1. **Data**

In order for there to be any meaningful analysis later on in the paper, data will have to be collected. The data that needs to be found includes the locations of the boroughs in Amsterdam, their population densities, safety ratings, and list of venues in and around each borough. The name of the boroughs, as well as their population densities, are easily found on [Wikipedia](https://en.wikipedia.org/wiki/Boroughs_of_Amsterdam) and do not need to be converted further. The locations (coordinates) of each borough were found on the City of Amsterdam’s very own [government website](https://maps.amsterdam.nl/open_geodata/?LANG=en). As you can see, the government website refers to the boroughs as “districts” while Wikipedia lists them as boroughs. In this paper, I will use the terms “borough”, “district”, and “neighborhood” interchangeably. The table below (Table 1) shows the names, coordinates, and densities of the eight districts in Amsterdam after the data has been turned into a dataframe on Jupyter Notebooks.

**Table 1: Amsterdam’s Eight Districts**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Neighborhood** | **Latitude** | **Longitude** | **Population Density (people/km2)** |
| **0** | Centrum | 52.37321 | 4.903712 | 13748 |
| **1** | Westpoort | 52.41095 | 4.803871 | 10 |
| **2** | Zuidoost | 52.30465 | 4.974994 | 4391 |
| **3** | West | 52.37611 | 4.86452 | 15252 |
| **4** | Nieuw-West | 52.36407 | 4.802676 | 4478 |
| **5** | Oost | 52.34981 | 4.956049 | 7635 |
| **6** | Zuid | 52.34172 | 4.86605 | 9349 |
| **7** | Noord | 52.388 | 4.917663 | 2269 |

The next variable that needs to be found is each boroughs’ safety index. Fortunately, I speak Dutch, so I was able to, again, go to the Government of Amsterdam’s website and read through their 2019 publication aptly titled “[City District in Numbers 2019](https://data.amsterdam.nl/publicaties/publicatie/stadsdelen-in-cijfers-2019/f4b55c38-6b3c-4143-bea5-94c4f1cb0887/)”. This comprehensive publication lists and outlines many different variables and statistics for the eight districts in Amsterdam, including the safety ratings. These ratings go from low (safe) to high (dangerous). The report uses a simple, binary, way of indexing whether or not a neighborhood is considered “safe”: a Safety Index under 100 is considered safe, while a Safety Index over 100 is considered dangerous. As such, we will simply put a “1” for every neighborhood that is considered safe, and a “0” for ones that are not.

This brings us to our final variable: the number of venues in each district. Much like in the lab with New York, or the peer reviewed assignment with Toronto; here too we will be using [Foursquare](https://developer.foursquare.com/docs/) and its API function via Python to retrieve a list of venues nearby. I will not get into the specifics of this, as I am sure everyone taking this course has had plenty of experience by now doing this, but the code (with explanatory markdowns) can be found in the Jupyter Notebook that was shared along with this report on my [GitHub repository](https://github.com/rb2925a/Coursera_Capstone). The important things to note are that I used a radius of 500 meters (given Amsterdam’s relatively small size) and a limit of 200 venues per district.

**Table 2: Most Common Venues by Neighborhood**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Neighborhood** | **1st Most Common Venue** | **2nd Most Common Venue** | **3rd Most Common Venue** |
| **0** | Centrum | Bar | Hotel | Coffee Shop |
| **1** | Nieuw-West | Soccer Field | Theater | Gym / Fitness Center |
| **2** | Noord | Plaza | Yoga Studio | Bed & Breakfast |
| **3** | Oost | Soccer Field | Hockey Field | Playground |
| **4** | West | Nightclub | Snack Place | Supermarket |
| **5** | Westpoort | Heliport | Harbor / Marina | Yoga Studio |
| **6** | Zuid | Bagel Shop | Office | Supermarket |
| **7** | Zuidoost | Farm | Harbor / Marina | Park |

Before we go ahead and add the number of venues per district to our data table, I thought it would be interesting to take a quick look at the most common venues found in each district, according to Foursquare. As can be seen in the table above (Table 2), the Centrum (Dutch for center, yes!) has bars, hotels, and coffee shops as the top three most common types of venues. This makes a lot of sense, given that the Centrum of Amsterdam is the biggest tourist hotspot, catering largely to people looking to drink, smoke, and stay the night. Taking quick looks at the data like this can be helpful, I find, to verify that our assumptions regarding the data are, so far, correct.

Another part of the data table that stands out is the Westpoort (West Haven) district, given that the first most common type of venue is listed as “heliport”. Having lived in Amsterdam for over a decade, this does not come as a surprise to me, as Westpoort is almost entirely a business district / industrial park that serves as the corporate headquarters for many, many companies operating in Amsterdam. As could be seen in Table 1, the population density for Westpoort is a mere 10 residents per square kilometer, so having heliports and harbors as the most common types of venues is not entirely surprising. What will be interesting in the analysis and results section of this paper is to see whether Westpoort will have a significant impact, given how different it is from the other seven districts.

**Table 3: Complete Data Table**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Neighborhood** | **Latitude** | **Longitude** | **Venues** | **Population Density** | **Safety Index** | **Safe** |
| **0** | Centrum | 52.37321 | 4.903712 | 88 | 13748 | 69 | 1 |
| **1** | Westpoort | 52.41095 | 4.803871 | 2 | 10 | 0 | 1 |
| **2** | Zuidoost | 52.30465 | 4.974994 | 4 | 4391 | 132 | 0 |
| **3** | West | 52.37611 | 4.86452 | 17 | 15252 | 79 | 1 |
| **4** | Nieuw-West | 52.36407 | 4.802676 | 10 | 4478 | 130 | 0 |
| **5** | Oost | 52.34981 | 4.956049 | 6 | 7635 | 80 | 1 |
| **6** | Zuid | 52.34172 | 4.86605 | 9 | 9349 | 56 | 1 |
| **7** | Noord | 52.388 | 4.917663 | 13 | 2269 | 112 | 0 |

Now that all of the data has been collected, we can gather it all together into one data table (Table 3). New additions are the “Venues” column, which shows the number of venues found via Foursquare in each district, as well as the “Safety Index” and “Safe” binary variable. Unsurprisingly, the number of venues is greatest in Centrum, as that is the tourist hotspot. West, Nieuw-West (New-West), and Noord (North) are all in the double digits, while very few venues were found for Westpoort, Zuidoost (Southeast), Oost (East), and Zuid (South). As stated previously, any neighborhood with a safety index value below 100 is considered “safe”, while anything above is considered “unsafe” by the government’s guidelines. Only three of the eight neighborhoods are considered unsafe: Zuidoost, Nieuw-West, and Noord.