Average house and venues per neighbourhood in Dublin, Ireland

Table of Contents

[1. Background 2](#_Toc44857854)

[2. Data Description 2](#_Toc44857855)

[2.1 Data Sources 2](#_Toc44857856)

[2.2 Data Cleaning & Preparation 2](#_Toc44857857)

[3. Methodology 3](#_Toc44857858)

[4. Conclusion 4](#_Toc44857859)

[5. Reference 5](#_Toc44857860)

## 1. Background

Dublin is the capital and largest city of Ireland. The population of Dublin is approximately 1.4 million and the population density is 4,811/km2. The city is divided into 22 districts that are enumerated from 1 to 24, the south side comprises the even-number districts, which are from 2 to 24 and 6W, and the north part has the odd-number districts, which are from 1 to 17. [1]

According to Mercer’s cost of living survey, **Dublin is the most expensive city for expatriates in the Eurozone and the 46th most expensive country out of 209 cities.** The cost of rental accommodation was named one of the main reasons for the high cost of living. [2]

The city residents will naturally prefer regions in which the cost of renting is lower and where the neighbourhood most fit their lifestyle. From the investor perspective, it’s expected that they also prefer districts where the real estate cost is lower and the kind of business, they want to create is less intense.

Typically, the cost of renting is a small percentage of the house market price, therefore, analysing the average house price per district can indicate the average rental price.

Taking all of that into account**, the project aim to develop a map and an information chart informing the average house price and the most common venues per district** to assist investors and residents in making their decisions either on where to invest or to live.

## 2. Data Description

### 2.1 Data Sources

The data used to evaluate the problem is listed below:

* The average house price data was obtained from the **2019, 2018 and 2017 Residential Property Price Register** in .csv format, which was built by the Property Services Regulatory Authority (PRSA), this data contains the following information: [3]
  + Date of Sale
  + Address
  + Postal Code
  + County
  + Price (€)
  + Not Full Market Price
  + VAT Exclusive
  + Description of Property
  + Property Size Description
* The coordinates for each county will be scrapped from the **GeoNames, Ireland – Postal Codes** website [4]
* It was used **Foursquare API** to acquire the most common venues of each county of Dublin [5].

### 2.2 Data Cleaning & Preparation

Data downloaded and scraped from different sources were combined into one table. Rows that had No Answer (NA) were removed from the dataset. I decided to use only the house prices from 2017 to 2019 because the price has not significantly changed during those years, so as to minimise the impact of inflation on the final result.

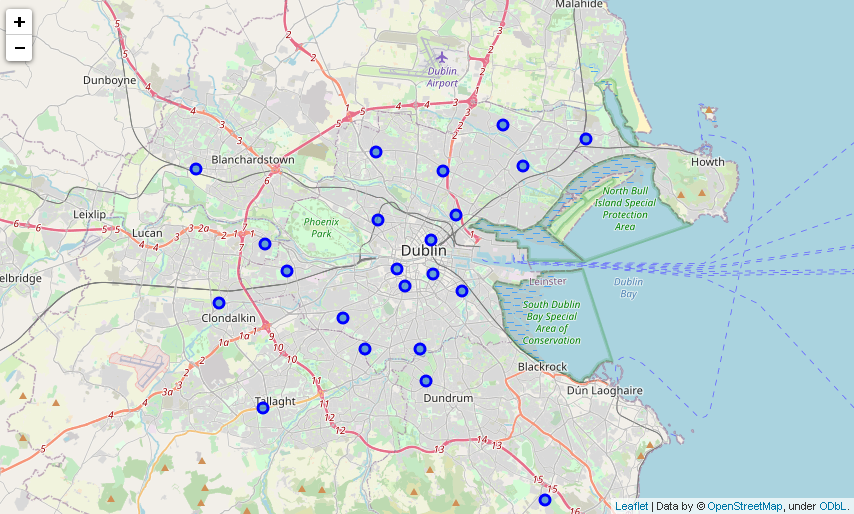
Moreover, the following columns obtained from scrapped data were removed as they are not relevant for the analysis: Unnamed: 0, Country, Admin1, Admin2, Admin3. From the downloaded data, the date of sales, address, county, not full market price, vat exclusive and description were also removed, as the focus is on the postal code and the price of each property.

Counties such as Dún Laoghaire, Blackrock, Lucan or Swords don’t use the numbering system and are not part of the local government area, so they will be removed from the table as the focus is on urban area.

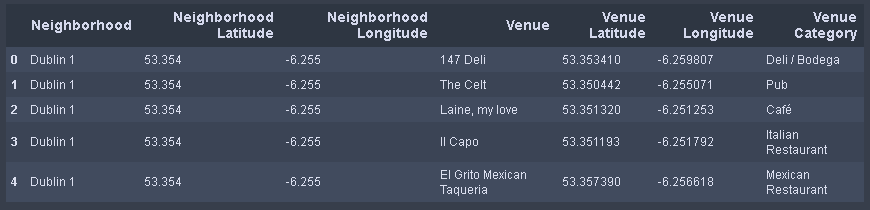
It’ll be used a radius of 750m from the centre of each district to find venues in the area on Four Square,.

## 3. Methodology

The final data used was composed of Postal Code, Average House Price, Latitude and Longitude. I used python folium to visualize the geographic details of Dublin and its regions, a circle marker and a label showing the name of each region was included in the map using latitude and longitude values to get the visuals, as below:



Next, I utilized the Foursquare API to explore each region and compared them in order to find similarities. It was considered a limit of 200 venues and a radius of 750m around the centre of each neighbourhood given by the scrapped latitude and longitude, an excerpt of the information obtained is below:



In summary, 138 unique categories were returned by Foursquare. A table showing the top 10 most common venue category for each region was created as below:



Next, I merged this table with the average house price table for obtaining a more complete understanding of the data and it was created a map showing the average house price and the 3 most common venues.



## 4. Conclusion

As a result, Dublin 22 really captures our attention as it has the lowest average house, which can be appealing to Dublin residents, and it doesn’t have a pharmacy nor a pub in the top 10 venues, which are potential businesses for investors, so they should proceed with further market analysis, evaluating other factors can impact on the final business decision.



## 5. Reference

[1] Information about Dublin, Wikipedia: <https://en.wikipedia.org/wiki/Dublin>

[2] Cost of Living Survey, Mercer: <https://www.mercer.ie/newsroom/dublin-is-ranked-46th-out-of-209-cities-around-the-world-in-the-mercer-2020-cost-of-living-survey.html>

[3] Residential Property Price Register, PRSA: <https://www.propertypriceregister.ie/website/npsra/pprweb.nsf/page/ppr-home-en>

[4] Postal Codes, Geo Names: <http://www.geonames.org/postalcode-search.html?q=&country=IE>

[5] Venues, Four Square API: <https://developer.foursquare.com/>