The Battle of Neighbourhoods: IBM Data Science Capstone Project

Topic: Exploring the Food Culture in Melbourne
using Foursquare API
and K-Means Clustering

Submitted By: Tanay Arora

Date: 11/01/2021

1. Introduction

1.1. Background

Melbourne is known as the food capital of Australia. It is considered to be a multicultural melting pot as it is a home to more than 200 nationalities, hence similar number of cuisine inspirations. Quoting from an article on the food culture of Melbourne:
"There is no authentic Melbourne dish or cuisine. Rather, the city's food options are limitless because it is a multicultural melting pot. Food is a portal into culture, and Melbourne's vibrant immigration history remains at the forefront of its culture partly because of its undeniably international cuisine."

1.2. Problem

While on one hand having a plethora of food options is advantageous in an ethnically diverse city, however on the other hand exploring a wide array of food places might be equally time consuming for someone not so well verse with the city. Moreover, there is a fierce competition among the eateries to attract the foodies in the city, and hence new eateries keep popping on and off on Melbourne's radar.

The aim of this project is to explore the food preferences/taste of various suburbs in Melbourne by segmenting eateries in various suburbs based on the cuisines which they offer. In this project, the Foursquare's 'Places API' will be utilised to fetch venues in Each sub-urban location and further 'K-Means' algorithm will be used to segment these venues into clusters of similar cuisines. Moreover, a profile for each suburb location will be prepared describing the most common type of eateries using Exploratory Data Analysis (EDA), which would help to discover further about the culture and diversity of the neighborhood.

1.3 Stakeholders

Individuals

- The results from this analysis will be useful in understanding the distribution of diverse food cultures in Melbourne, which might making it easier for individuals to choose or navigate to their desired food destinations.

Businesses

- A high level understanding of the distribution of food culture of various suburbs can also be utilised by various business owners who might be planning to expand their ventures to other suburbs or open a new venture in any of the suburbs in Melbourne.

2. Data Acquisition

2.1 Melbourne City Dataset

The Melbourne City Dataset was retrieved from the following link: https://github.com/matthewproctor/australianpostcodes. The data is available in various Formats, however for the sake of simplicity the data was downloaded in form of .csv format.

The table below provides the description of the fields in the Melbourne City dataset along with an example of each field:

Field	Description	Example
id	Primary Key from source database	1
postcode	The postcode numerical format 0000-9999	3000, 3924, 3008
locality	The locality of the postcode - typically the city/suburb or postal distribution centre	Melbourne, Hawthorn, etc
state	The Australian state in which the locality is situated	VIC
long	The longitude of the locality - defaults to 0 when not available	144.956776
lat	The latitude of the locality - defaults to 0 when not available	-37.817403
dc	The Australia Post distribution Centre servicing this postcode - defaults to blank when not available	MELBOURNE
type	The type of locality, such as a delivery area, post office or a "Large Volume Recipient" such as a GPO, defaults to blank when not available	LVR
SA3	The SA3 Statistical Area code	215
SA3 Name	The name of the SA3	Melbourne City

	Statistical Area	
SA4	The SA4 Statistical Area code	21501
SA4 Name	The name of the SA4 Statistical Area	Melbourne - Inner
Region	Designated Region Area	R1
status	A note indicating whether the data is new, removed or updated - new column Nov 2018	Updated

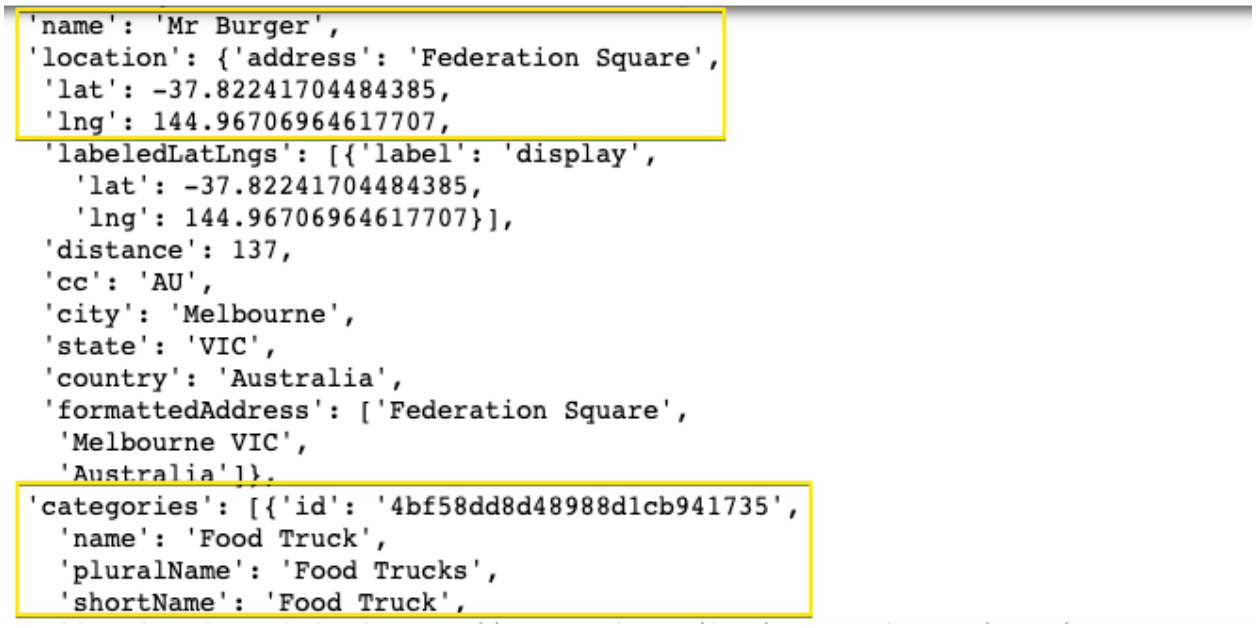
2.2 Foursquare Places API

The Places API offers real-time access to Foursquare's global database of rich venue data and user content. Link: https://developer.foursquare.com/docs/places-api/

In this project, RESTFUL API calls will be made to the Foursquare places to retrieve nearby venues corresponding to a particular geo-location. In particular, the 'explore' endpoint of the API

will be used, which returns a list of recommended venues in a json format near a particular location.

A sample response from one of the requests is shown below, showing a food venue near SouthBank, Melbourne:



Some of the key fields to note here are the name, latitude, longitude values. Moreover, the API returns the categorical information of the venue. These fields will be useful in categorising the venues into different types/cuisines. More on Foursquare categories: https://developer.foursquare.com/docs/build-with-foursquare/categories