

Predicting the Improvement of NBA players

Sarika Jain

16th May, 2020

Introduction:

In this project I will try to find an optimal location for a Vegan restaurant. Specifically, this report will be targeted to stakeholders interested in opening a Vegan/Vegetarian restaurant in Perth, Australia.

Since there are lots of restaurants in Perth, I will try to detect locations that are as close to city center (Hayatt Regency Hotel) as possible.

I will use my data science knowledge to generate a few most promising neighborhoods based on this criteria. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

Data :

Based on definition of this problem, factors that will influence my decision could be as follows :

No. of existing Vegetarian restaurants in the neighborhood.

No. of and distance to other vegetarian/Vegan restaurants in the neighborhood, as well as from Hayatt Hotel.

Following data sources will be needed to extract/generate the required information: Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using Foursquare API interface.

No. of restaurants and their location in neighborhood will be obtained using Foursquare API coordinate of Perth center will be obtained using geocoder.



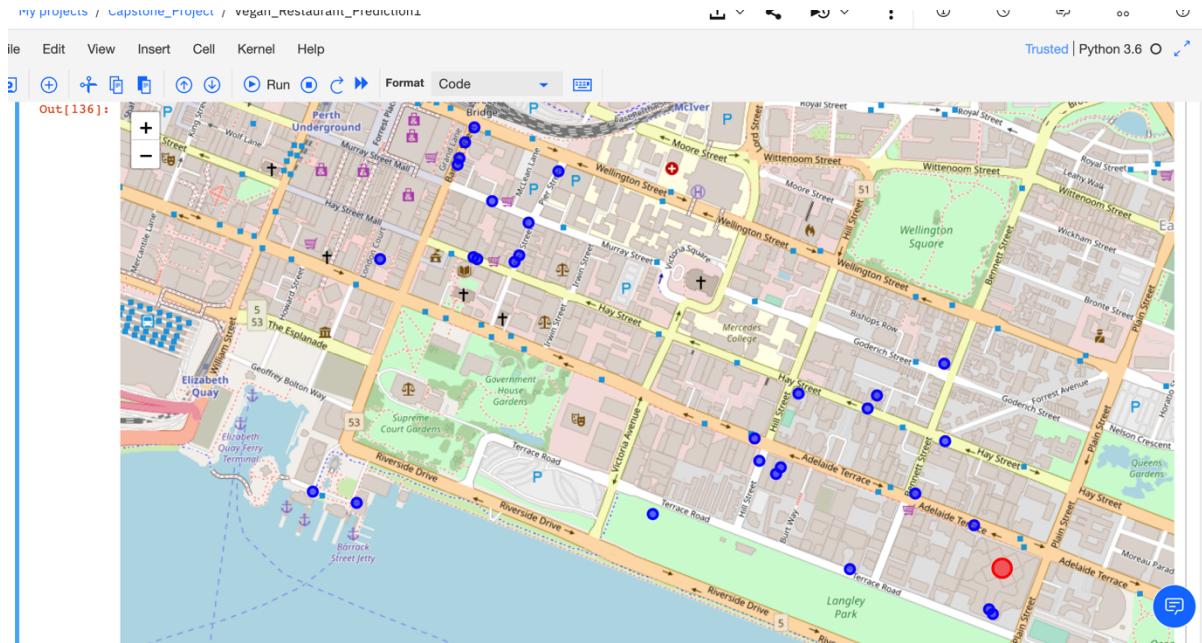
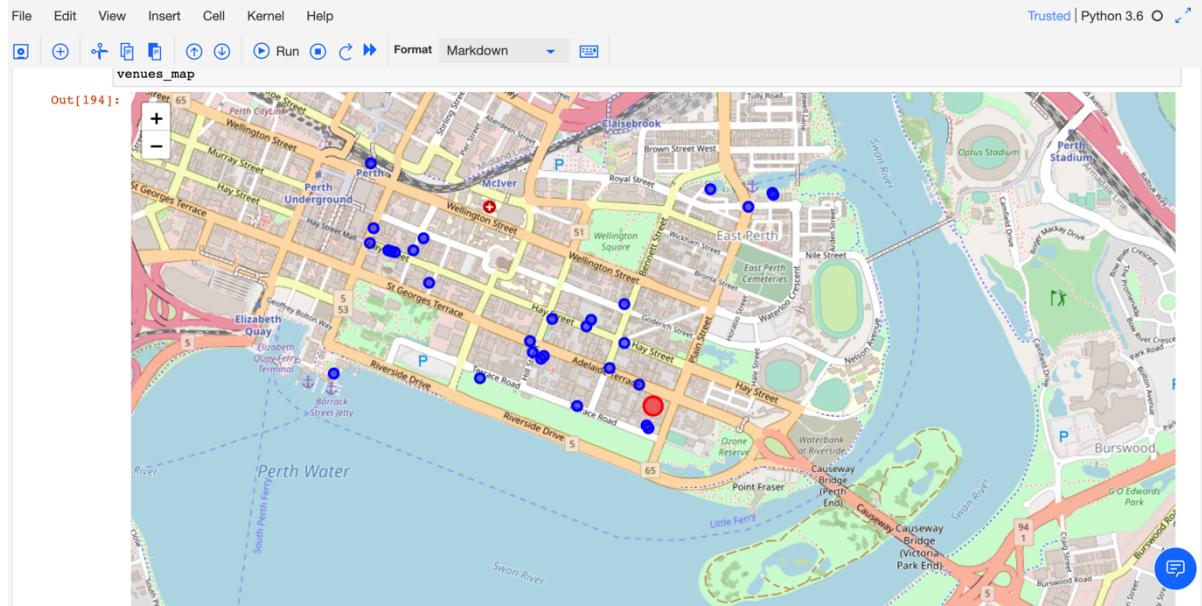
Methodology :

In this project we will direct our efforts on detecting areas of Perth that have low restaurant density, particularly those with low number of Vegetarian/Vegan restaurants. We will limit our analysis to area 1.5km around city center.

In first step we have collected the required data: location and type (category) of every restaurant within 1.5km from Perth center (Hayatt Regency Hotel). We have also identified Vegetarian restaurants (according to Foursquare categorization).

Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of Perth - we will use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Vegetarian/Vegan restaurants in vicinity) and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two restaurants in radius of 250 meters, and we want locations without Vegetarian/Vegan restaurants in radius of 400 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

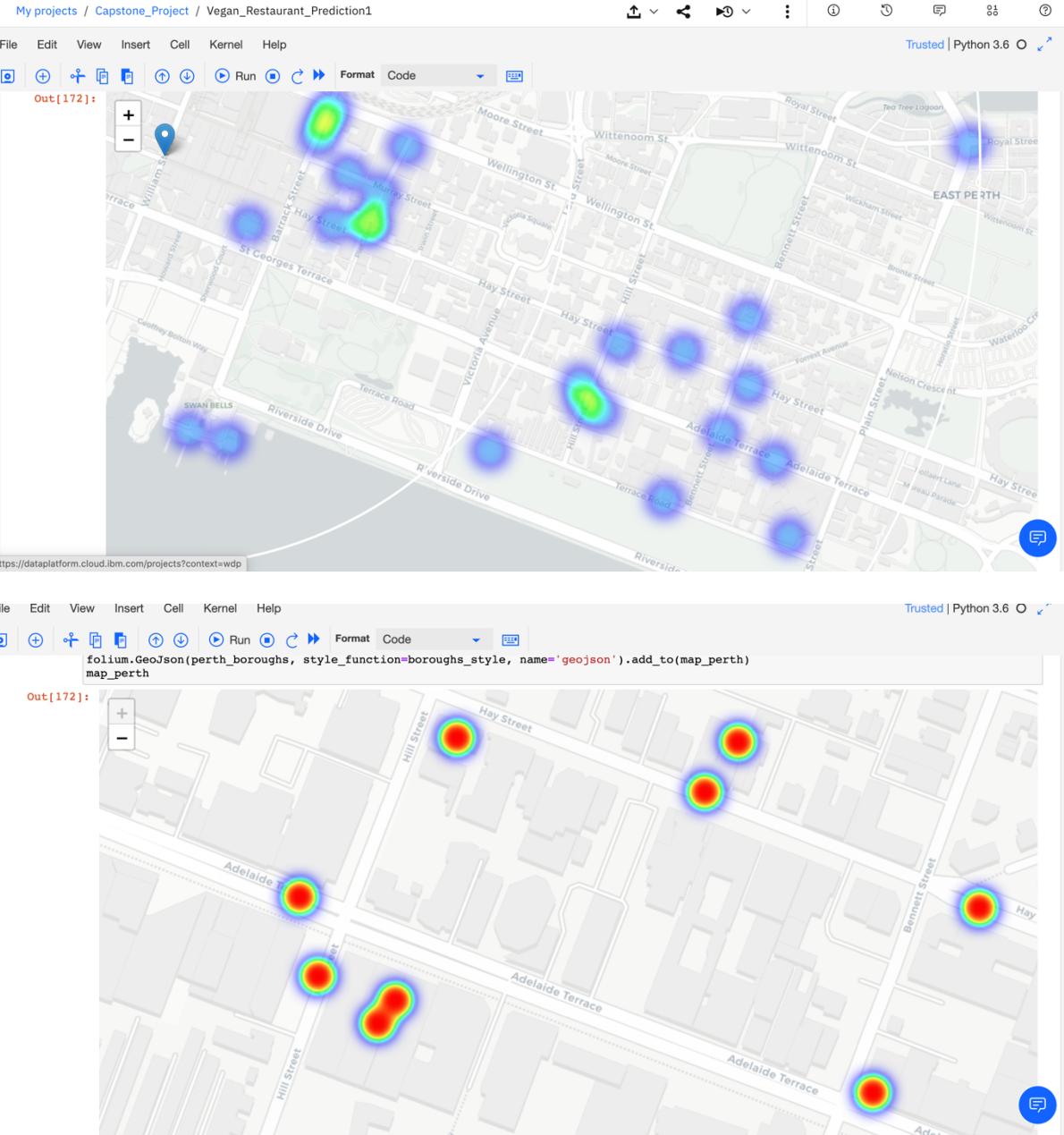


Analysis :

I tried to do clustering of the areas on the basis of the density of the restaurants. And performed some basic explanatory data analysis and derived some additional info from our raw data. First counted the number of restaurants in every area candidate. Then plotted the clusters there.

Then tried to find out other affecting factors in the particular Area.

I used K-means, and Density based clustering methods.



Results and Discussion :

Our analysis shows that although there is a great number of vegetarian/vegan restaurants in Perth, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north and west from Hayatt Hotel, so we focused our attention to areas south, south-east and east. Our attention was focused on Forrest Chase which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low restaurant density.

After directing our attention to this more narrow area of interest (covering approx. 2x2km south-east from Forrest Chase) we first created a dense grid of location candidates (spaced

100m apart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Italian restaurant closer than 400m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Based on largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Vegan restaurants particularly are expected. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Perth center but not crowded with existing restaurants - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area.

Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion :

Purpose of this project was to identify Perth areas close to center with low number of restaurants (particularly vegetarian/Vegan restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Vegetarian restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general Areas that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.