Capstone Project - The Battle of Neighbourhoods (Week 1)

Introduction

For a person like me, to relocate to a new place the following questions pop up in the mind:

1. What would be my rental expenditure?

2. How safe is the area?

3. How well it is connected by public transportation (bus/ train)?

4. How close it is from all the venues?

In quest of the answers, I have done lots of research online and finally came up with a thought to start working on finding solutions by doing some data analysis. As the problem is a very complex one, I am planning on doing it in stages. This article gives a detailed description of my first stage of work.

I chose to tackle two part of the above requirements i.e., connectivity by train and the details of top venues nearby the suburb of choice.

Problem Statement

Part-I

Clustering suburbs of Sydney and Melbourne on the bases of their connectivity to public transport (by train) and comparing how well Sydney and Melbourne are connected by train.

Part-II

Analysing the most common/popular venue details for all suburbs and their comparison.

Part-III

It is a comparison study of the top venues in major cities in Australia with the ones in New York.

Data Collection

I am planning to collect the data from the following sources:

- Data related to suburb and train stations from by scraping Wikipedia.
- Venue details from Foursquare develop API.

Scope of the Project

The analysis aiming to be done as a part of the project can be useful for people who are planning to relocate to Melbourne/Sydney.

They can get an information of how far off the train stations are to the suburbs and based on that initial information one can plan to make further decisions.

The analysis also gives an overview of the top venues in the suburbs

Also, from the third part of the project, one can get some insights into the Australian culture based on the details of the most popular venues.