Moving to Dublin, Ireland

### Coursera Python Capstone | Michael Culligan

# Introduction / business problem

Dublin is the capital city of Ireland, and by far the largest city in the country. Over the past two decades Dublin has emerged as a hub for tech companies, with the European Headquarters of Google, Facebook, Twitter, LinkedIn and many others based in the [Silicon Docks](https://en.wikipedia.org/wiki/Silicon_Docks) area of the city. Thus, it is not surprising that there are many Data Science positions available in Dublin.

As someone beginning their career as a Data Analyst in Ireland, my next role is likely to be in Dublin and I will need to relocate to the city. Dublin City and the surrounding area is divided into many small neighbourhoods, called Electoral Districts (EDs). To give myself the best quality of life, I wish to live in the neighbourhood that is best suited to my needs. I have created a list of criteria which my new neighbourhood needs to satisfy:

* Dublin City is surrounded by a motorway (highway) ring-road to the north, west and south called the M50. I wish to live inside this ring-road so that the city centre is easily reachable from my new neighbourhood.
* I do not own a car, so it is essential that I am within walking distance of a supermarket.
* Getting exercise and spending time outdoors is important. So, I wish to live close to a park and a gym.
* Dublin's buses can be overcrowded and get delayed by rush hour traffic. Therefore, I wish to live close to a train or tram station.
* I enjoy coffee, so I'd like to have a Café in my neighbourhood.
* I'd like to have a nice bar or pub in my neighbourhood.
* I enjoy going to see the latest movies, so I'd like to live close to a cinema (movie theatre).
* My favourite two cuisines are Italian food and Indian food. I'd like my neighbourhood to offer both cuisines.

Ideally, I'd like to live near the water, as it more scenic and tranquil. Dublin is a maritime city and has canals. I would also prefer to have a sports club in my neighbourhood.

In addition, there are a few amenities which I don't want in my new neighbourhood. I will not live near an airport, a nightclub, or a casino, as these will be too noisy at night.

Using Foursquare location data, I will identify the Dublin neighbourhoods which best satisfy my criteria. I will then cluster these remaining neighbourhoods and look at the most common amenities in each cluster. I will select which cluster seems most attractive to me, and then choose which neighbourhood in that cluster I'd most like to live in.

My analysis will identify the most attractive and liveable neighbourhoods in Dublin City. This will assist me and other like-minded people to know which areas to search for accommodation in when moving to Dublin. Let the Battle of the Neighbourhoods begin!

# DATA

I will be primarily using two datasets in this project

#### 1. Dublin Electoral District Data

County Dublin, in which Dublin City is located, is divided into 353 smaller areas or neighbourhoods which are called Electoral Districts (EDs). The website [townlands.ie](https://www.townlands.ie/page/download/) provides spatial data for Electoral Districts, and is derived from OpenStreetMap. The GeoJson version of the dataset will be used in a similar manner as the 'newyork\_data.json' dataset was used in the Manhattan Neighbourhood's project. This dataset will allow for Folium Maps showing clusters of neighbourhoods and their location relative to each other and the city centre. The longitude and latitude fields will enable the identification of neighbourhoods outside of the city's surrounding ring-road. There are 6 fields of interest in this dataset:

* geometry.coordinates: This contains a list of the co-ordinates (latitude and longitude) which define the boundary of the neighbourhood.
* properties.LATITUDE: The latitude of the centre point of the neighbourhood.
* properties.LONGITUDE: The longitude of the centre point of the neighbourhood.
* properties.NAME\_TAG: The name of the neighbourhood.
* 'properties.CO\_NAME': The county in which the neighbourhood is located. Since the dataset covers the entire country of Ireland, this field is needed so that only neighbourhoods in Dublin are included in the analysis.

#### 2. Foursquare Location data

Foursquare location data will be used to identify the amenities (venue) in each Dublin neighbourhood. It is important to be aware that Foursquare does not contain every business or amenity in Dublin but does provide a comprehensive list of facilities. On initial data exploration, it was found that there are 272 unique categories of amenity in County Dublin. Some of these categories are very similar to each other (e.g. Indian Restaurant and North Indian Restaurant) so some may need to be grouped manually. This will give each neighbourhood the best chance of achieving all the criteria outlined in the Introduction section. The Foursquare location data will also be used to cluster the neighbourhoods which satisfy the criteria, and to identify which neighbourhood best suits my interests and needs. There are 4 fields which are required about each amenity:

* name: The name of the business or amenity
* lat: The latitude of the amenity
* lng: The longitude of the amenity
* categories: The category of the amenity

Foursquare also offers reviews of venues which would be desirable knowledge to have when choosing a neighbourhood. This will not be included in this project as it would greatly increase the complexity of the project.

# Methodology

I will find my most suitable Dublin neighbourhood by taking the following steps:

* Identifying all electoral districts in County Dublin
* Filtering by location to only include those inside the M50 ring-road
* Using Foursquare data to find all venues and amenities within 1km of the centre of each neighbourhood
* Using the venue categories, find the neighbourhoods which best satisfy my criteria
* Group the remaining neighbourhoods into clusters using k-means unsupervised learning algorithm.
* Identifying the cluster or neighbourhood which best suits me by looking at most common venues and criteria satisfied.

# analysis

## Condition 1: Inside the M50

The first condition specified in the introduction section is that the neighbourhoods must be located inside the M50 Ring-Road. Here, the neighbourhoods are filtered so that all neighbourhoods north, west, or south of the most extreme co-ordinates of the ring road are removed. Thus, only neighbourhoods close to central Dublin are included. The map below shows all of County Dublin. It highlights in dark green all the neighbourhoods which are partially or completely inside the M50. In total, there are 276 neighbourhoods which satisfy the condition.

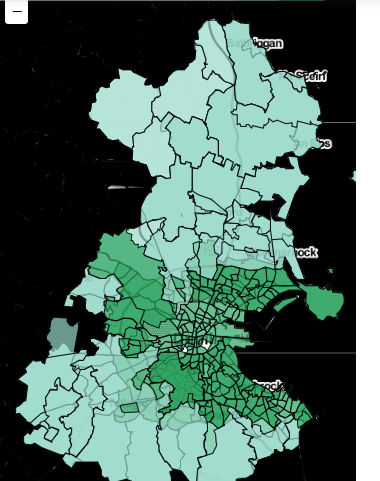


Figure 1: Neighbourhoods in Dark Green are inside M50

## Foursquare location data

For each neighbourhood, we find up to 100 different venues or businesses located within 1km of the centre of the neighbourhood. We take 4 pieces of data about each venue:

* Venue
* Venue category
* Venue latitude
* Venue longitude

Perhaps unsurprisingly in Ireland, the most common amenity is a pub, which appears 798 times. It is important to point out than an amenity can be attributed to more than one neighbourhood, especially in the case of small, central areas. Overall, there are 9186 amenities attributed to the 276 neighbourhoods. These amenities fall into 255 unique categories. A list of the most common occurring amenities is provided in Figure 2.

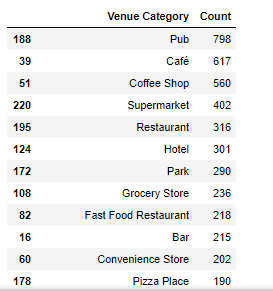


Figure 2: Dublin's most common Venues

## Criteria

As well as a neighbourhood inside the ring-road, I had 9 further criteria. These need to be located within 1km of the centre of the neighbourhood.

* Supermarket
* Park
* Gym
* Train or Tram Station
* Café
* Bar
* Movie Theatre
* Italian Restaurant
* Indian Restaurant

I will not live near an airport, a nightclub, or a casino. Neighbourhoods with these venues are removed.

There are two optional extra conditions. These are having sports facilities in the neighbourhood, and being close to the sea, river, or canal.

9 neighbourhoods satisfied all 9 criteria, with 2 of these also satisfying the optional extra conditions (Rathmines West B and Rathmines East D). In total, 23 neighbourhoods satisfied at least 8 of the 9 criteria, and the search for my most suitable Dublin neighbourhood will be narrowed down to these 23.

The list of these 23 neighbourhoods is given, and they are shown in blue on the map in Figure 3. Between them, they contain 1666 amenities. 4 of the 23 are on the north side of the city, and many of the neighbourhoods appear to form clusters. Those central neighbourhoods in white contain an amenity to avoid, such as a nightclub or arcade, and are excluded from further analysis.

The 23 neighbourhoods are:

* Ballinteer-Meadowbroads
* Botanic B
* Botanic C
* Clontarf West D
* Dun Laoghaire-East Central
* Dun Laoghaire-Glenageary
* Dun Laoghaire-West Central
* Dundrum
* Dundrum-Kilmacud
* Dundrum-Sweetmount
* Dundrum-Taney
* Pembroke East
* Pembroke East B
* Rathmines East D
* Rathmines West A
* Rathmines West B
* Rathmines West C
* Rathmines West D
* Rathmines West F
* Rathmines and Rathgar East
* Rathmines and Rathgar West
* Rotunda A
* Stillorgan-Kilmacud

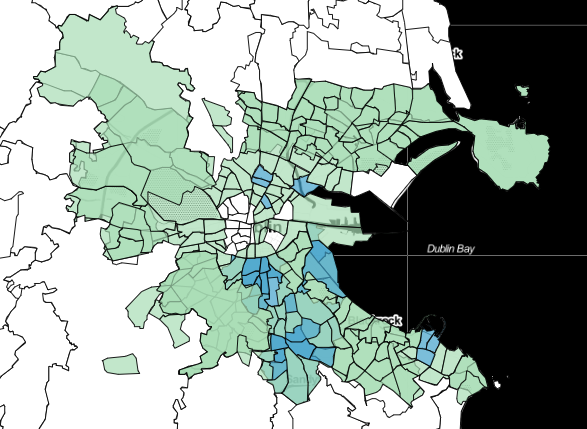


Figure 3: Neighbourhoods by criteria satisfied

## Clustering

The proportion of each type of amenity in each of the 23 neighbourhoods is used to cluster the neighbourhoods. The clustering algorithm used is k-means clustering, an unsupervised machine learning algorithm. From experimentation, I found that 6 was the optimal number of clusters. The locations and colour of each cluster are shown on the map in Figure 4 (see findings section)

# Findings

## Clusters

The clusters are broken down by colour and shown in Figure 4.

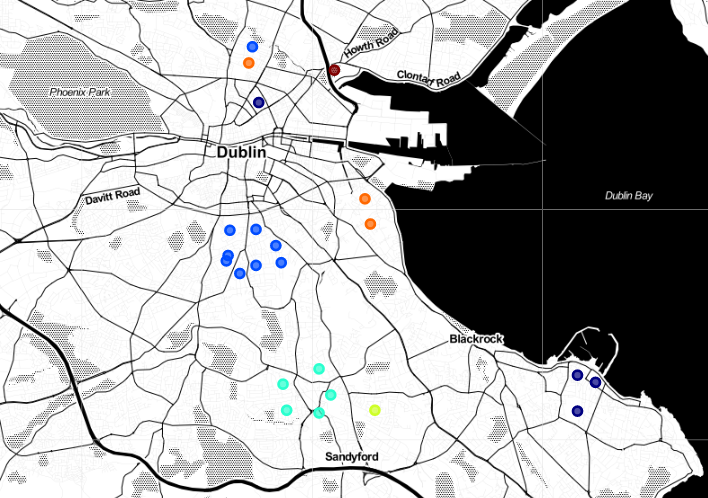


Figure 4: Dublin's Clusters

Many of the neighbourhoods which are located close to each other have been clustered together, which is not surprising. 4 clear clusters have emerged in the southside of the city. It is interesting that the 4 neighbourhoods on the northside of the city have been assigned to 4 different clusters. The 6 clusters can be roughly defined as:

* **Dark Blue:** Dún Laoghaire (and Rotunda)
* **Blue:** Rathmines (and Botanic B)
* **Teal:** Dundrum
* **Yellow:** Stillorgan
* **Orange:** Pembroke (and Botanic C)
* **Maroon:** West Clontarf

Let’s look at each cluster in more detail and discover their distinctive traits.

## Dún Laoghaire

Dún Laoghaire (pronounced Done-Leery) is a suburban coastal town, so it is unusual that is would be paired with the central neighbourhood of Rotunda A. These neighbourhoods have a very high number of cafés and coffee shops.



Figure 5: Dún Laoghaire's Most Common Venues

## Rathmines

The villagelike, inner suburb of Rathmines centres on a cluster of upscale gastropubs, global bistros, restaurants, and coffee shops, which dominate the most common venues list. It is for this reason that the northside neighbourhood Botanic B (named after the nearby National Botanic Gardens) is clustered with these 8 Rathmines districts.



Figure 6: Rathmines's most common venues

## Dundrum

Dundrum is an outer suburb on the southside of Dublin City. Home to Dundrum Town Centre, Ireland's largest shopping mall, it has a higher proportion of Clothing Stores, Department Stores and Pharmacies, while Cafés and Coffee Shops are also common.



Figure 7: Dundrum's most common venues

## Stillorgan

Stillorgan is a suburban area of Dublin City. This neighbourhood has a higher proportion of convenience stores and supermarkets.

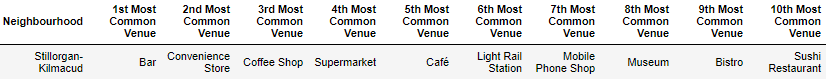


Figure 8: Stillorgan's most common venues

## Pembroke and Botanic

The Pembroke East and Pembroke East B neighbourhoods are two of Ireland's most upscale, expensive neighbourhoods, and have a higher proportion of hotels and pubs than the other neighbourhoods on this list. Therefore, Botanic C, located many kilometres away to the north east, has been clustered with them. It is interesting that the neighbouring areas of Botanic B and Botanic C are in different clusters.

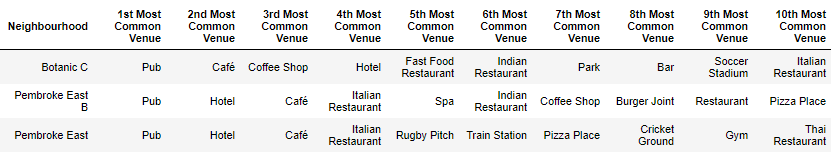


Figure 9: Pembroke / Botanic's most common venues

## Clontarf West

Clontarf West, located to the north east of Dublin city centre, has a high proportion of Cafés and Convenience Stores when compared to the other clusters.

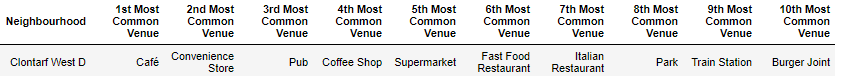


Figure 10: Clontarf West's most common venues

## Criteria in each neighbourhood

In this table, the number of venues which fall into one of the categories given in the introduction section, as well as the number of waterfronts and sports facilities in each neighbourhood, are listed. The neighbourhoods are in descending order, by number of venues which satisfy the criteria.

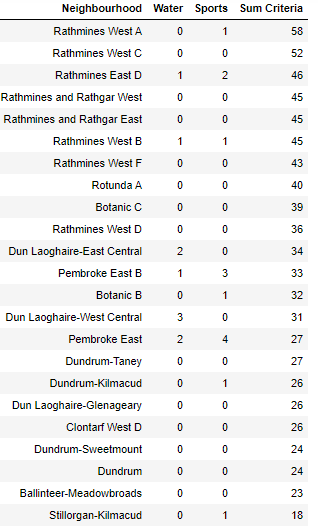


Figure 11: How many venues satisfy criteria

# Conclusion

Rathmines, Dún Laoghaire and Pembroke/Botanic seem to offer the best selection of Bars, Cafes and Restaurants *(looking at Figures 5 to 10)*.They appear to be the most favourable clusters. Dún Laoghaire offers a seaside lifestyle, while the other clusters border Dublin's Grand Canal and have a closer proximity to the city centre.

When considering Figures 5 to 10, together with Figure 11, the winning neighbourhood is Rathmines East D! The area 42 venues and amenities which satisfy the criteria, as well as a waterfront and sports facilities. Without question, Rathmines is the most suitable cluster, with the top 8 neighbourhoods in this list all from this cluster. A map of the neighbourhood, with pins marking each venue, is given in Figure 12.

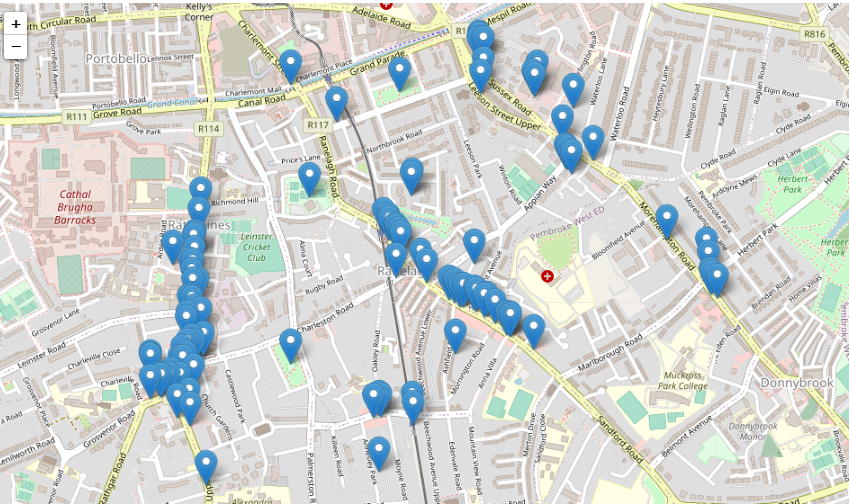


Figure 12: Rathmines East, Dublin's best neighbourhood (according to this report)

The villagelike nature, tram station, green spaces, and wide array of eateries and bars led to Rathmines, and specifically the Rathmines East D neighbourhood, emerging as the most suitable neighbourhood for me in Dublin City.

While not included in this analysis, the clusters which emerged as the most desirable neighbourhoods also have some of the highest average housing prices, in a city which already has one of the most expensive rental markets in Europe. In a future phase of this project, it would be worthwhile to include rent price data, and to try and strike a balance between affordable cost of housing and having a great range of amenities on your doorstep.