

Suburb Venue Profiles and the Sydney Housing Market

IBM Data Science Professional Certificate - Capstone Project

Cameron Brown

Introduction

This project aims to investigate the correlation between the types and quantities of various venues in Sydney's suburbs and the median house price within the area. Investors, city planners, appraisers, and real estate agents are always seeking new ways of predicting the market value of houses and determining which factors may influence market movements. The information presented from this investigation will aid them in achieving a greater understanding of how neighbourhood venues impact real estate value in Sydney.

Data

There are four types of data that will be used in this project. The first is a list of suburbs in Sydney, Australia. This can be scraped from a webpage like Wikipedia:

https://en.wikipedia.org/wiki/List_of_Sydney_suburbs.

The second is the geographical coordinates of these suburbs, which can be acquired from a python package like geocoder or <https://www.geonames.org/>.

The third set of data that is required will come from Foursquare. This is a list of the venues within each Sydney suburb, including their type (e.g. restaurant, clothes shop, bank) and quantity.

The final set of information that is needed for this project is the median house price of each Sydney Suburb. This can be found on a number of websites online like

<http://house.speakingsame.com/suburbtop.php?sta=nsw&cat=HomePrice&name=&page=0>