

# **Coursera Capstone**

## **IBM Applied Data Science Capstone Project – The Battle of the Neighbourhoods**

### **Acorn Technologies moving to Trondheim, Norway**

By: Edward Parkin

March 2021



## **The Battle of Neighbourhoods - Introduction**

### **Introduction 1.1 Background -**

A start-up technology company (Acorn Technologies) wants to set up headquarters and begin trading in Trondheim, Norway. The owner wants to evaluate all of the options available to him to minimise risk and build for the future. It is important to him that the new headquarters be situated near other firms in the same industry because he predicts that there will be close collaboration with these companies, sharing talent and establishing relationships with people. Additionally, he only has a limited budget to work with and recognises that this is an expensive city so would like to maximise the value he gets from the neighbourhood he decides on. Especially considering that cash flow tends to be the biggest challenge to a new business. Whilst these are his main priorities for the near future, he would also like to prepare for future growth and the ability to attract talented employees. This means he is interested in the facilities a neighbourhood has to offer, such as transport links, cafes, and shops.

### **1.2 Problem -**

The purpose of this project is to build an analysis of what the different neighbourhoods in Trondheim have to offer Acorn Technologies, with relevance to the three main areas the owner has highlighted as important to making the correct decision.

1. Proximity to NTNU university campuses in the city
2. Neighbourhood facilities available

This breakdown, into the two areas of interest, will allow the report to build a detailed analysis of each individually and then enable the owner to make a value judgement on the neighbourhoods, considering his weighting of their importance.

### **1.3 Location -**

Trondheim is a city on the northern coast of Norway situated next to Trondheim Fjord. It is a highly student-oriented city with 1/5th of its population being comprised of students and is home to the Norwegian University of Science and Technology (NTNU).

### **1.4 Interest -**

This report has been commissioned by the owner of Acorn Technologies and thus will likely be kept private but shared with whomever they decide will be able to help in making a decision of neighbourhood.

## The Battle of Neighbourhoods - Data

### Data Description 2.1 Data Sources

#### Neighbourhood Source

The data pertaining to the neighbourhoods of Trondheim will be taken from the url: <https://en.wikipedia.org/wiki/Lerkendal>. The information in this link is contained in a table towards the bottom, it includes the four boroughs the city is divided into with their respective neighbourhoods alongside. The data will have to be transformed as it needs other data appending to it.

#### Property price source

Upon reflection and the search for data, obtaining commercial property prices by neighbourhood for Trondheim is not going to be possible. The housing market in Trondheim is complex and very little data is compiled within one place to be able to scrape and have confidence it is accurate. Few properties are bought and sold in the city and its especially rare to get listings for anything other than apartments which would not be suitable for this business start-up. In this regard, an alternative is availability of transport in certain areas,

#### Foursquare API

The source of data to be taken from the Foursquare API includes:

- availability of restaurants and stores etc.

The client would like to have information to plan for the future which involves quality of life for future employees. For this, the Foursquare API is going to be used to analyse the number of services available in each borough.

### 2.2 Data Cleaning

In order to obtain the latitude and longitude data I used the python package geopy. The data fed into this package did not contain the postal codes of the locations so the data was fed in following the format: '{} , Trondheim, NORWAY'.format(Neighbourhood, Borough). The iteration of this list output five erroneous results which were obviously not neighbourhood coordinates in Trondheim. There were two districts in Denmark, one in Finland, one in Italy and the fifth in the USA. Noticing this, I went back to input these locations through the geopy package manually and replaced the latitudes and longitudes in the compiled list. If the list were longer and number of errors larger, it would be much more difficult to spot these results and fix them manually. Plotting the locations on a graph also helped to spot these though and had it been more difficult to fix manually, I would have devised another method to rectify the issue, perhaps spending an extra step finding the postcodes of the neighbourhoods.

## **The Battle of the Neighbourhoods - Methodology**

### **3.1 Methodology**

Following this, it was important to visualise the four different boroughs on the map more easily as this helped to identify the clustering of neighbourhoods in the city. I divided the data frame into four separate tables and plotted them onto the map with each in a different colour:

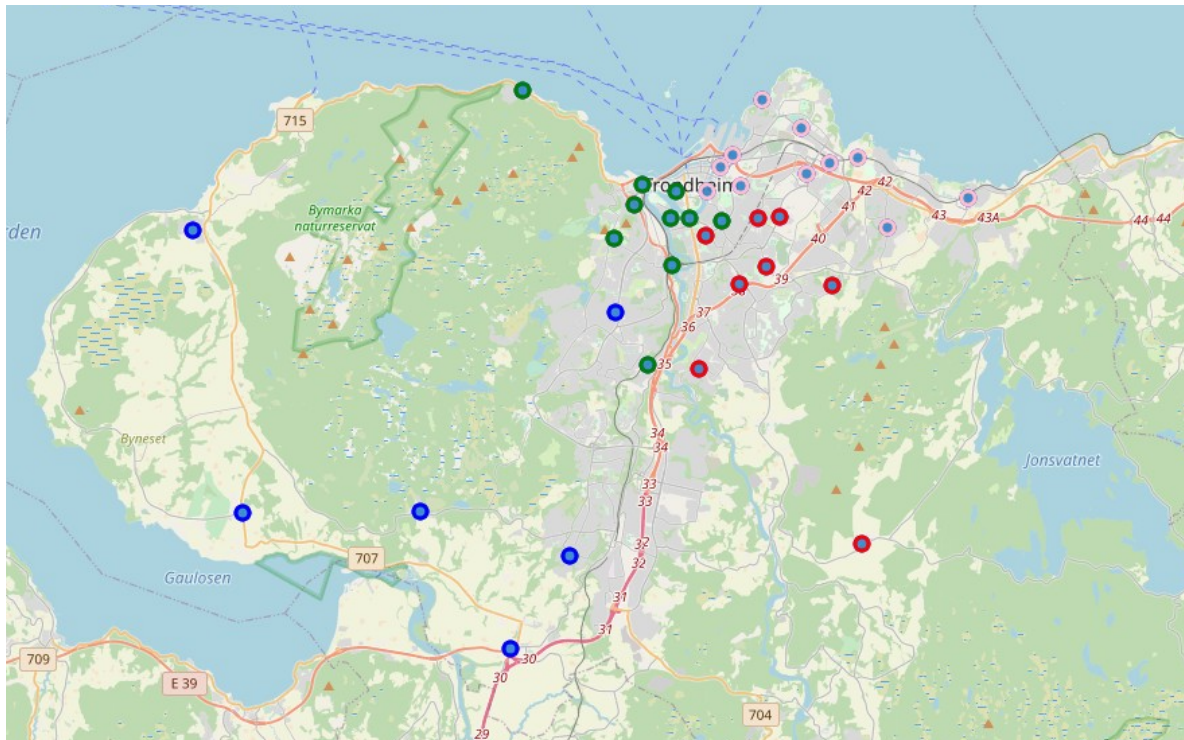


Figure 1- Trondheim with Neighbourhoods plotted

Østbyen – pink

Midtbyen – green

Lerkendal – red

Heimdal – blue

As we can see from the map, the green (Midtbyen) and pink (Østbyen) markers are more densely clustered over the city centre. The red (Lerkendal) and blue (Heimdal) markers are the boroughs which represent more rural and residential areas of the city. The CEO of Acorn technologies will be able to see from this that much of the economic activity, with access to products and services occurs within the Midtbyen and Østbyen boroughs. This is not to say that he should exclude the other two when making a decision, but he can come to the conclusion that if he chose to situate there, then his access to networks and future services for his employees will be restricted.



It was important to acorn technologies that they be able to recruit from the university which would involve attending career days to demonstrate to pupils what the company is involved in, the skills they are looking for and career paths. For this reason, ease of networking with students is a big priority, which includes accessibility to campuses. I plotted the two campuses onto the existing

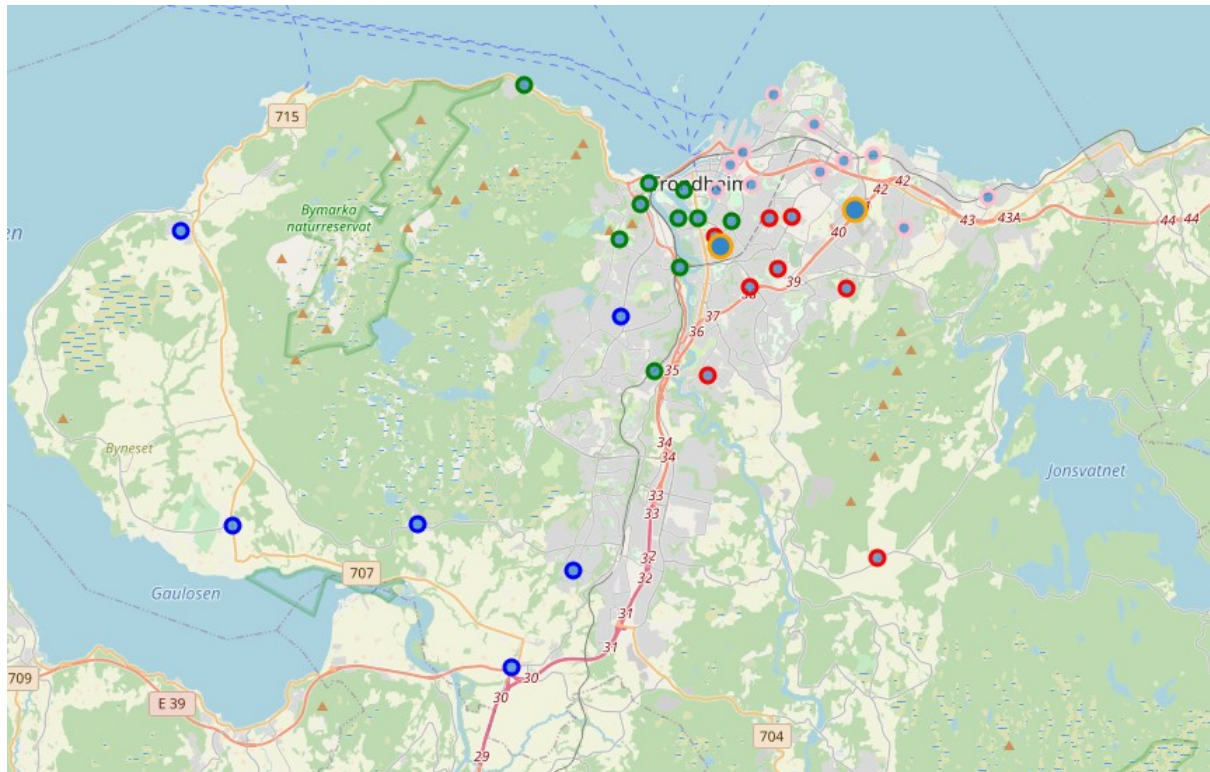


Figure 2 - Trondheim with Neighbourhoods and NTNU plotted

borough map of Trondheim in orange in order for him to visualise their locations in relation to the boroughs. In the python notebook, these pins are interactable so he can quickly see which are nearest and whereabouts in the city. This map also allows the user to zoom in to view the local area around the neighbourhoods. This is likely something that the owner of Acorn Technologies will do after using the data analysis to narrow down the few areas he wants to decide between.

Out[17]:

	Neighbourhood	Latitude	Longitude	VenueName	location.lat	location.lng	categories
0	Heimdal	63.401512	10.356507	Trondheim Klatresenter	63.434459	10.462260	Climbing Gym
1	Heimdal	63.401512	10.356507	Leangen Ishall	63.427669	10.465469	Stadium
2	Heimdal	63.401512	10.356507	Vinmonopolet (Valentinlyst)	63.425011	10.442527	Wine Shop
3	Heimdal	63.401512	10.356507	Plantasjen	63.431369	10.451830	Garden Center
4	Heimdal	63.401512	10.356507	IKEA	63.428764	10.473686	Furniture / Home Store
...	...	...	...	...	...	...	...
520	Østbyen	63.432918	10.453973	Kiwi	63.418663	10.464790	Grocery Store
521	Østbyen	63.432918	10.453973	SATS ELIXIA Valentinlyst SATS	63.424233	10.443001	Gym / Fitness Center
522	Østbyen	63.432918	10.453973	Mega	63.424333	10.441615	Grocery Store
523	Østbyen	63.432918	10.453973	Tregården Restaurant	63.431271	10.443668	Italian Restaurant
524	Østbyen	63.432918	10.453973	Coop Prix	63.433078	10.446825	Grocery Store

525 rows × 7 columns

Figure 3 - Foursquare API Venue Dataframe

I made a call to the Foursquare API to return venues with a radius of 1000 around the four boroughs of Trondheim. This returned 525 results with 12 unique categories included in the dataframe.

I then displayed the number of venues in this dataframe corresponding to each borough. As we can see, this reaffirms the finding from plotting the boroughs on the map, the two more densely clustered over the centre of the city have a greater number of venues to choose from. Additionally, we can conclude that the walking distance between these venues will be smaller because there are more of them in a smaller area. Quality of life for future employees, measured by access to venues within walking distance, should be higher in Midtbyen and Østbyen.

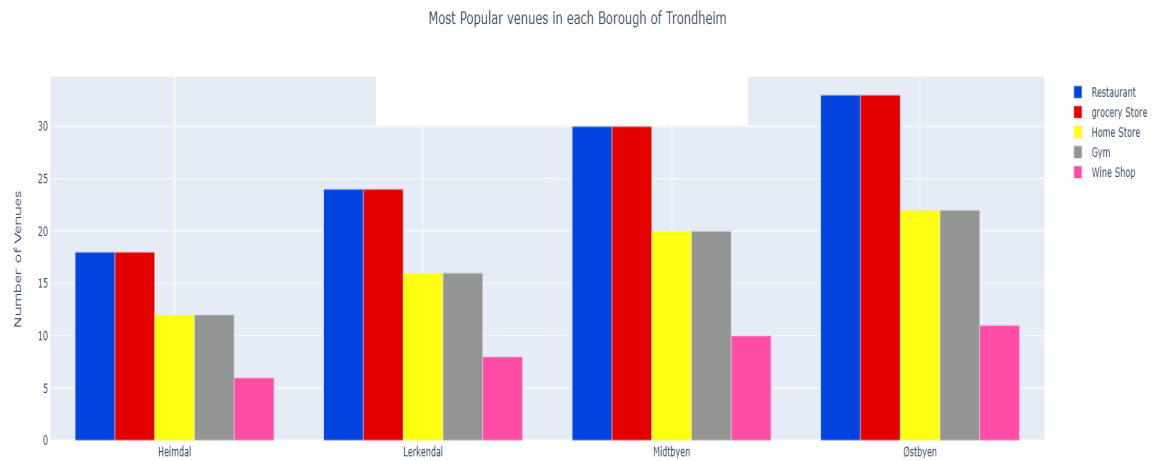


Figure 5 - Graph showing the five most popular venues in each Borough

I listed the most popular venues in order in the total venues dataframe to see what there was most of throughout the city. Then I found the values for each Borough. Figure 5 shows the five most popular venues in each Borough. We can see from this that all boroughs have a similar level of 'restaurants' and 'grocery stores', similarly, 'home stores' and 'gyms'. With 'wine shop' being the lowest in each. It is odd that there isn't more variation on this front but considering these are the most popular venues we can deduce that there are only a certain number needed in any given area to satisfy the local population.



## 4. Results

In this project I analysed the neighbourhoods of Trondheim and how they are situated in relation to one another. I also analysed their proximity to the two campuses of NTNU as a helpful guide to the owner of Acorn Technologies who wanted to recruit students from this university and would have to arrange for career days and networking opportunities. Finally, I used foursquare API to help him build a picture of which neighbourhood in Trondheim would be the most exciting to help retain these new employees.

## 5. Discussion

Examining the data I have collected, I would say it comes down to either a neighbourhood in Østbyen or one in Midtbyen. This, in my opinion is a matter of personal preference between the two. Midtbyen is closer to the main university campus in the west of the city, which would make it easier to bring students to see the operations taking place there, get a feel for what is happening and so may mean Acorn are better able to employ newer students. However, when it comes to retaining these students by the criteria Acorn have listed as important, Østbyen has a higher availability of the top five venues in the city. Personally, I would say that Midtbyen is the better choice because the difference in venue offerings in Østbyen is minor and it is also a more spread-out borough of the city. Additionally, the offerings of venues don't, to my mind, make or break whether a student would choose to work for the company. The major factor would be generating the excitement of what they could be a part of which would be through invitation to the premises.

## 6. Conclusion

Ideally, this data science study could've been improved with better information sources, I was able to provide a good insight into locations which would be good for Acorn Technologies based on the factors requested. However, this could've been improved with property prices by region which were not available in a large enough data set to use accurately.