

Capstone Project

Smart Living Sydney Providing Housing Services
Boon Fooi Cheng

Living in Sydney



Introduction

- Immigration service company, Smart Living Sydney is planning to expand its business to provide housing advises and services to clients in addition to immigration services
- As a start, the company is focusing on providing housing services for clients who are interested to live in Sydney

Problem Description

- Provide an overview of the Sydney based on venues available in all suburbs
- Based on the information acquired, the company is able to help clients from different cultural background to live in a suburb best suit them
- This study is focused on venue the following venue categories: restaurants, grocery stores, schools and public transports
- The company is focusing on clients from Chinese, Malaysian, Italian and Indian background

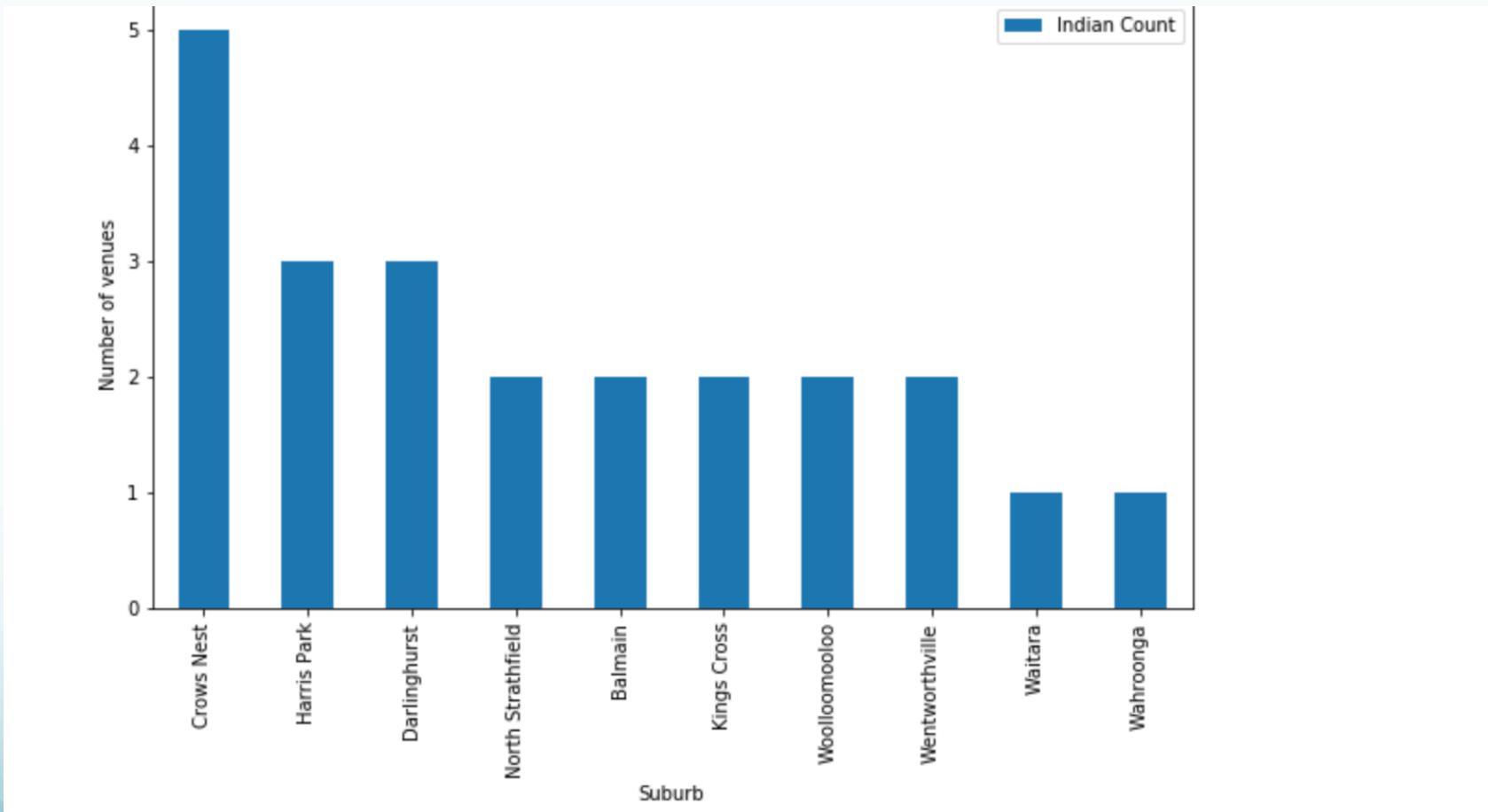
Data Source and Preparation

- The full list of Sydney suburbs is retrieved from this link:
<https://www.intosydneydirectory.com.au/sydney-postcodes.php>
- This list include columns: Suburb, State and Postcode
- As “State” is a constant value “NSW”, it is redundant data and it is discarded
- Python Geocoder is used to retrieve location (Latitude and Longitude) for each suburb
- FourSquare is used to retrieve venue information for each suburb
- Venues related to Chinese, Indian, Malay and Italian are extracted
- Other important venues extracted including schools and public transports

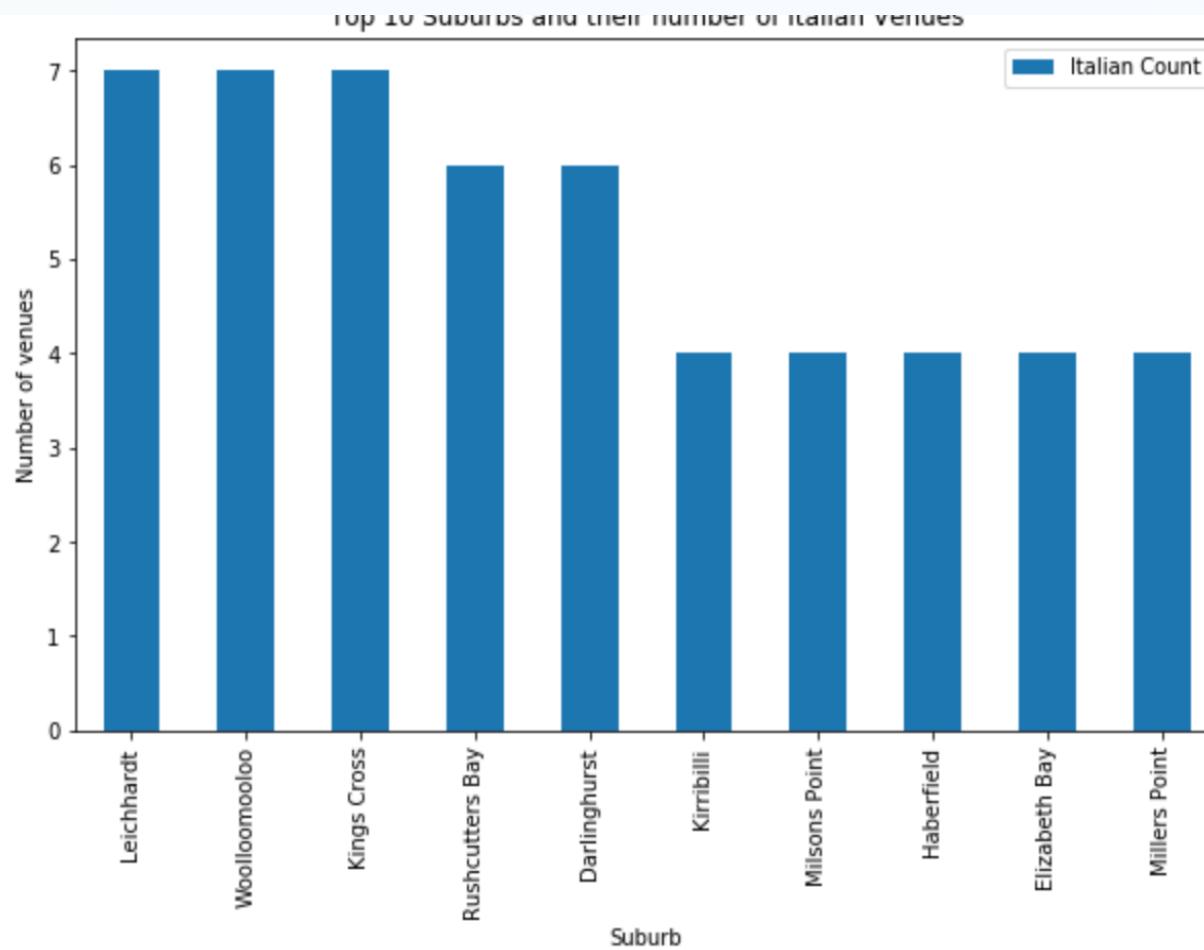
Missing Data

- No school venues returned by FourSquare
- FourSquare is not able to provide different type of grocery store other than the general store

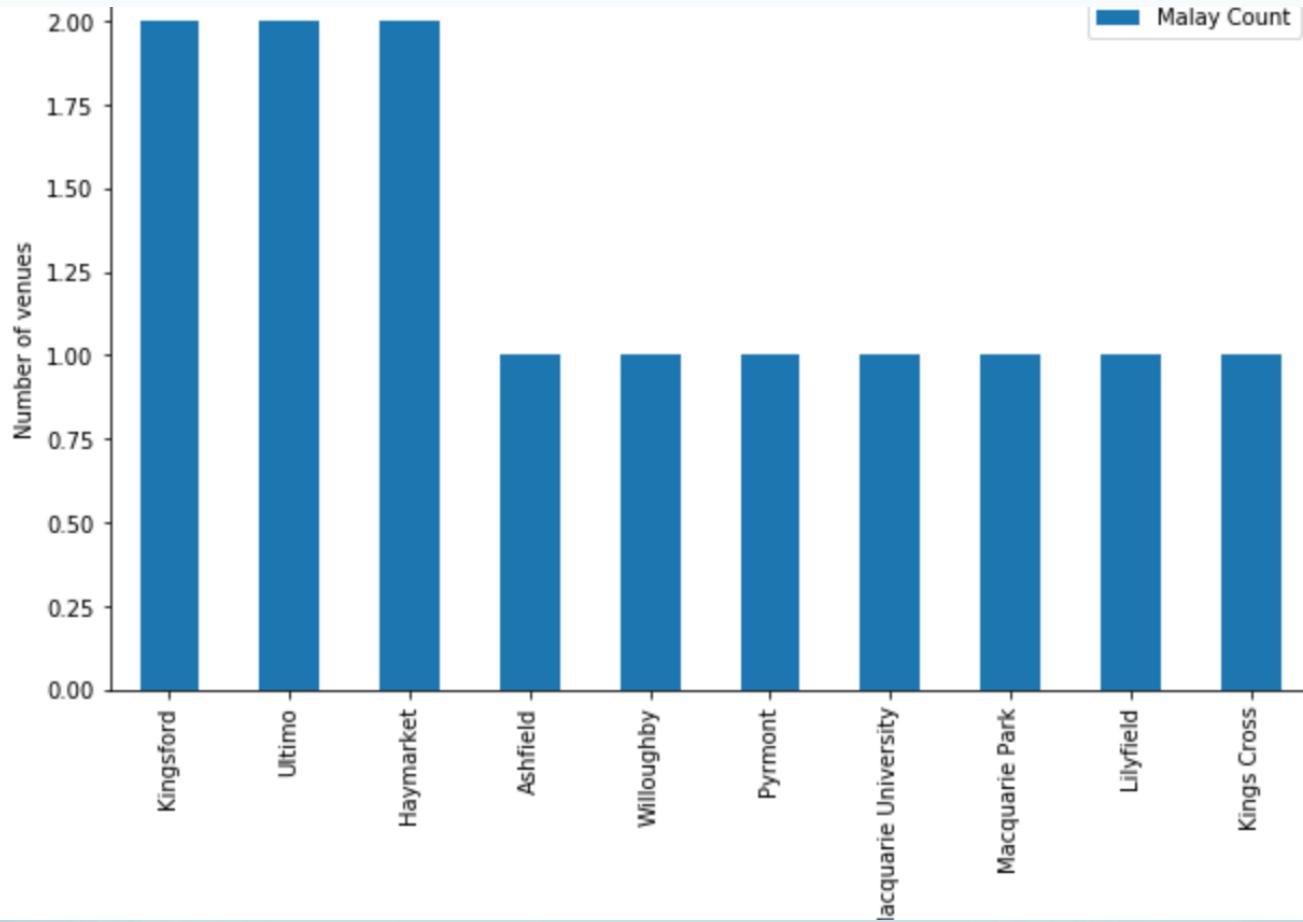
Top 10 Suburbs with Indian Venues



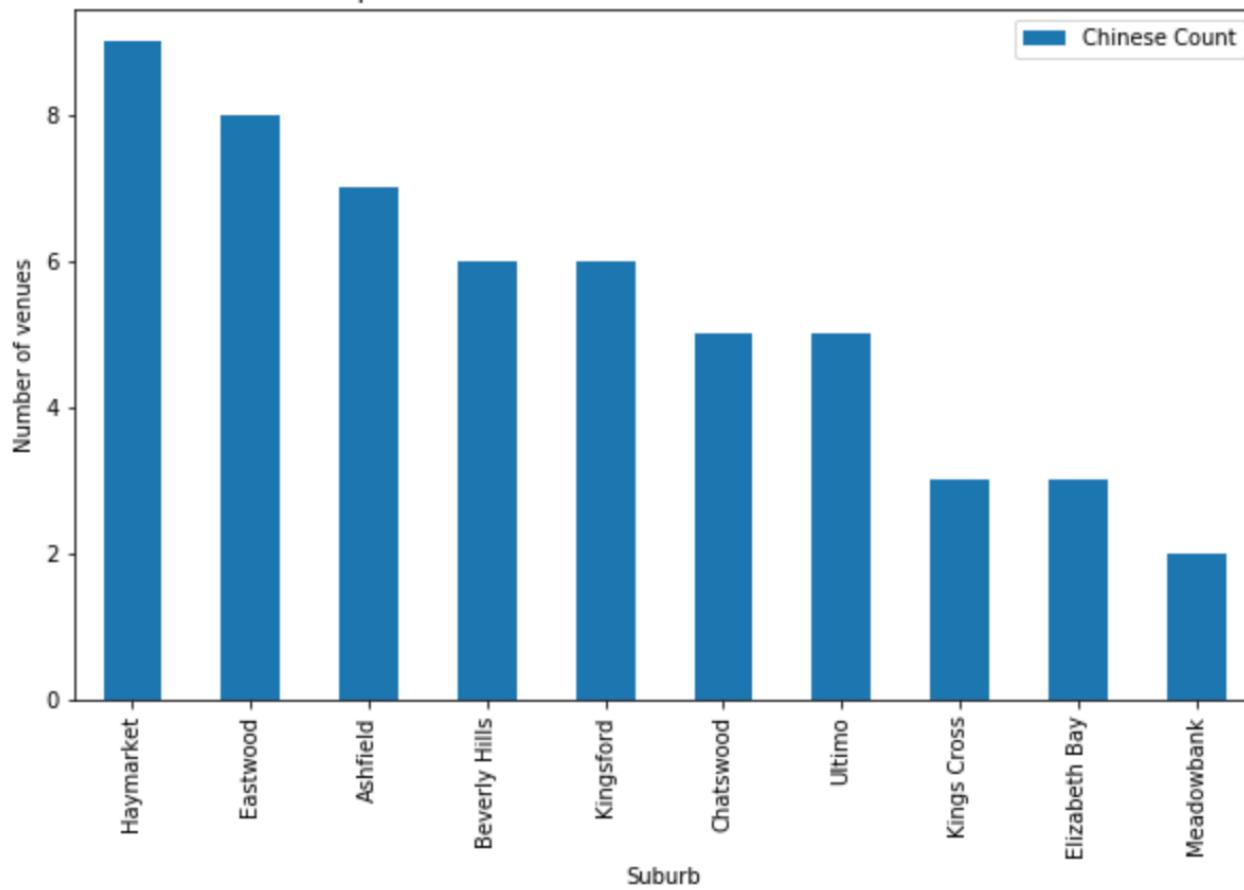
Top 10 Suburbs with Italian Venues



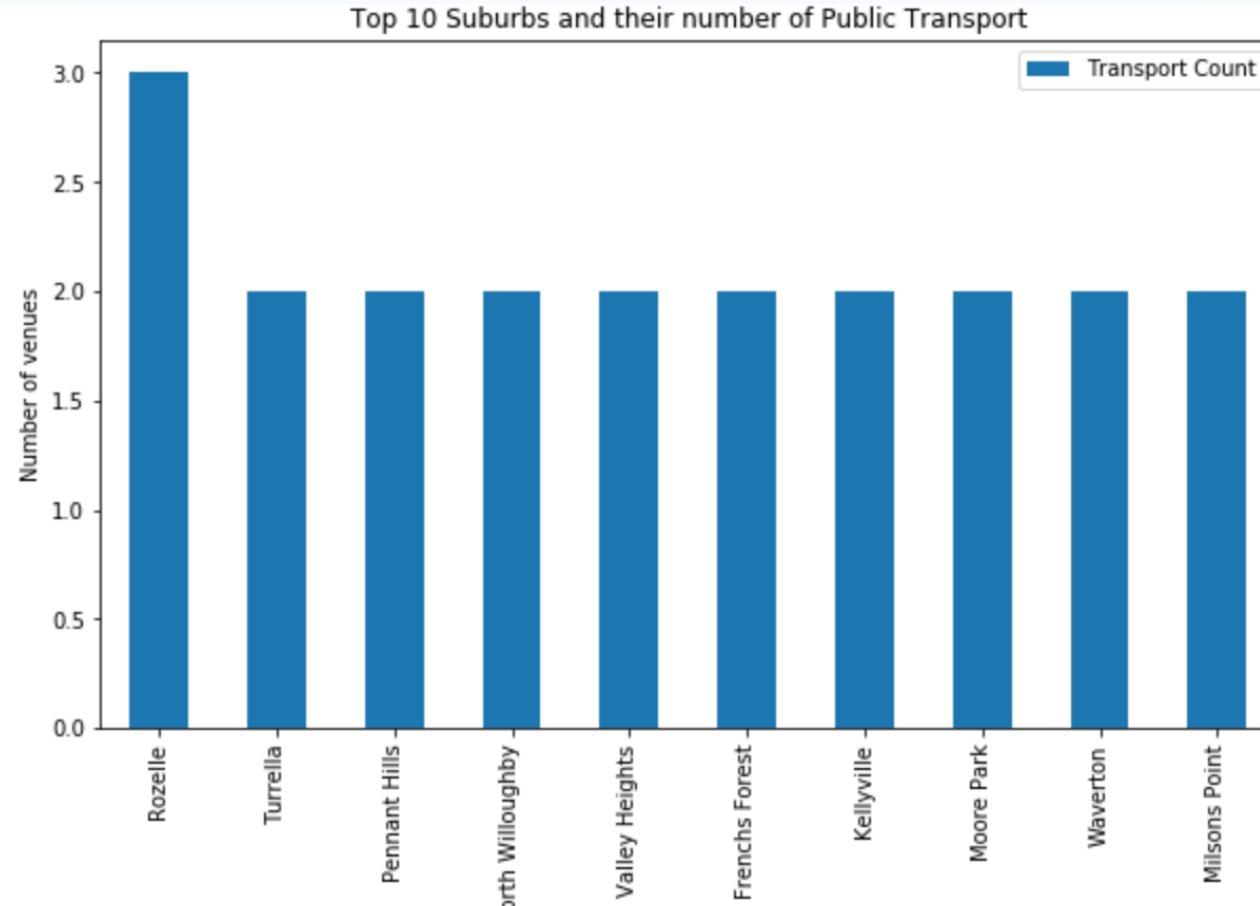
Top 10 Suburbs with Malaysian Venues



Top 10 Suburbs with Chinese Venues



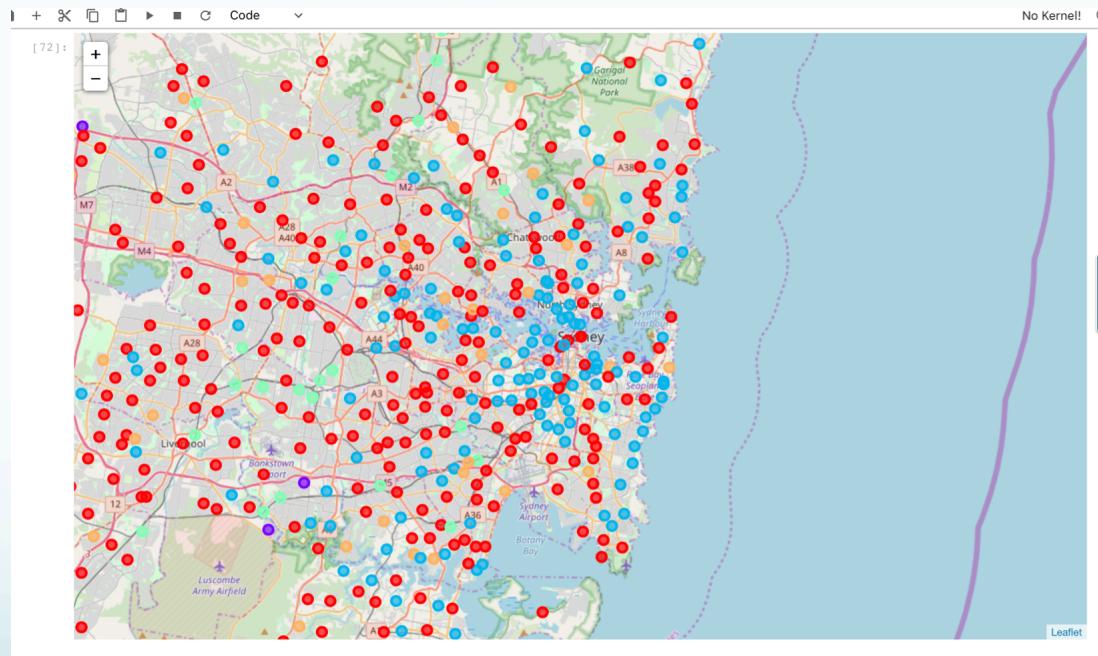
Top 10 Suburbs with Public Transport



Clustering of Sydney Suburbs

- K-Means modeling is used to perform clustering of Sydney suburbs
- Suburbs are divided into 5 clusters
- Housing cost of some suburbs is high even though they are ideal place to live. With the finding of the K-means clustering, we can propose other similar suburbs to live in with lower housing cost

Sydney Map K-Means Clusters Map



Discussion and Investigation of Results

- Results presented are heavily based on information provided by FourSquare
- Has FourSquare provided complete set of venues?
- Venues returned is depends on LIMIT and RADIUS set when setting up FourSquare url, different values of these two parameters might give different results
- Venues returned by FourSquare related to Indian, Chinese etc. are mainly restaurants. We have no information of other important venues such as grocery stores
- No school venue given by FourSquare, this is one of the important information for clients who has children of school age
- The public transport venues returned by FourSquare are limited. We have found the top suburb has 3 venues and other suburbs are having 2 venues. Therefore based on the limited data we have, it is hard for us to make conclusion which suburbs provide better public transport

Future Direction

- Acquire and provide school information to clients
- More detailed research of public transport facilities including frequencies and traveling time to major business districts
- Provide type of grocery stores, for example Indian or Italian?
- More detailed research of housing cost for cluster of suburbs in order to provide different housing options for clients