Peer-graded Assignment:

Capstone Project - The Battle of Neighbourhoods

Report

# Introduction/Business Problem (wk 1)

Dublin City is one of the most expensive places to live in Europe in terms of property prices and costs of living. A number of key pressures exist for individuals seeking to purchase a home in the city, in particular in the areas of the city closest to known commercial areas, business districts and where there are good public transport links. But what are the key factors that influence property prices in Dublin? Key areas include schools, commuting and public transport. Dublin is divided into townlands property prices can vary greatly between townlands. This report assesses which townlands are best for purchasing property having regard to influencing factors.

**Research question:**

What is the best townland in Dublin to buy a property in terms of price to influencing factors? We assess this by creain relation to townlands and influencing factors. We create segments of Dublin City by property price and then create segments of other factors, and establish the relevant correlations to answer the question.

For this task I will leverage the Foursquare location data in relation to the location of its wealth of data on commercial businesses as well as other data sources. The audience for this research would be someone wishing to purchase a property in Dublin. If links between property prices and other factors can be established we can understand what facilities are available as related to property prices. The individual therefore can make a decision based on their requirements.

# ****Data (wk 1)****

The data I will be using to solve this problem will come from two sources:

1. The data available on Foursquare in relation to the location of schools, and
2. Open data published by the Irish government and other public bodies.

There are two locations where this data is available:

1. The Irish Government’s open data portal: <https://data.gov.ie/> .
2. Dublin City’s local authority open data store, called “Dublinked”: https://data.smartdublin.ie/

A number of open datasets are available on the above sites that I will be using. These include:

1. A GeoJSON shapefile of the relevant areas of Dublin, including: <https://data.smartdublin.ie/dataset/dlr-townlands> .
2. Other GeoJSON files showing various explored factors such as transport links.

I will first use visualisation methods to show the area at interest and then assess the relevance of the various factors to property prices. I will use choropleth maps, k-means clustering and KNN classification.

# ****Methodology****

Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, if any, and what machine learnings were used and why.

# ****Results****

discuss the results

# ****Discussion****

discuss any observations you noted and any recommendations you can make based on the results.

# ****Conclusion****

conclude the report