USING FOURSQUARE INFORMATION TO HELP ASSESS A NEW LOCATION FOR A HOTEL

CAPSTONE PROJECT - THE BATTLE OF NEIGHBORHOODS

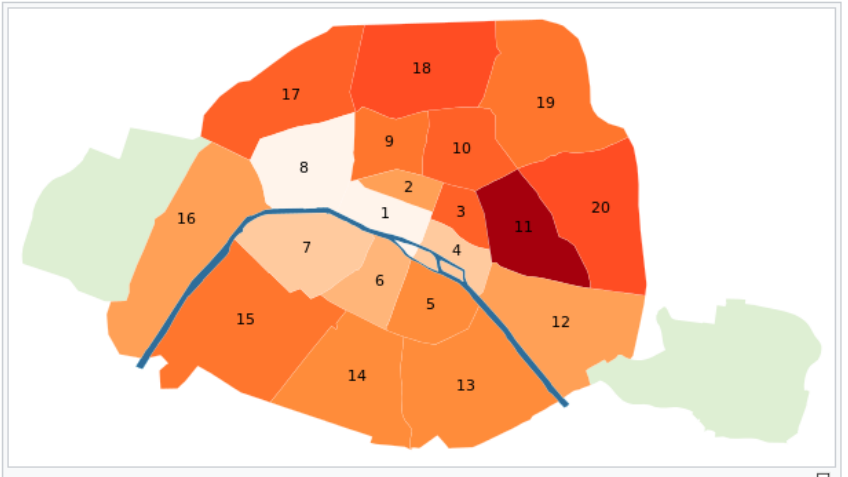
VILLE RAUTOJA

INTRODUCTION

- A boutique hotel located in the area of Atocha in Madrid, Spain would like to open up a new hotel branch in Paris, France.
- As their clientele chooses their hotel based on the services nearby, the hotel owners would like to find a similar neighborhood in Paris so that the new hotel would attract the same segment of customers.
- This exercise takes advantage of venue information available from the Foursquare API, Wikipedia and various modules available for Python, such as Geopy, Folium, Requests and K-means from Sklearn to help process the data and derive an analysis of the possible matching locations.

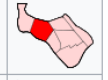
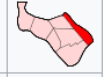
DATA & METHODOLOGY

- Neighborhood information is scraped from Wikipedia sites listing the areas in Paris and Madrid.
- The coordinates of the selected areas are then defined with Geopy
- A table of venues in each area is then found with the Foursquare API
- The unique locations per location are reviewed and then one hot encoding is done to standardize the venue results, and finally top 10 common venues is created.
- K-means clustering is used to assess locations similar to Atocha area
- Venues of Atocha and suggested area in Paris are plotted with Folium.



Population density map of Paris in 2012.

Arganzuela

21	Imperial	
22	Acacias	
23	Chopera	
24	Legazpi	
25	Delicias	
26	Palos de Moguer	
27	Atocha	

RESULTS AND DISCUSSION

- Using a K-means clustering with 3 groups gives us 2 areas in Paris that could be similar to Atocha.
- After further assessment, Batignolles-Monceau is taken into further analysis.
- Batignolles-Monceau has similar variety of restaurants, bars and other venues compared to Atocha available. Additionally, after assessing the geographical location, it is noticeable that both areas are in fact near train stations. This would potentially indicate that areas near major train stations in large cities have similar venues.
- The hotel owners are then provided the outcome to take into consideration when they decide their new location.
- The approach can be used to assess other cities in other countries with a minimal investment, saving time and effort while providing meaningful insights.

```
1 df_merged.loc[df_merged['Cluster Labels'] == 1]
```

	District	Neighborhood	lat	lng	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th I Com Ve
0	Arganzuela	Atocha, Madrid, ES	40.400687	-3.682391	1	Spanish Restaurant	Bar	Café	Italian Restaurant	Restaurant	Hotel	Asian Restaurant	Museum	Bi
4	4th (IVe) R	Hôtel-de-Ville, Paris, FR	48.856426	2.352528	1	French Restaurant	Ice Cream Shop	Art Gallery	Plaza	Wine Bar	Pub	Park	Hotel	
17	17th (XVIIe) R	Batignolles-Monceau, Paris, FR	48.881312	2.315750	1	French Restaurant	Hotel	Wine Bar	Italian Restaurant	Bar	Thai Restaurant	Pastry Shop	Theater	C