

Capstone Project of IBM Data Science Specialisation

By Andrea Di Sarra

Select optimum location to start up a new wine cellar shop in the City of Milan

Introduction

The city of Milan is one of the most interesting Business Case in Italy for its proximity to other major European capitals and for an undoubtable ecosystem which attract investors, therefore creating an hub whereas most of the Italians, especially the southern, are willing to move.

Milan is also famous all over the world for the fashion industry, the city centre through the famous “Quadrilatero della Moda” has a variety of fashion shops of the most important brands (i.e. Prada, Gucci etc.) which is home for travellers and fashion bloggers.

The city is hosting the headquarters of many companies, included but not limited to those of the energy or engineering sectors, the financial / stock market, legal firms etc. Milan is also home of top ranked universities which enhance the variety of the population which is established in the city.

It is also well-known that **tourists** from abroad are attracted from **Italian cuisine** and **wines**, and those three components are key pillars of the Italian economy.

Business Problem

The objective of this capstone is to analyse, study and select the most attractive place where to establish a new wine cellar shop where to showcase Italians' wines.

The business question behind this objective is therefore the following: if a private investor in the field of the wine production is interested to open up a new shop in the city of Milan, where do you would recommend to start up the activity?

Target Audience

There are so many excellent wine producers in Italy, as well as company which are active in the distribution of such products. Any of those players, the owner of a selected brand, re-seller or private investor who want to start up a new business, might be willing to understand optimum location where to establish their shop in one of the most diverse city of Italy.

Data Description & Sources

- The nine boroughs of Milan
https://en.wikipedia.org/wiki/Municipalities_of_Milan
- The districts of Milan
https://en.wikipedia.org/wiki/Category:Districts_of_Milan
- The Latitude and Longitude of boroughs and districts as required.
- Venue data, paying particular attention to the restaurants & wine shops to cluster the city of Milan accordingly.

Methods and Application of Data Science packages

The data as resulted from web scarping, are imported in Python and data wrangling, cleaning and transformation is performed through the help of beautiful soap package. Locations are being added via Python Geocoder package and venues information are taken by leveraging on Foursquare API. Clustering methods through application of K-means algorithm is uded to characterise the city and its districts, therefore through ranking of main activities to understand the nature of existing businesses. Foilum library is applied to facilitate the data visualization to make faster and most informed decisions.