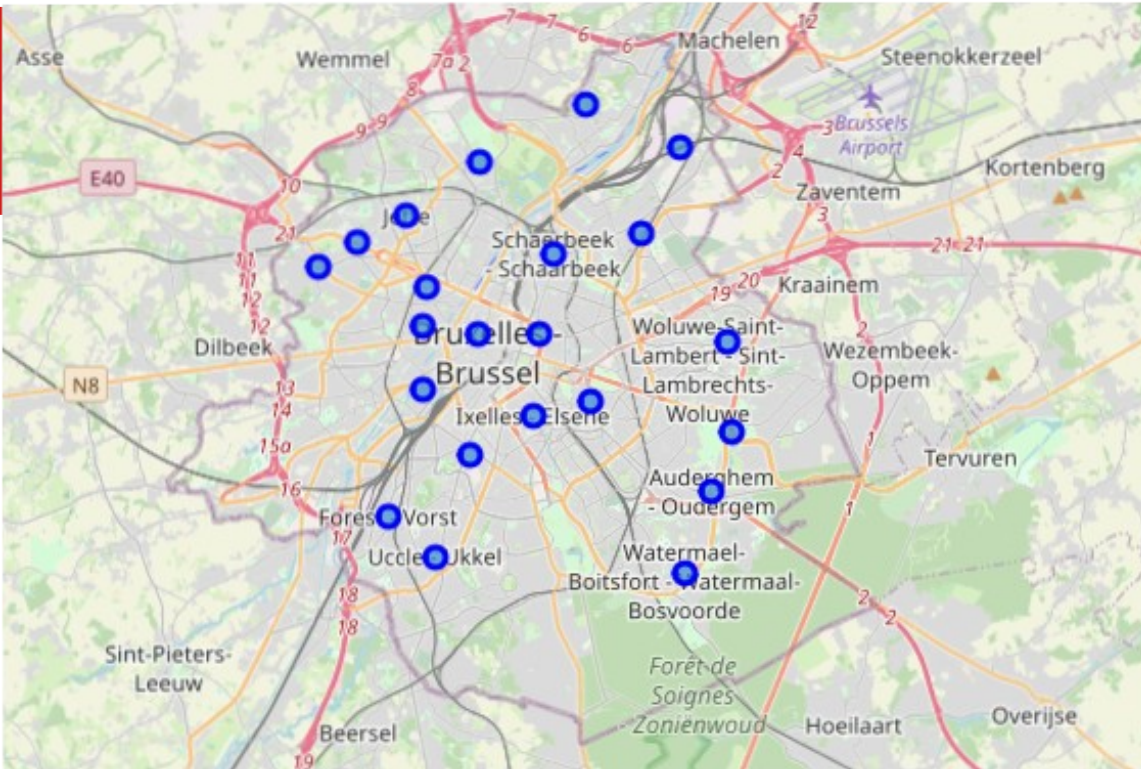


# Battle of Neighbourhoods of Brussels Capital Region



Nimmy Samson Attipetty  
01/04/2021

# Executive Summary

**The problem:** Choosing the best neighbourhood in Brussels for a partner of startup company who wishes to work in a co-working space.

**Important criteria:** Presence of farmers market, supermarket and French Restaurant along with co-working space.

**Data :** Location data from Foursquare, postcodes and municipality names of Brussels Capital Region and data for co-working space.

**Machine learning techniques:** K-means clustering, Elbow method, Silhouette method

**Result:** Etterbeek and Uccle could be the places most desired by the client.

# CONTENTS

Introduction: Business Problem

Data

Methodology

Analysis

Results and Discussion

Conclusion

# Introduction

- **Relocation is often challenging to a person.**
- **Unprepared move could make life unpleasant in the new location.**
- **Presence/Lack of venues similar to neighbourhood.**
- **Different preferences for everyone.**

## **The Business Problem**

**Pierre is a partner of an IT startup company.**

**Relocating to Brussels.**

**Requirements: Presence of Farmers market, Supermarket and French Restaurant.**

**Deciding Factor: Municipalities hosting co-working spaces.**

# The target audience

- Deal with analysis of municipalities in Brussels to find the best place where Pierre can relocate to, based on his personal preferences.
- Could also be extended to others with specific needs such as, an exchange/international student arriving in the city or for people who are searching for a new home within Brussels itself.

## Required Data

**As per the wishes of Pierre the following data should be retrieved for the analysis:**

- The postcodes and the names of municipalities in the Brussels Capital Region
- The list of venues in all of the municipalities in Brussels.
- The number of co-working spaces available in every municipality of Brussels Capital Region

# Data Sources

## Data for postcodes and municipalities

<https://www.ccc-ggc.brussels/sites/default/files/documents/graphics/bruxelles-en-cartes/tableaux-codes.jpg>

<http://pouvoirs-locaux.brussels/fr/acteurs/les-communes>

<https://www.brussels.be/newcomer>

## Location data

<https://opencagedata.com>

<https://foursquare.com>

## Data for co-working spaces

<https://data.gov.be/fr/dataset/310f1955-2326-47f4-96f3-5da9899e54bc>



# Preparing the data

## **Cleaning the data:**

Clearing Postal code values, splitting the postal codes for the municipality of Brussels (Ville de Bruxelles).

Inserting 4 new places for Ville de Bruxelles into the data set.

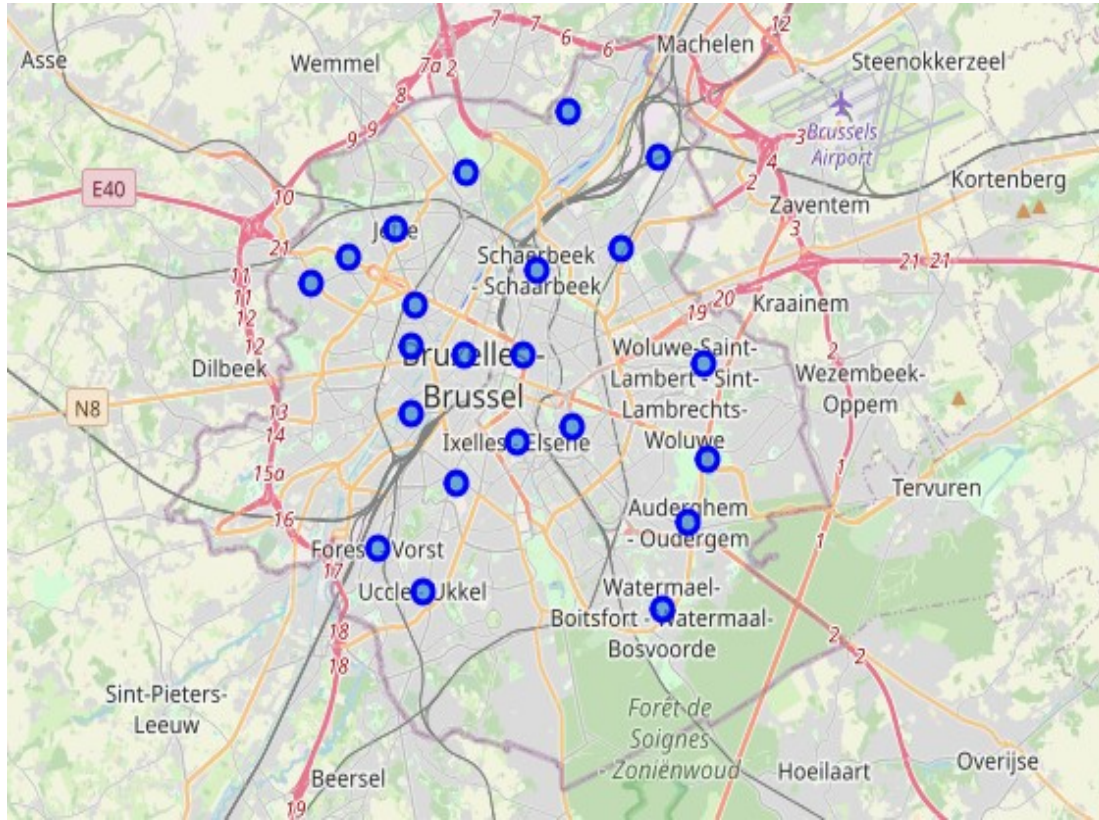
## **Obtaining the Location data:**

Using Opencage Geocoder to obtain latitudes and longitudes.

## **Obtaining the venues:**

Using the latitudes and longitudes along with postcodes and names to obtain venues and information related to it

# Map of Brussels



Mapped using the information obtained from Opencage Geocoder.

Brussels Capital Region consists of 19 municipalities/communes

Ville de Bruxelles contains 7 districts with overlapping postcodes.

Four districts added to the dataset for betterment of analysis.

# Using Foursquare API

```
LIMIT = 100 # limit of number of venues returned by Foursquare API
radius = 500 # define radius
url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll=
      CLIENT_ID,
      CLIENT_SECRET,
      VERSION,
      bru_latitude,
      bru_longitude,
      radius,
      LIMIT)
url # display URL
```

**Considers venues in 500 m radius and 100 top venues (user defined) at the latitude and longitude specified**

# Methodology

## **Exploratory Data Analysis**

- To obtain the feel of the data such as number of venues, types of venues
- Number of co-working spaces in each municipality.

## **One-hot encoding**

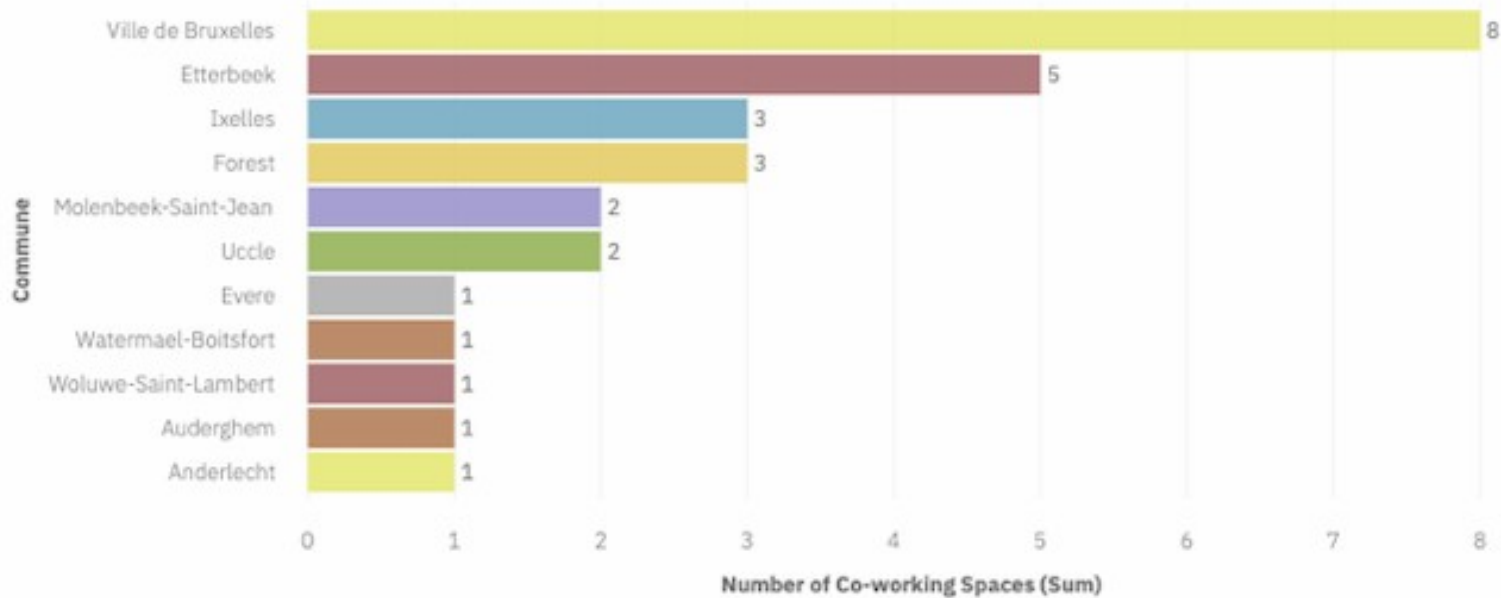
- Converting categorical data points to binary form.
- Returns 1 only if index value is present.

## **Clustering using machine learning techniques**

- Using a simple clustering algorithm called k-means.
- Optimal value of clusters determined by Elbow method/Silhouette method.

# Number of co-working spaces in municipalities

Number of Co-Working Spaces by Commune in the Brussels Capital Region



# Venues in each Municipality

## 3.1.2 Exploring the number of venues in each commune

```
print(brussels_venues.shape)
brussels_venues.tail()
```

(729, 7)

	Commune	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
724	Haren	50.890099	4.416321	Haren-Sud Gare / Haren-Zuid Station (MIVB / ST...	50.889810	4.414930	Bus Stop
725	Haren	50.890099	4.416321	Snack Verdun	50.891731	4.417738	Snack Place
726	Haren	50.890099	4.416321	Kortenbach (MIVB)	50.892075	4.417218	Bus Stop
727	Haren	50.890099	4.416321	Kortenbachplein / Place Cortenbach	50.892389	4.417450	Plaza
728	Haren	50.890099	4.416321	Dienstpad Schaarbeek Group R - Haren-Zuid	50.890915	4.410647	Trail

# One-hot encoding

	Commune	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
698	Haren	50.890099	4.416321	Snack Verdun	50.891731	4.417738	Snack Place
699	Haren	50.890099	4.416321	Kortenbach (MIVB)	50.892075	4.417218	Bus Stop
700	Haren	50.890099	4.416321	Kortenbachplein / Place Cortenbach	50.892389	4.417450	Plaza
701	Haren	50.890099	4.416321	Spoor / Voie 1 (Spoor 1)	50.888499	4.419596	Platform
702	Haren	50.890099	4.416321	Dienstpad Schaarbeek Group R - Haren-Zuid	50.890915	4.410647	Trail

**Further Analysis requires categorical data to be in binary form.  
One-hot encoding: 1 only if the index is true. Else returns 0**

## After one-hot encoding

	Commune	African Restaurant	Argentinian Restaurant	Art Gallery	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Bagel Shop	Bakery	Bar	Bath House	Beer Bar	Beer Store	Belgian Restaurant	Big Box Store	Bike Rental / Bike Share	Bistro
0	Anderlecht	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
1	Auderghem	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0
2	Berchem-Sainte-Agathe	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
3	Bruxelles	0	0	0	1	1	0	0	1	1	0	0	0	1	0	0	0
4	Etterbeek	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1



# Clustering using k means

## **k-means clustering.**

This algorithm iteratively tries to group the data into predefined k number of non-overlapping clusters with one dataset unique to one cluster only.

### **Optimal k-value for clustering - Using Elbow method**

Inertia is the sum of squared distances of samples to their closest cluster center.

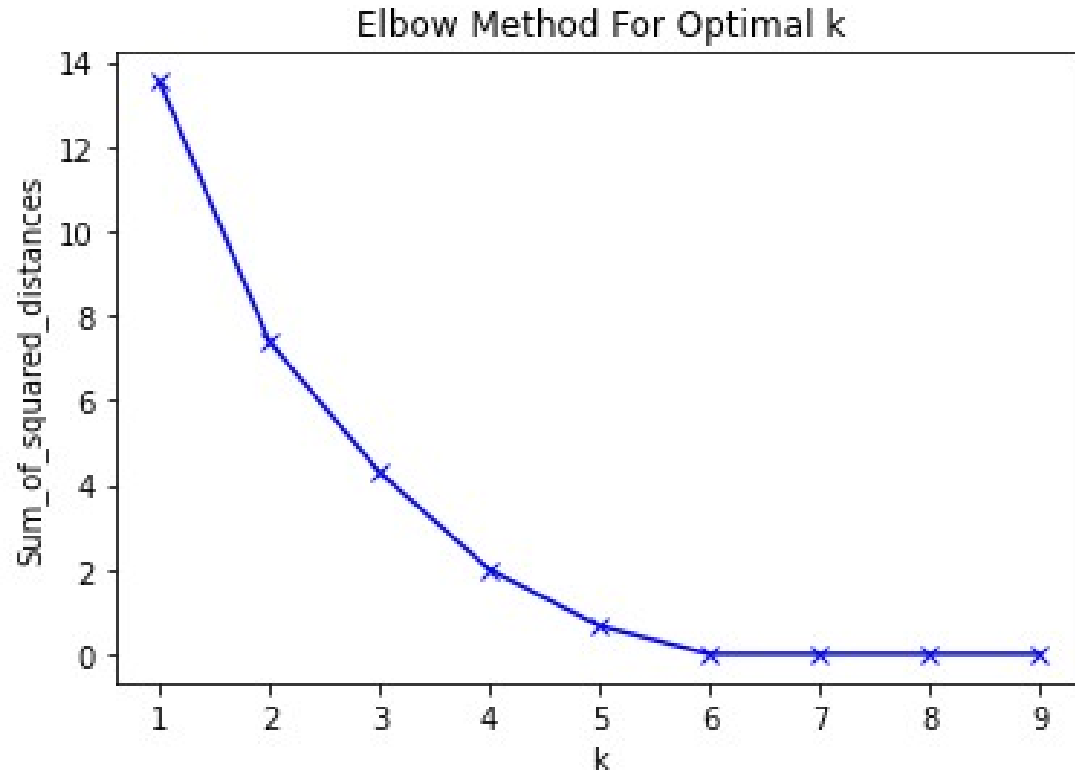
To determine the optimal number of clusters, we have to select the value of k at the “elbow” ie the point after which the distortion/inertia start decreasing in a linear fashion

### **Optimal k-value for clustering - Using Silhouette Score**

Silhouette value is superior measure than elbow method

measure the similarity an object is to its own cluster compared to other clusters.

# Optimal clusters using Elbow method



- Optimal value of k comes at  $k=5$
- This is verified using Silhouette Score as well

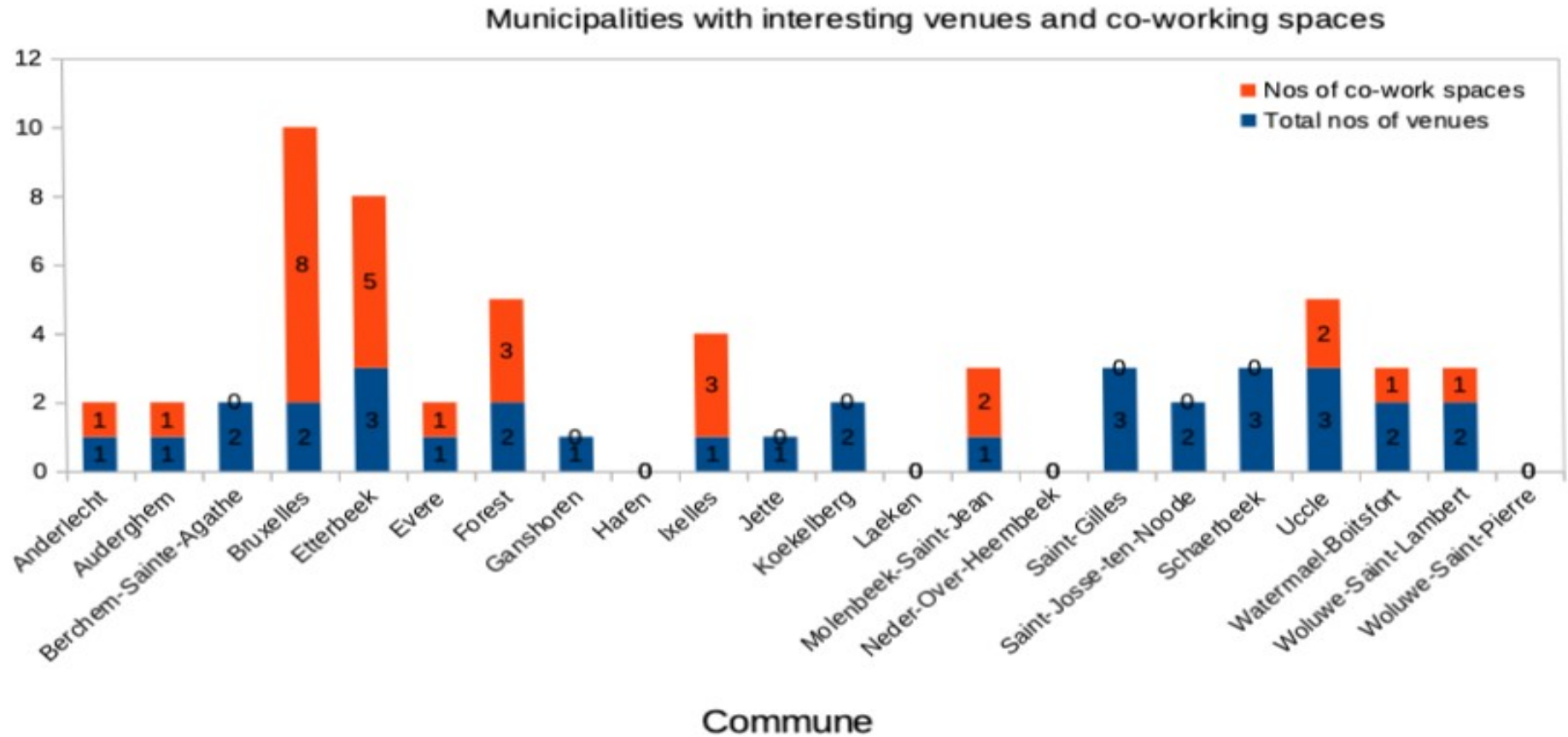
# Municipalities of interest

```
brussels_merged.loc[brussels_merged['Cluster Labels'] == 4, brussels_merged.columns[[1] + list(range(5, brussels_mer
```

	Commune	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Etterbeek	4	Ice Cream Shop	Seafood Restaurant	Massage Studio	Event Service	Modern European Restaurant	Doner Restaurant	Diner	Deli / Bodega	Concert Hall	Comedy Club
11	Saint-Gilles	4	Pizza Place	Friterie	Brasserie	Supermarket	French Restaurant	Brewery	Latin American Restaurant	Sandwich Place	Café	Farmers Market
13	Schaerbeek	4	Plaza	Convenience Store	Middle Eastern Restaurant	Snack Place	Bus Line	Spanish Restaurant	Coffee Shop	Supermarket	French Restaurant	Brasserie
14	Uccle	4	Food Service	Breakfast Spot	Thai Restaurant	Chinese Restaurant	Juice Bar	Swiss Restaurant	Plaza	Supermarket	French Restaurant	Theme Restaurant

Do they hosts co-working spaces as well?

# Results and Discussion



# Ideal locations with co-working spaces

<b>Ideal Locations (3 venues)</b>	<b>Secondary options (2 venues)</b>
Etterbeek	Forest
Uccle	Bruxelles
	Watermael Boistfort
	Woluwe Saint Lambert

Etterbeek (5 co-working spaces) and Uccle (1 co-working space) -Determining factor??  
Apartment rents and apartment availability  
Slightly higher rents in Uccle.

# Conclusion

- **Etterbeek and Uccle are found to be good locations for Pierre.**
- **However several other factors also influence ‘the definition of ideal location’.**
- **Availability of apartments is a major factor.**
- **Brussels region has highest rents in all of Belgium.**
- **Apartment rents are above 1000 euros in Brussels area.**
- **If unavailability of apartments arise, secondary options are to be considered.**