Open a Restaurant in Rome - Where's better?

Introduction

If someone is looking to open a restaurant, where would you recommend that they open it?

Rome is the capital city of Italy as well as the capital of the Lazio region. The city has been a major human settlement for over two millennial. With 2,860,009 residents in 1,285 km2 (496.1 sq mi),it is also the country's most populated comune. It is the third most populous city in the European Union by population within city limits. It offers many business opportunities for new bees so is competitive. The analysis from this report help for the new businesses strategically target the market and help in a high return on investment so is low risk.

Description

Opening a new restaurant can be daunting. There are so many things to consider - from choosing the right location to finding financing to selecting the right name. And, of course, buying equipment and hiring staff. The first step in opening a new restaurant is deciding what type of restaurant it is going to be. Do you want to open a high-end fine dining restaurant? A casual 1950s-style diner? Do you have a specific type of cuisine in mind, such as Italian, French, or Indian? But a good question could be: **Is the restaurant located in a good position?** For example, is the location busy with plenty of foot traffic? If not, is there enough parking? Have several other restaurants opened and closed in the same spot?

We will focus on the types of different restaurants which are opened or closed in a particular place and then decide if it is a good place to open a new restaurant based on the popular cuisine around that place.

Target Audience

The objective of this project is to recommend which neighborhood of Rome is a good choice for a new restaurant business to open.

Data

We will be using a list of Rome Neighborhood data taken from Wikipedia (https://en.wikipedia.org/wiki/Quarters_of_Rome) for analysis. There is a total of 35 neighborhoods in Rome. We will find latitude and longitude of each neighborhood, then we will use Rome geographical data as input to the Foursquare location data and fetch top 50 restaurants nearby to each neighborhood within 1000 meters. Then we will make a decision examining each cluster of neighborhoods.

Sample records from Wikipedia: (Q. I, Q. II, etc.. stands for Neighborhood 1, Neighborhood 2, and so on...)

- **Q. I** Flaminio
- **Q. II** Parioli
- **Q. III** Pinciano
- **Q. IV** Salario
- **Q. V** Nomentano
- **Q. VI** Tiburtino
- **Q. VII** Prenestino-Labicano
- **Q. VIII** Tuscolano

- **Q. IX** Appio-Latino
- **Q. X** Ostiense
- **Q. XI** Portuense
- **Q. XII** Gianicolense
- **Q. XIII** Aurelio
- **Q. XIV** Trionfale
- **Q. XV** Della Vittoria
- **Q. XVI** Monte Sacro
- **Q. XVII** Trieste

- **Q. XVIII** Tor di Quinto
- **Q. XIX** Prenestino-Centocelle
- **Q. XX** Ardeatino
- **Q. XXI** Pietralata
- **Q. XXII** Collatino
- **Q. XXIII** Alessandrino
- **Q. XXIV** Don Bosco
- **Q. XXV** Appio Claudio
- **Q. XXVI** Appio-Pignatelli

- **Q. XXVII** Primavalle
- **Q. XXVIII** Monte Sacro Alto
- **Q. XXIX** Ponte Mammolo
- **Q. XXX** San Basilio
- **Q. XXXI** Giuliano-Dalmata
- **Q. XXXII** Europa
- **Q. XXXIII** Lido di Ostia Ponente
- **Q. XXXIV** Lido di Ostia Levante
- **Q. XXXV** Lido di Castel Fusano

Then we will find the latitude and longitude of each neighborhood and map to the data frame.

	Neighborhood	Name	Latitude	Longitude
1	Q. II	Parioli	41.925949	12.487320
2	Q. III	Pinciano	41.918781	12.485332
3	Q. IV	Salario	41.940398	12.508896
4	Q. V	Nomentano	41.917002	12.512918
5	Q. VI	Tiburtino	41.896633	12.511238
6	Q. VII	Prenestino-Labicano	41.891549	12.516829
7	Q. VIII	Tuscolano	41.867638	12.539368

We will then use this Rome geographical data as input to the Foursquare location data and fetch top 50 restaurants nearby to each neighborhood within 1000 meters. The sample foursquare location data looks like this:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
1	Q. II	41.925949	12.487320	Metamorfosi Restaurant	41.924077	12.484641	Italian Restaurant
6	Q. II	41.925949	12.487320	La Pariolina	41.928589	12.486173	Italian Restaurant
7	Q. II	41.925949	12.487320	Ristorante Molto	41.927481	12.484284	Italian Restaurant
12	Q. II	41.925949	12.487320	Il Caminetto	41.928184	12.486310	Italian Restaurant
13	Q. II	41.925949	12.487320	Chez Cocò	41.929623	12.485371	Italian Restaurant
...	...	...	...	...	...	...	...
959	Q. XXXV	41.716500	12.323709	Sporting Beach	41.713333	12.317398	Italian Restaurant
960	Q. XXXV	41.716500	12.323709	La Mariposa	41.709623	12.326375	Italian Restaurant
961	Q. XXXV	41.716500	12.323709	Kursaal	41.715174	12.315824	Italian Restaurant
962	Q. XXXV	41.716500	12.323709	kursal	41.715134	12.315835	Italian Restaurant
964	Q. XXXV	41.716500	12.323709	Da rutilio a mare	41.715249	12.314477	Italian Restaurant

Methodology

We will follow this step:

1. We first take the neighborhood information from Wikipedia and load that into a data frame dropping unnecessary columns.
2. Then we perform data wrangling to convert the data into the analysis-ready form.
3. Then we use the geopy library to get the latitude and longitude information for each neighborhood.
4. Then we use this data to get the 50 nearby restaurants within 1000 meters of a neighborhood using foursquare location data.
5. Calculate and perform unique restaurant categories in each neighborhood and find the top 10 most common restaurants.
6. Then we will cluster all the restaurants in all the neighborhoods using the k-means clustering algorithm.