Hotel’s surroundings

# **Introduction**

Although there are many travel search engines available on the internet, for example: Trivago, Hotwire, Kayak and others. Most of them allow the user to search for hotels based on travel dates and on price. However, sometimes the user is more interested in the features around the hotel and on its accessibility. In this case, the user, who might not be familiar with the location, have to read reviews to try to understand the area, which is a time-consuming activity. In some cases, the user might not have time to do that. This study aims at showing the applicability of a search engine that takes into account what the user expects from the surrounding area around a hotel. The aim of this report is not to produce a final guideline for an application, it will instead show a case study of the functionality of such tool.

The case study will be based in Montreal. A hypothetical user is looking for a hotel in Montreal, a place it he’s never been for a business trip. Since his company is paying for the stay, he is not very concerned about the price, he is more concerned about the hotel surroundings. He is looking for the following three features:

* Easy access to public transit, specially by metro, because it is easier to ride it than to use buses when a person doesn’t know the city. He is willing to walk 500m to a metro station.
* He wants to have access to a grocery store because he hates hotel food and he always forgets to take important hygiene items. He is willing to walk 500m to a grocery store.
* He wants easy access to a post office, because he loves sending post cards when he travels. E is willing to walk 300m to a post office.

Although it is a single user case, the study will show that it is possible based on a person’s preference to choose a better location for a hotel.

# **Data**

For this project the data will be collected exclusively from Foursquares. However, the data will be assembled to form a larger dataset. Four searches will be performed: the first one to find all the listed hotels in an 8 km radius from downtown Montreal, then three other searches using the same radius will be made to find groceries, metro stations and post offices. Once all information is available, a database will be built with all the hotels and with a field indicating if in the acceptable radius around the hotel there is a venue that the user wants. For example, if in a radius of 500m around the hotel there is a grocery store, the column ‘n\_grocery’ will have the value ‘1’for that hotel.

# **Methodology**

The methodology can be split in the following 8 steps (thee same steps were indicated in the Jupyter Notebook)

1. Import Required Libraries
2. Extract Hotel, Grocery, Metro Stations and Post Office Data from Foursquares
3. Organize JSON in data frames
4. Overview geographical data with Folium
5. Build main data frame based on distances between hotels and venues.Distances were calculated using Pythagoras. There are many problems with this approach such as:

* It is 'as the crow flies' distance and not the real distance that he will walk (maybe a Manhattan approach should have been used)
* It is a rough estimate (doesn’t consider the earth curvature).

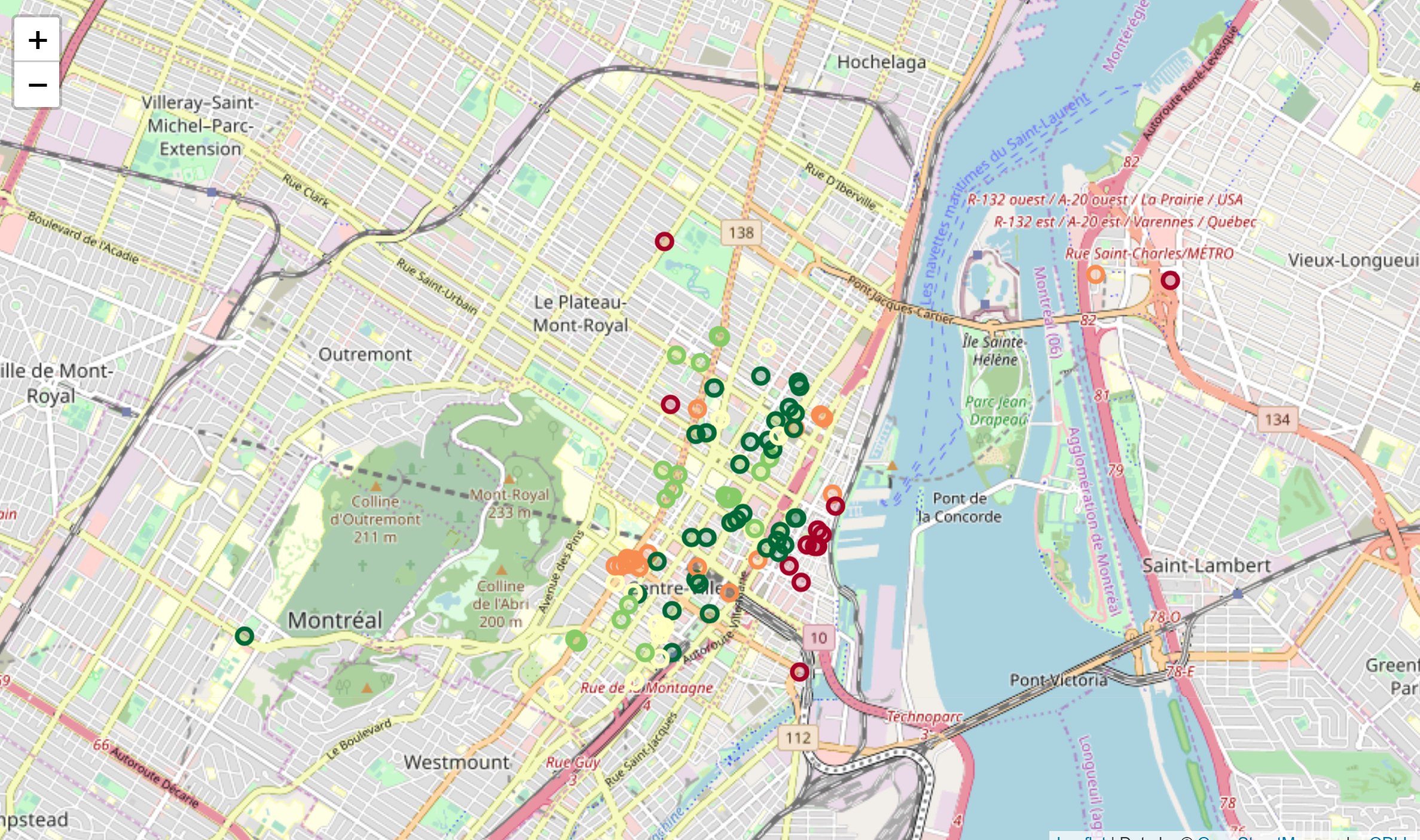
However, it was assumed that for the purpose of this study, Pythagoras was sufficient to exemplify the problem.

1. Use K-means to classify hotels in 5 groups
2. Create grades based on clusters (clusters don't respect user preference, so we need to evaluate which ratings are best for the user)
3. Create a color-coded map, to show in green the best hotels and in red the worse hotels to the user

# **Results**

The final result is shown in the image below. It is a color-coded map where the dark green hotels are better for the user and the red ones are the worse. The map also shows hotels with intermediate ratings according to its surroundings.

The user can now restrict its searches for the best hotel.



# **Discussion**

This case study showed that with easily available data on Foursquares it is possible to use a simple and know algorithm to restrict a person choice based on their preference without knowing anything about a city. Ideally less values would be shown so that the person could make a choice easily. However, the study shows that better search engines can exist. Its purpose was not to build the tool, but to show its capabilities.

# **Conclusion**

It is possible to enhance travel search engines so that the surroundings of hotel can be evaluated while choosing it. A clustering algorithm was used to classify all hotels in a 8km radius from Downtown Montreal based on three criteria determined by a hypothetical user.