Data Science Capstone Project

An Assessment of the Potential to Open a Shopping Mall

in Perth, Western Australia

INTRODUCTION

Perth is the fastest growing city in Australia. At the same time, it is also the most recent city in the world being some 3000 miles from the next city of similar size, Adelaide. This provides a potentially unique market for businesses. The growth in population means that the demand for goods and services in the city and its surrounding areas continues to increase. However, the city must be self-sufficient in meeting this demand, as distances preclude the cost-effective supply and movement of goods and services from other locations.

This unique situation, provides a potentially lucrative opportunity to establish a business providing essential goods and services to new and expanding communities. New communities in the Perth region are being developed to a similar plan, where residential zones are supported by retail centres and industrial parks. The centralization of retail outlets, the relatively fixed customer base

and the need to provide a wide-range of goods and services to fully service all the community's needs, suggests a retail solution based around a shopping mall construct.

Example Perth Suburb Development Plan

A shopping mall is more than just a place to buy groceries. It is also a place where relaxation and recreational activities can also be conducted for individuals, families and communities. This can involve grocery shopping, dining at restaurants, fashion outlets, watching films and many more activities. Shopping malls are a potential win-win opportunity for both customers and retailers. Customers can find everything they need under one roof, while retailers have a mass market and the opportunity to distribute and market their products and services efficiently and effectively. For property developers, the building of shopping malls allows them to earn consistent rental income and achieve economies of scale in terms of operational and management costs. However, the building of shopping malls is not a new concept and there are many companies looking for opportunities to build more. There are many factors involved in the selecting where a new mall should be built; however, as with many property developments schemes one the main criteria for success is 'location, location, location'.

BUSINESS PROBLEM

Project Objective.

The objective of this capstone project is to analyse and select the best locations in the city of Perth, Western Australia to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, this project aims to provide the information to answer the question: In the city of Perth, Western Australia, if a property developer is looking to open a new shopping mall, where would you recommend to do it?

Target Audience.

This capstone project will be useful to property developers looking to open a new mall in the city of Perth, Western Australia, investors seeking a new business opportunity, or retailers looking to expand. This project is timely as the city is forecasting a levelling off in the rate of its growth and the time available to establish new businesses may be limited. A fall in the population could even result in an over-provision of retail centres and the rationalization of even the existing business footprint in the city.

DATA SECTION

Data Requirements

To resolve the problem, we will need the following data:

- List of neighbourhoods in Greater Perth. This defines the scope of this project which includes the City of Perth and its suburbs within the region of Western Australia.
- Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to shopping malls. We will use this data to perform clustering on the neighbourhoods to analyse the scope of existing retail coverage.

- Population data for each neighbourhood, to help identify whether the clusters can provide sufficient business demand to make the investment in a shopping centre profitable.

Data sources and extraction methods

Data in respect of the City of Perth has been drawn from several sources. The main ones being: https://en.wikipedia.org/wiki/List_of_Perth_suburbs and https://quickstats.censusdata.abs.gov.au/census_servicesgetproduct/census/2016/quickstat/SSC51218. These data sources give a complete listing of the current suburbs of Perth with census information relating to 2016. Web scraping techniques to extract the data from these pages, with the help of Python requests and Beautifulsoup packages. Then we will get the geographical coordinates of the neighbourhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighbourhoods. After that, we will use Foursquare API to get the venue data for those neighbourhoods. Foursquare has one of the largest databases of 105+ million places and is used by over 125,000 developers. Foursquare API will provide many categories of the venue data, we are particularly interested in the Shopping Mall category in order to help us to solve the business problem put forward. This is a project that will make use of many data science skills, from web scraping (Wikipedia), working with API (Foursquare), data cleaning, data wrangling, to machine learning (K-means clustering) and map visualization (Folium).