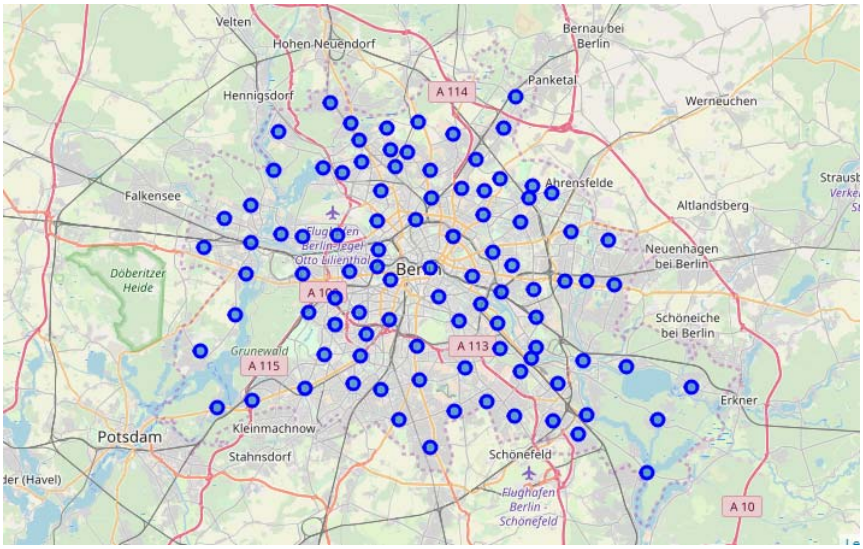
# *Berlin Neighborhoods*




***-Capstone project-***

***Nov, 2018***

***Author: Djordje Dajkovic***

# *Table of content*

## Contents

***Table of content*** ..... 1

***Introduction*** ..... 2

***Methodology*** ..... 3

***Results*** ..... 4

***Discussion*** ..... 5

***Conclusion*** ..... 5

# *Introduction*

Big Drugstore Chain from the USA wants to start a business in Germany. They want to start at least three Drugstores in Berlin but they need to choose the best locations for their Stores. After they made their SWAT analysis they decided to find locations with as much as possible population, not high competition around and not big office rent price. The datascientists job is to find such locations that fulfil that criteria.

First, data about Berlin boroughs and neigborhoods must be found, also, population and price data for boroughs and neighborhoods. Foursquare APIs will be used to find venues in neighborhoods and select those where is as little Drugstores as possible and filter by low prices.

# *Methodology*

Data about Berlin sources are collected from various Internet sources. Data about Boroughs, Neighborhoods are collected from Wikipedia (https://en.wikipedia.org/wiki/Boroughs_and_neighborhoods_of_Berlin) and prices for Office space rent are collected from following page: https://setting.io/blog/2016/06/13/startup-office-rent-map-berlin. Population and Prices data are from 2016.

After that geographical coordinates of Neighborhoods were necessary and they are obtained using Geocoder package. After this data obtained following map was generated using Folium package:

In order to get venues of each Neighborhood, Foursquare APIs were used and after data processing new dataset was created, containing 20 most common venues by Neighborhood. All Neighborhoods that contained Drugstores were eliminated from further analysis. This was done in order to have only Neighborhoods with very low competition in Drugstore business.

After that, two filters are applied to resulting dataset. First is to eliminate Neighborhoods with less than 50000 population and another to eliminate all Neighborhoods that have maximum rental price higher than Berlin median.

# *Results*

Result obtained by this research are following:

| Borough | Neigborhood | Population | Price(eur/m2) |
|---|---|---|---|
| Lichtenberg | Friedrichsfelde | 50010 | 8.5 |
| Lichtenberg | Neu-Hohenschönhausen | 53698 | 8.5 |
| Marzahn-Hellersdorf | Hellersdorf | 72602 | 10.0 |
| Marzahn-Hellersdorf | Marzahn | 102398 | 10.0 |
| Tempelhof-Schöneberg | Schöneberg | 116743 | 10.0 |
| Treptow-Köpenick | Köpenick | 59201 | 11.0 |
| Neukölln | Neukölln | 154127 | 11.0 |

On Berlin map it looks like this:

# *Discussion*

As seen on map, there are 7 Neighborhoods fulfill criteria for this business case. None of this Neighborhoods are close to city Center what was expected considering prices for rent of Office space. Positive thing is that 3 of Neighborhoods have population over 100000. On Business owner is to decide where their ventures will be positioned.

# *Conclusion*

After viewing results of this research it is clear that there is a lot of space in Berlin to develop successful Drugstore business.