

Finding Best Locations for Restaurant Business In The Hague City In The Netherlands

Solomon Edwin Musa

21 April 2019

INTRODUCTION

The Hague City in the Netherlands (South Holland) is the target city of this project. The project is about assessing the best locations for a restaurant within the eight urban districts (Boroughs) comprising the Hague City in South Holland.

BACKGROUND AND PROBLEM

A group of immigrants arrived from Asia and South America to seek asylum in the Netherlands and have been given a stay in the Hague City by the government of Netherlands and City Council of the Hague. These asylum seekers have made it known to the authorities that they will be better integrated into the Dutch society if they are provided opportunity to start and engage in restaurant business, which they claimed to be their profession. However, the Hague City, like other parts of the Netherlands, is already a scene of diverse type of restaurants everywhere, and with possible restrictive regulations on business locations, there is need to guide immigrants who want to establish restaurant businesses. The City Council contracted me as a Data Scientist and Analyst to investigate the best possible locations for restaurants, and to provide them recommendations on possible locations in a presentation, the result of which will be used to determine whether or not council will provide financial assistance to these immigrants in order to start their food restaurant businesses of interest.

DATA ACQUISITION/SOURCES

1. For description of the Hague City, Districts/Boroughs, Neighborhoods, and Population, data was manually extracted from this link: <https://www.angloinfo.com/blogs/netherlands/south-holland/expatinfo-holland/city-districts-the-hague/>. There was no need for web scraping because data was readily available
2. Similarly, latitude and longitude values for boroughs (referred to as districts in the Netherlands) and neighborhoods were also manually acquired from this link: <https://www.distancesfrom.com/Den-Haag-latitude-longitude-Den-Haag-latitude-Den-Haag-longitude/LatLongHistory/14049.aspx?IsHistory=1&LocationID=14049> . The reasons for manual extraction are due to cost chargeable for such data on proprietary Web sites, because they're alternatively freely available on other Dutch sites and small number of districts and neighborhoods
3. Proprietary Web sites include the following as examples: <https://www.postcode.nl/services/adresdata/api> ; and,
4. <https://api.postcode.eu/nl/v1/addresses/latlon/>

METHODOLOGY/PROCEDURE FOR DATA ANALYSIS

The procedure for segmenting and clustering neighborhoods was used in this research and the steps are here simply itemized and which were followed for each of the three districts or boroughs investigated, including, Hague Centrum, Escamp, and Scheveningen. The research was limited to this three as a matter of convenience and they represent examples of a district or borough as a seat of government, the district or borough with largest population, and a district or borough with relatively lower population but with the highest number of neighborhoods. The Hague Centrum is seat of government, with six neighborhoods, and a population of 103,279, Escamp has four neighborhoods with the highest population of 122,825 compared to all districts, and Scheveningen has eight neighborhoods with a population of 55,976. Additionally, due consideration was given also to resource use optimization as well as the limitation posed by foursquare on the calls per day for venues.

The procedures used are as follows after importing all necessary libraries for the analysis:

1. The Hague City dataframe (thcty_data) was created manually with data obtained from the sources mentioned above;
2. Hague Centrum data (hgcentrum_data) was created by slicing the Hague City dataframe;
3. Used geopy to get the Hague Centrum coordinates but creating map with folium failed;
4. Used Foursquare API procedures, then explored the first neighborhood of Hague Centrum, and getting its latitude and longitude values;
5. To get the top 10 venues in this neighborhood and within a radius of 500m and a limit of 100, a url was defined and used with the GET requests to retrieve venues results from foursquare;
6. A get_category_type function from foursquare was used that extracts the categories of each venue;
7. The resulting json file was cleaned and structured into a Pandas dataframe and to get nearby_venues;
8. To explore all neighborhoods in Hague Centrum, a function to repeat the procedure applied to exploring the first neighborhood was used that returned the nearby_venues;
9. A code to run the function above (in 8) was used on each neighborhood and then a new dataframe hgcentrum_venues, and how many venues returned per neighborhood a groupby method on neighborhood with count function was used, followed by getting unique categories returned by all venues;
10. To analyze each neighborhood and prepare for clustering later, a onehot encoding was applied to create hgcentrum_onehot dataframe;
11. The onehot dataframe was then used to group rows by neighborhood and by taking the mean of the frequency of occurrence of each category, defined as hgcentrum_grouped;
12. Each neighborhood was printed along with the top 5 most common venues;
13. The above was put into a dataframe by first using a function to sort the venues in descending order, then a new dataframe hgcentrum_venues_sorted was created;
14. The neighborhoods were then clustered by running k-means into a number of clusters, 2, 4, or 5 depending on the number of venues earlier returned for each neighborhood;
15. Then a new dataframe including the clusters and the top 10 venues for each neighborhood hgcentrum_merged was created, and again, attempt to visualize a map with folium failed; and,
16. The clusters were examined, but in all cases only the first cluster returned a result, probably due to low number of venues returned by each neighborhood.

Therefore, the most important results returned and presented are the hgcentrum_venues_sorted, hgcentrum_merged, and hgcentrum_centrum cluster 1 for each district or borough investigated.

DATA CLEANING

Since almost all the data required for this research was free and readily available and the fact that there are a small number of districts or boroughs as well as neighborhoods comprising the Hague City, the dataframe was created manually for convenience and named hgcty_data. Data cleaning was therefore not an issue in this case, except for the foursquare venues data acquired for first and the rest of neighborhoods for each of the three districts or boroughs investigated, including, Hague Centrum, Escamp, and Scheveningen districts or boroughs and their neighborhoods.

For the cleaning of the foursquare data returned, the get_category-type function was used firstly and followed by the normal procedure for cleaning the returned json file and to structure it into a pandas dataframe, with respect to the procedures for segmenting and clustering neighborhoods, and with reference to the procedures itemized above.

RESULTS

THE HAGUE CITY DATA

Table 1: thcty_data

(38, 4)

[38]:

	Borough	Neighborhood	Latitude	Longitude
0	Hague Centrum	Archipelbuurt-Willemspark	52.097090	4.300937
1	Hague Centrum	Zeeheldenkwartier	52.082520	4.299584
2	Hague Centrum	Kortenbos	52.077037	4.302460
3	Hague Centrum	Transvaalkwartier	52.066831	4.291122
4	Hague Centrum	Schildersbuurt	52.068487	4.300321
5	Hague Centrum	Stationsbuurt	52.071510	4.316363
6	Escamp	Rustenbuurt-Oostbroek	52.060426	4.284248
7	Escamp	Wateringse Veld	52.027100	4.289700
8	Escamp	Moerwijk	52.047836	4.289200
9	Escamp	Bouwlust-Vrederust	52.037498	4.256879
10	Haagse Hout	Benoordenhout	52.093245	4.322308
11	Haagse Hout	Marlot	52.099139	4.351493
12	Haagse Hout	Haagse Rose	52.086100	4.310900
13	Haagse Hout	Mariahoeve	52.093644	4.359234
14	Haagse Hout	Bezuidenhout	52.084096	4.339665
15	Haagse Hout	Beatrixkwartier	52.080367	4.334579
16	Laak	Binckhorst	52.067695	4.340051
17	Laak	Spoorwijk	52.053541	4.315263
18	Laak	Laakkwartier	52.056142	4.320906

The dataframe above is the result of manually creating it using data acquired from sources and reasons for choosing to do so and limiting the research to only three districts or boroughs given above. There are 38 rows and 4 columns. The 38 rows mean all the neighborhoods column values, including latitude and longitude column values. There are actually eight boroughs/districts, but each district has a number of neighborhoods and a total of 38 neighborhoods in the Hague city. It was impossible to use postal codes because in the Netherlands grouping them is not feasible as postal codes are based on streets and also, names of neighborhoods or streets may be the same for a few cases in other cities and even in the same city in different boroughs is possible. Only part of the full table is provided above due to space.

HAGUE CENTRUM RESULTS

Table 2: hgcentrum_venue_sorted

[29]:

	Neighborhood	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
0	Archipelbuurt-Willemspark	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
1	Kortenbos	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
2	Schildersbuurt	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
3	Stationsbuurt	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
4	Transvaalkwartier	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
5	Zeeheldenkwartier	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop

Table 3: hgcentrum_merged

[33]:

	Borough	Neighborhood	Latitude	Longitude	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
0	Hague Centrum	Archipelbuurt-Willemspark	52.097090	4.300937	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
1	Hague Centrum	Zeeheldenkwartier	52.082520	4.299584	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
2	Hague Centrum	Kortenbos	52.077037	4.302460	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
3	Hague Centrum	Transvaalkwartier	52.066831	4.291122	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
4	Hague Centrum	Schildersbuurt	52.068487	4.300321	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
5	Hague Centrum	Stationsbuurt	52.071510	4.316363	0	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop

Table 4: hgcentrum cluster 1

[34]:

	Latitude	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
0	52.097090	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
1	52.082520	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
2	52.077037	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
3	52.066831	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
4	52.068487	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop
5	52.071510	Restaurant	Theme Park	Tram Station	Theme Park Ride / Attraction	Tennis Court	Park	Gift Shop

ESCAMP RESULTS

Table 5: escamp_venues_sorted

[68]:

	Neighborhood	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
0	Bouwlust-Vrederust	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
1	Moerwijk	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
2	Rustenbuurt-Oostbroek	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
3	Wateringse Veld	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop

Table 6: escamp_merged

[72]:

	Borough	Neighborhood	Latitude	Longitude	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
0	Escamp	Rustenbuurt-Oostbroek	52.060426	4.284248	0	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
1	Escamp	Wateringse Veld	52.027100	4.289700	0	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
2	Escamp	Moerwijk	52.047836	4.289200	0	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
3	Escamp	Bouwlust-Vrederust	52.037498	4.256879	0	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop

Table 7: escamp_cluster 1

[73]:

	Latitude	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
0	52.060426	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
1	52.027100	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
2	52.047836	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop
3	52.037498	Park	Bus Stop	Soccer Field	Soba Restaurant	Snack Place	Mini Golf	Fish & Chips Shop

SCHEVENINGEN RESULTS

Table 8: schev_venues_sorted

[56]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue
0	Belgische Park	Café	Monument / Landmark
1	Duindorp	Café	Monument / Landmark
2	Hof van Schreveningen	Café	Monument / Landmark
3	Intl Zone	Café	Monument / Landmark
4	Oostduinen	Café	Monument / Landmark

Table 9: schev_merged

[60]:	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue
0	Scheveningen	Oostduinen	52.115508	4.303549	0	Café	Monument / Landmark
1	Scheveningen	Belgische Park	52.109199	4.294428	0	Café	Monument / Landmark
2	Scheveningen	Westbroekpark	52.103916	4.293610	0	Café	Monument / Landmark
3	Scheveningen	Van Stolkpark	52.098634	4.292792	0	Café	Monument / Landmark
4	Scheveningen	Hof van Schreveningen	52.092787	4.261920	0	Café	Monument / Landmark
5	Scheveningen	Statenkwartier	52.094523	4.279591	0	Café	Monument / Landmark
6	Scheveningen	Intl Zone	52.059572	4.221827	0	Café	Monument / Landmark
7	Scheveningen	Duindorp	52.090618	4.259287	0	Café	Monument / Landmark

Table: 10 schev_cluster 1

[61]:	Latitude	1st Most Common Venue	2nd Most Common Venue
0	52.115508	Café	Monument / Landmark
1	52.109199	Café	Monument / Landmark
2	52.103916	Café	Monument / Landmark
3	52.098634	Café	Monument / Landmark
4	52.092787	Café	Monument / Landmark
5	52.094523	Café	Monument / Landmark
6	52.059572	Café	Monument / Landmark
7	52.090618	Café	Monument / Landmark

DISCUSSIONS AND CONCLUSION

Table: 2 hgcentrum shows the first most common venues in the Hague Centrum and its 6 neighborhoods are restaurants, second are theme parks providing different types of amusements and attractions for adults, teenagers, and children, third is tram stations meaning a transport hub, fourth is another type of theme park for riding attractions, fifth are tennis courts, sixth are normal nature parks, and seventh are gift shops. In summary a lot of humans are attracted to this district, especially in summer when activities peak. These could likely be the reason for attracting the establishment of restaurant businesses. This is evident in all results presented for Hague Centrum, Tables: 3 and 4.

Tables: 5 – 7 are results of the same analysis procedures followed for the Escamp district or borough and the 4 neighborhoods. The first most common venues are normal nature parks, second are bus stops, third are soccer fields, fourth are the specific private restaurants named Soba, fifth are snack places, sixth are mini Golf fields, and seventh are fish chips shops. It indicates different location for mixed type of foods and restaurant businesses. This could be as a result of the type of venues around quite different from the theme parks in the Hague Centrum which provide a big center of amusements and

attractions. Here could be rated second possible location for mixed type of restaurant or food businesses.

Tables: 8 – 10 are results for Scheveningen district or borough with the highest number of 8 neighborhoods, compared to the rest of the other districts or boroughs making up the Hague City. The results show the first most common venues to be cafes, followed by Monuments or Landmarks. This is a totally different location and mostly cafes are found and not much venues of attractions compared to other districts or boroughs. That does not mean restaurants may not be present or cannot be established if law or government regulations does not disallow that.

FUTURE DIRECTIONS

Only three districts or boroughs were investigated with Hague Centrum showing the best locations for a restaurant business, followed by Escamp district or borough, and lastly Scheveningen. The other districts or boroughs, five in total, not investigated by this research could be possible additional locations for restaurant businesses and could even surpass Hague Centrum, one never can tell without investigating.

The next steps to further improve on this research are:

1. To investigate whether any strong relationship exists between restaurants establishment and the presence of theme parks as well as other centers of human activities around and which types, through regression and predictive models using Machine Learning techniques.
2. Other districts could add to the understanding of possible suitable locations if further investigation is done on the remaining districts or borough not investigated by this research.
3. It will be good also to investigate what regulations or policies, if any, are in place to guide the location of restaurant businesses, and perhaps for other human activities in each district or borough.