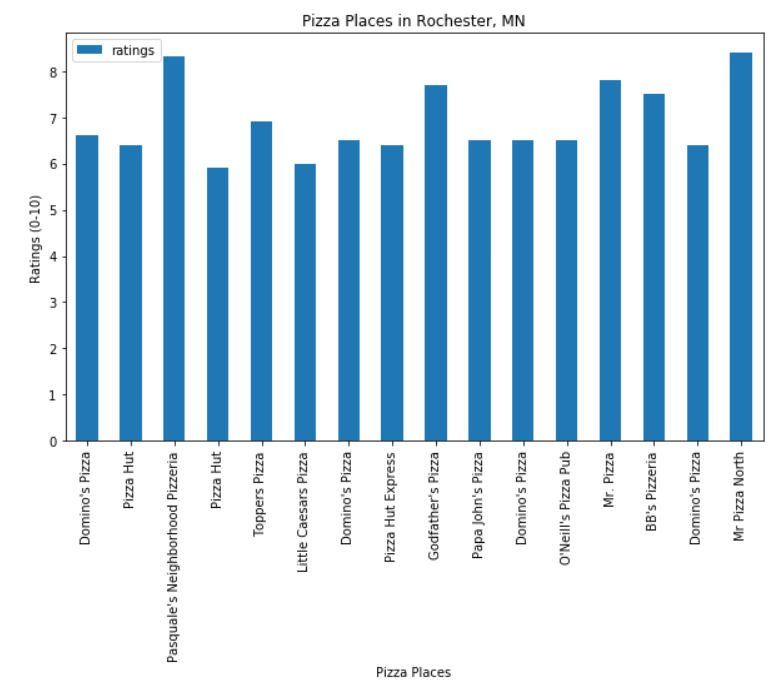
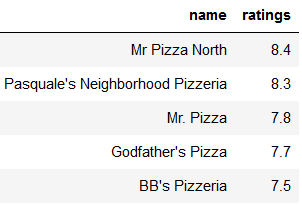
From there I filtered the resulting data frame down to some of the data points that mattered to me, including the venue name, category, latitude and longitude values, and ID among other things. The ID was crucial because I had to do a separate API call on each ID individually to compile each pizza place’s ratings and number of likes into a useable format. There were thirty restaurants that were pulled initially, however I removed any elements without a rating, which reduced my data to sixteen elements. Unfortunately when I initially pulled the number of likes for each restaurant it only returned sixteen results which did not match my initial thirty, so I continued without this information. My assumption is that the values I did pull corresponded to the sixteen values that remained after removing restaurants without ratings, but I don’t know for sure.

1. **Analysis**

Given the nature of my problem, I did not require the use of any machine learning or modeling. In fact by the time I got to my analysis stage, the answer was basically right in front of me! During my data cleaning, I plotted the pizza restaurants on a map of Rochester, MN in case I wanted to take distance into account for a more personalized recommendation at a later date. The map I generated is shown below.

*Map of pizza restaurants shown as blue dots in Rochester, MN*

After retrieving and cleaning the data, I essentially just needed to sort the values to see which restaurant the best user rating from those who use Foursquare.



It looks like Mr. Pizza North is the winner by a narrow margin at this standpoint. I thought that Pasquale’s Neighborhood Pizzeria had a close enough rating to warrant a tiebreaker though, so I tried pulling each of their individual number of likes from Foursquare. Luckily I was able to get both and Mr. Pizza North still won out with forty likes compared to Pasquale’s Neighborhood Pizzeria’s eighteen likes. By this analysis, Mr. Pizza North has the best pizza in Rochester, MN.

1. **Conclusion**

According to the process I used for my analysis Mr. Pizza North supposedly has the best pizza in town for Rochester, MN, however this comes with many caveats. Ideally, I would have liked to also include the number of dislikes and possibly user tips to see what people said about the pizza as well. There were also a few restaurants that came up a few times because they’re chains such as Pizza Hut and Domino’s Pizza, which could skew results. Another thing to consider that the ratings and likes are more likely tied to the restaurant as a whole rather than just the pizza itself, even if Pizza is primarily what they sell. Ratings and number of likes take things such as customer service, facility and other factors into account. This simple rating system could easily be applied to other types of restaurants in other cities and be used in a variety of use cases. At the end of the day, I know I need to go and check out Mr. Pizza North for my next pizza!