Opening new businesses in low density areas in Amsterdam

Introduction:

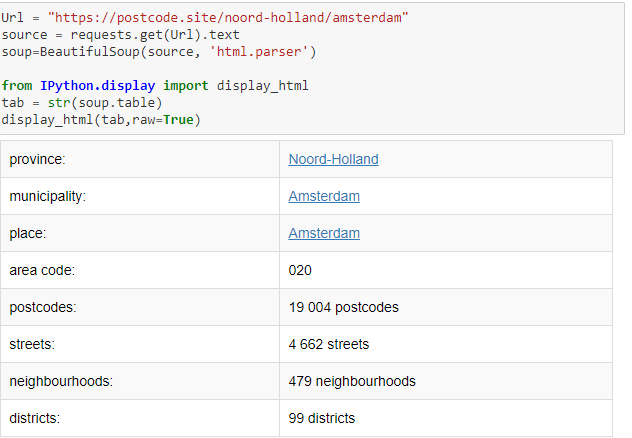
Amsterdam is a tourist attraction with many venues to choose from. My idea was for new entrepreneurs who would like to open a new business. Where should they open the business where they'll have low density in competition, but high density in population.

The used a few different sources for the location analysis.  '<https://postcode.site/noord-holland/amsterdam>' provided the costal codes for the addresses in Amsterdam, while <https://en.wikipedia.org/wiki/Boroughs_of_Amsterdam> provided the names of the districts and boroughs, that I would later cross-reference with the postal codes.

Finally, I would use Foursquare to locate the venues. I gave an example for coffee shops but any other business will work with this code. And used a ML algorithm (clustering) to determine where are the clustered and therefore where is the highest competition for our new business.

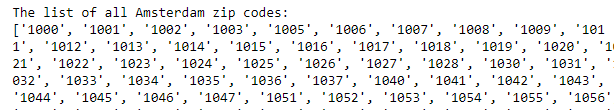
Pulling data:

First to get a rough estimation on the data set of how many neighborhood, districts and postal code I'll be working with

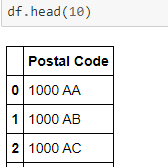


As you can see almost 100 districts, and 19 thousand postcodes.

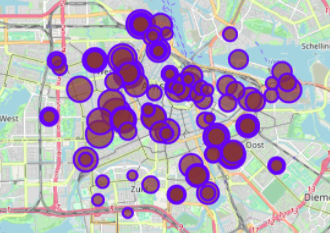
The list of postal code was producing only 4-digit numbers but I had to combine them with a 2-letter code as in the original postal codes.



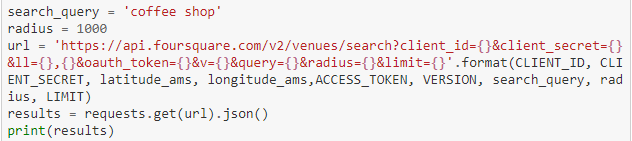
So the final result looked like this



I had 19,899 postal codes but with a lot of N/A's, so after cleaning the data I was left with 19,625.

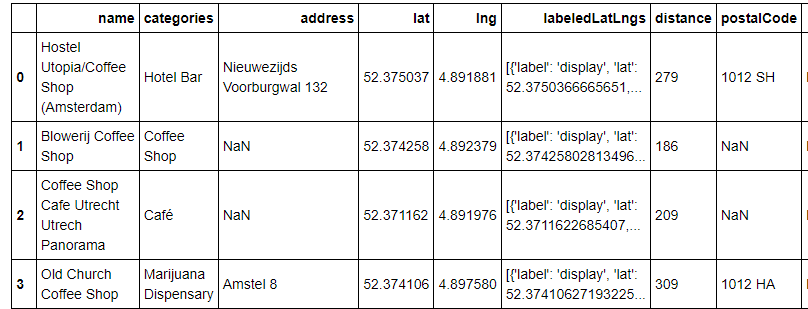


After looking at population density I turned to Foursquare to look at the businesses

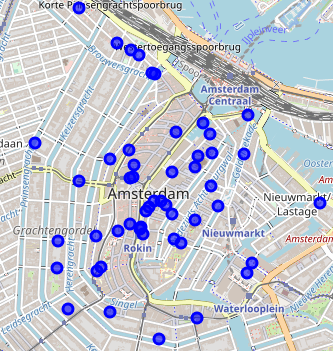


I used the search query "Coffee shop" but as I mentioned above, any other business type can be used.

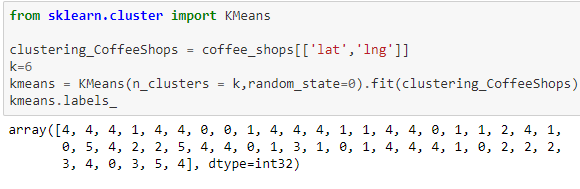
After cleaning the data this is a resulting table I got



Using the lat, lng columns I mapped the businesses across the city



And then I used K-Mean clustering the ground the data point together, so to better understand where is the centroids of those businesses.



And I inserted the clusters into the table and produced a new map color coding the clusters so to better understand the new groups.



Conclusions:

As you can see the blue dots are located on mostly the same line\street (Harrelemmenstraat). And therefore, all the are in the green hollow circle is with no competitors. Alost below the light green and above the teel dots we can see another hollow green circle the there is no coffee shops to compete with ours.

Implications:

This tool could service many entrepreneurs who are looking for low competition areas with high density of population