The best spot for a (new) Bike-Café in Copenhagen

Introduction

1. Background

Bikes are more than a mean of transportation in Europe: they are a culture. That means they are part of the State's planning and infrastructure projects, and people ride them the whole year (no breaks during winter).

Cyclists, moreover, share a passion for coffee. The energy benefits of this drink and its culture attract the athletes as a gathering point before a ride, for coffee pauses during long riding sessions or in the end of one, to recover the breath and finally go home.

In the study "Bicycle Cities Index 2019" conducted by Coya, 90 cities from all around the world were considered, ranging from traditional cycling cities to lesser-known locations with improving bike infrastructure. There were six main evaluation categories: weather, percentage of bike usage, crime & safety, infrastructure, sharing and events. The higher the total score, the better. Copenhagen and Amsterdam appear as the first 2 best Medium-size cities to ride a bike.

With that in mind, I decided to analyze the city of Copenhagen and its districts to understand what would be the best place to open a Bike-Café, a Café with the facilities of a bike shop.

2. Target public

This project is of interest to those who share both passions for coffee and bikes – both athletes and commuters. Additionally, it is of interest to those who want to start a business.

Data sources and data cleaning

A sure way to understand if a place is bike-friendly or not is to look at the statistics regarding accidents involving cyclists or to see how many kilometers a city has of built cycle ways. Looking for such information, I found the "Bicycle Cities Index 2019", to gather information about what should be the target city to analyze in order to establish a Bike-Café.

With the good ratings pointing to Copenhagen, I decided to explore ways to collect information on the neighborhoods of the city. After some attempts of creating a data set with only the neighborhoods of the city, I found the Postnord website, gathering all the Postal Codes from Copenhagen.

As a standard Places API in the IBM Data Science Capstone Project, Foursquare was used to explore the districts of Copenhagen to understand the most common venues in each area of the city. Foursquare lacked a specific category (and important to this project): bike cafés. So due to time and developer account limitations I used Google Maps to query the Bike Cafés in Copenhagen.

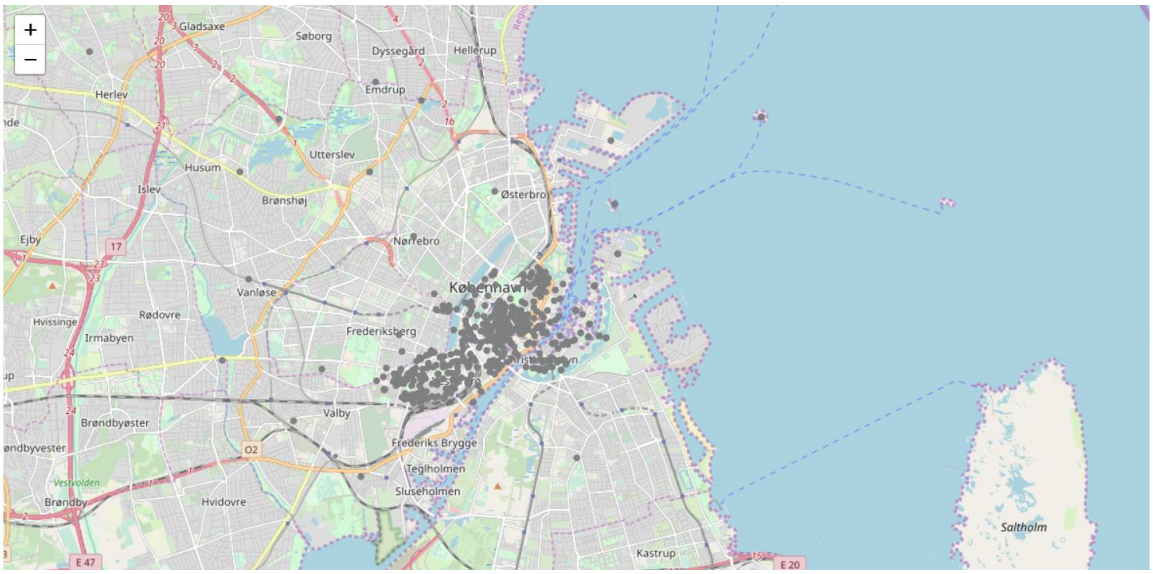
Finally, to give the investors an idea of rent prices in Copenhagen, I used this website to give an overview of the rent scene there.

Data cleaning was performed using NotePad++ to decode the text from the format UTF-8 to the format ASCII, and Excel to filter only postal codes from Copenhagen.

Methodology

The methodology used to complete this project can be divided in the following breakdown structure:

1. From the PostNord website, I have retrieved the information containing the Postal Codes for all the cities covered from the company. I then filtered only the ones corresponding to Copenhagen City and Region: that means, from 1001 to 3000 and 654 Postal Codes.



2. If you are working on a Data Science project, it is of utmost importance knowing how to convert general locations, such as addresses, into geographical coordinates. To succeed this task, I used geocoder.arcgis in Python to get the coordinates for each of the Postal Codes.

3. I have created a specific Notebook to query the venues in the surroundings of the Postal Codes using the Foursquare API, due to the daily limit of 950 calls for a Sandbox User. With the information I needed, I transformed the resulting data frame into a CSV file, so I would not have problems of exceeding daily limits in Foursquare Developer Portal every time I had to work in the project. The data set has a magnitude of (46459, 8). A snap of the data set you can see below:

```
[20]: print(copenhagen_venues.shape)
      copenhagen_venues.head()
```

(46459, 8)

[20]:

	District	PostalCode	Street Latitude	Street Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Kobenhavn K	1050	55.680453	12.58621	Det Kongelige Teater	55.679761	12.586075	Theater
1	Kobenhavn K	1050	55.680453	12.58621	Nyhavn	55.680247	12.589398	Harbor / Marina
2	Kobenhavn K	1050	55.680453	12.58621	Nebbiolo	55.681116	12.589191	Wine Bar
3	Kobenhavn K	1050	55.680453	12.58621	Kunsthal Charlottenborg	55.679926	12.588279	Art Gallery
4	Kobenhavn K	1050	55.680453	12.58621	Den Lille Fede	55.682199	12.585476	Mediterranean Restaurant

4. Back to work on my main Notebook, I used the one hot encoding technique to obtain the density of all the venues for each Postal Code in Copenhagen. After it, I grouped the sum of the venues for each District of the city, so I reduced my data set to a shape of (55, 236), with 235 unique venue's categories.

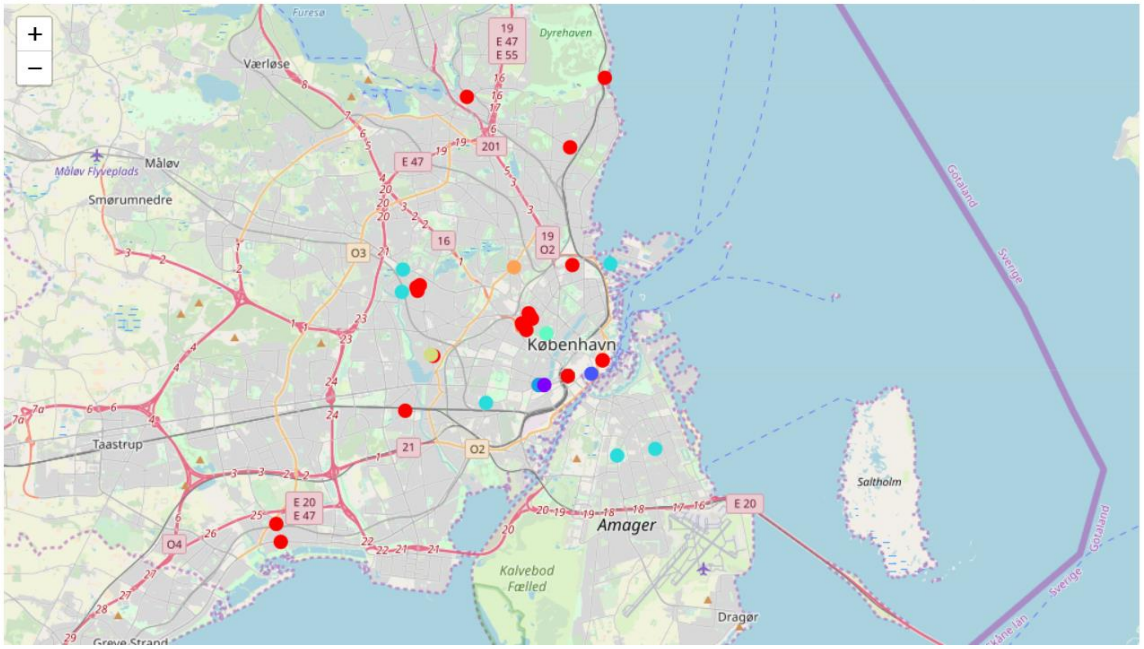
[7]:
```
# one hot encoding
copenhagen_onehot = pd.get_dummies(data[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
copenhagen_onehot['District'] = data['District']


# move neighborhood column to the first column
fixed_columns = [copenhagen_onehot.columns[-1]] + list(copenhagen_onehot.columns[:-1])
copenhagen_onehot = copenhagen_onehot[fixed_columns]

pd.set_option('display.max_columns', None)

copenhagen_onehot.head()
```

[7]:

	District	Advertising Agency	African Restaurant	American Restaurant	Antique Shop	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Australian Restaurant	BBQ Joint	Bagel Shop	Bakery	Bar	Basketball Court	Bath House
0	Kobenhavn K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Kobenhavn K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Kobenhavn K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Kobenhavn K	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
4	Kobenhavn K	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

5. Using the K-Nearest Neighbors, I have clustered the data set in 10 clusters. I decided to use this technique due to the simplicity of its construction and satisfactory processing time to find the children for each cluster.

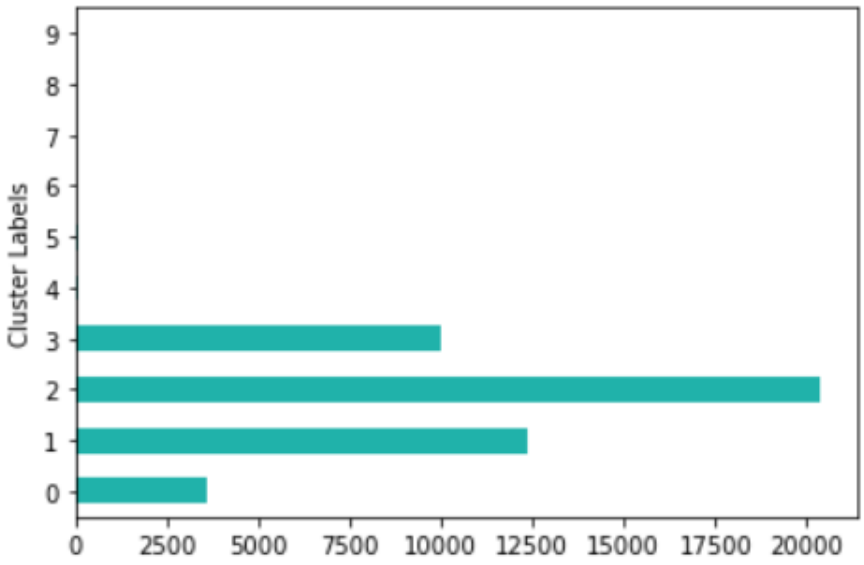
[26]:



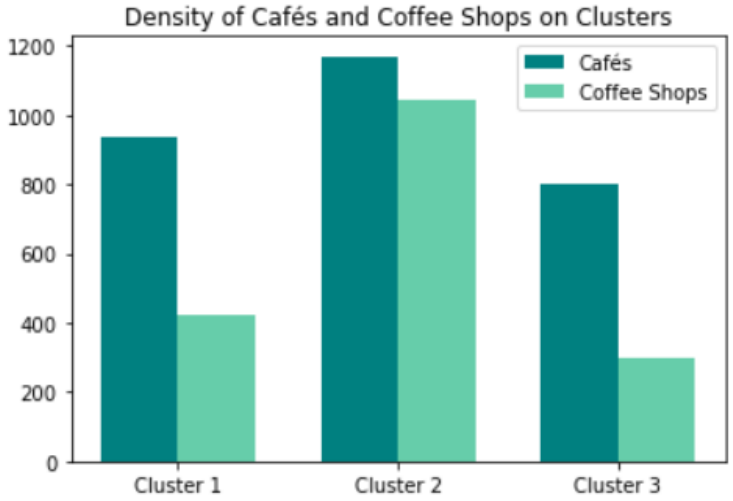
Discussion

1. The most populated clusters are the 1, 2 and 3, which correspond to the Districts of Kobenhavn V, Kobenhavn K and Frederiksberg C, respectively. I have focused my analysis in those 3 clusters.

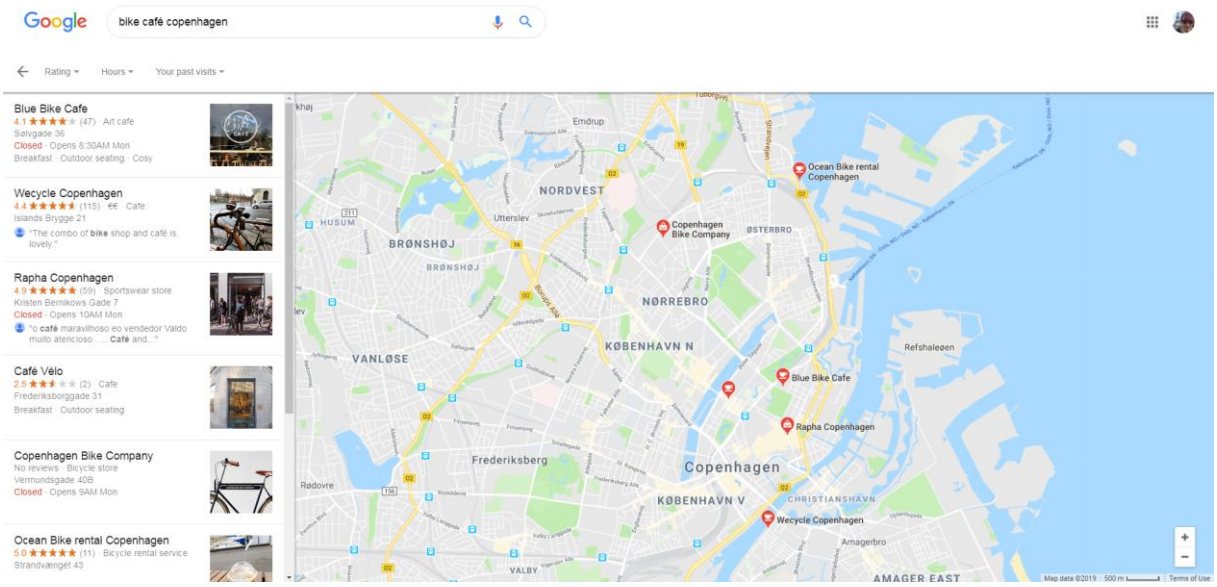
```
[44]: <matplotlib.axes._subplots.AxesSubplot at 0x7fba10b0fa20>
```



2. I plotted the Cafés and Coffee Shops for each of the 3 Clusters, and Frederiksberg C is the one with the least number for both venues and there is no Bike Shops according to Foursquare API. That means it is a great candidate to our Bike Café.



3. Furthermore, I Googled where the current Bike Cafés are in Copenhagen (Foursquare API lack this specific category) and it returned the following results:

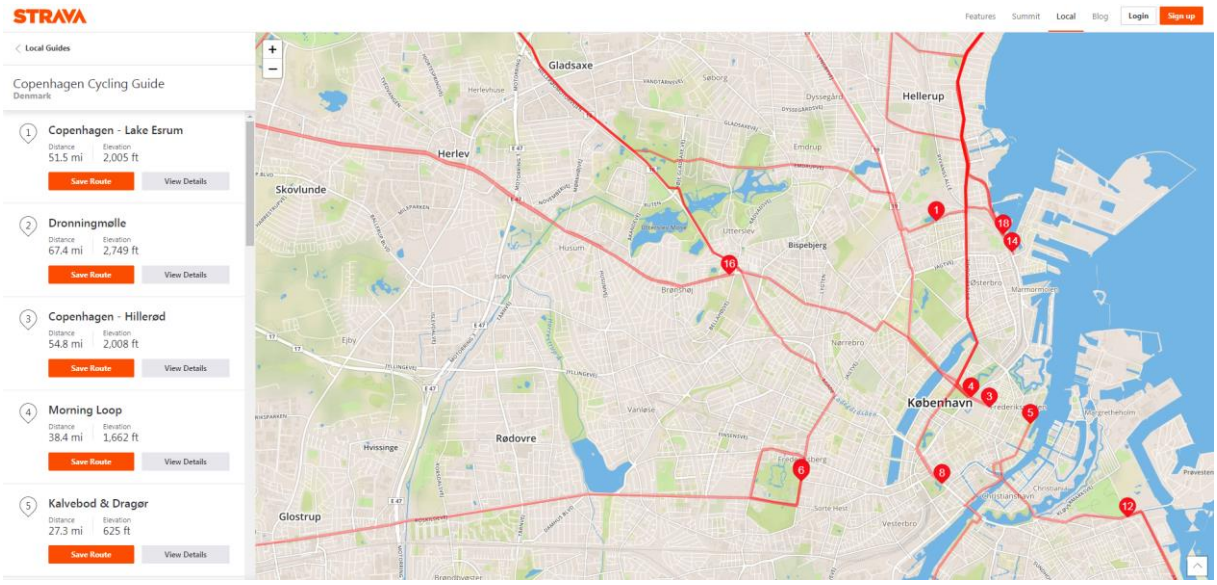


According to Google, there is no Bike Cafés in the District of Frederiksberg C.

The District of Frederiksberg is officially an independent Municipality of Copenhagen. It is a green area with important Education centers, such as a campus from the University of Copenhagen and the Copenhagen Business School.

Results

The target public of this project were people that love cycling and/or commute to work. A quick search at Strava for the cycling routes in Copenhagen shows Frederiksberg as a hot spot:



It means that it is not only easy to go for a training session from Frederiksberg, due to its location not directly in the center of Copenhagen, but it can also be used as a meeting point before/after a session. Moreover, this place could attract all the students and region commuters that gather in a Café to study and do network and meanwhile, would have the opportunity to repair their bikes and buy accessories. The posh profile of the District also corroborates the wish to open a Café with a bike shop facility: people with more money can afford expensive bikes that require special care.

Insights

A quick look at a local business rent website indicated that rents of spaces as big as 90 m^2 in Frederiksberg sum up to DKK 200.000 per year, which means roughly EUR 26.000 per year.

Conclusion

1. A future study could be to compare Copenhagen with other cities with potential in this direction, like Amsterdam – as the latter holds the second place as the best Medium sized city to cycle in the world, so to have a deeper understanding of which of the cities would be more interesting to open a Bike Café.

2. As for this analysis, a specific study on the financials needs to be further developed, to furnish consistent data regarding what would be the expected fixed costs and revenue for the business.