Applied Data Science Capstone Project

CHOOSING THE TOWN
TO LIVE IN SINGAPORE

BOB DJ

January 2021

IBM Data Science Professional Certificate

INTRODUCTION

" Life is about choices. Some we regret, some we're proud of. Some will haunt us forever "

Making a choice is always a difficult thing to do, because we don't know if we will be happy with our choice or regret it, days / weeks / months from when we made the choice. Let's face the truth, most of the times we regret our choices, saying I should have talked to more people before deciding on something or 'done more research'. The other side is always greener or the food on the next plate always looks more tastier etc. and we could go on with the quotes or proverbs that highlight the effect and perception of a choice that we make in life.

" Before you make a choice make sure you can live with it "

Especially if you have to choose your place of living, you have to literally live with your choice. Choosing your place of living is a big choice for almost everyone and we try to get as much details possible before we make that leap of faith.

Singapore (SG) the city state is an amazing place to live in, if only we knew where to make our choice of stay. It has a plethora of attractions, restaurants, amenities, venues, towns to choose from. That statement poses the first problem for anyone hunting for a place to stay. Fortunately that's where Data Science comes in, as we have abundant data and the proper knowledge, tools and means to sift through them to slice and dice it, and get meaningful results out of it in an easy to understand and visually attractive format.

PROBLEM TO SOLVE - TARGET AUDIENCE

A prospective homebuyer or a tenant wannabe decides to choose a place to live in the island of Singapore. The person has a few basic specifications or needs that cannot be compromised. Let's call the person with a gender neutral common name - Robin. Robin has a few specifications or conditions that are a must to be adhered to as a basic need. This Capstone project paper is an attempt to help Robin or any similar person to navigate that first essential step towards finalising the place to live. So anyone who wants to buy a property or find a place to live in Singapore is a target audience for this paper. Robin's basic requirements regarding the town/place of stay in Singapore is

1. There should not be too many people living around.

2. Some key amenities are required to be around the living area within a walking distance

3. There should be an open water area close by.

This paper aims to help solve the above requirements (addressed as problem to solve going forward) using the concepts learnt and practised in the Coursera IBM Data Science Professional Certificate course.

To elaborate a bit more on the background of the problem, Robin wants to live in a locality that has less number of people compared to surrounding areas and has amenities for doing exercises and workouts, do regular shopping and places to relax having a drink with friends and family. Robin has a preference towards particular food and drinks that need to be taken into account also. The final requirement is the presence of a waterbody in the vicinity to have a peaceful walk around or if there are sports options for a swim or canoeing etc.

The above 3 problems are objectively quantified in the below table

Problem statement	Interpretation	Variable	Quantification
1. There should not be too many people living around	The number of people living around should be less than average of people across the towns of Singapore	Population	Population < Average population across SG Towns
1. Some amenities are required to be around the living area in a walking distance	Requirement is to have exercising options like gym and Shopping Malls close by, Needs options for drinking with places serving whisky and wine in that order.	Venues	Venue choices to be within a 1 kilometre radius
1. There should be a open water area close by	Any kind of waterbody like a Pond, River, Stream or Sea to be in the vicinity.	Natural feature	Feature to be within a 1 kilometre radius

DESCRIPTION OF DATA

To help solve the above described problem, we need 3 different types of data, first for identifying the various towns in Singapore and their population, second for getting their coordinates to plot in the map and to third to help finding venues, analysing details regarding the venues and natural features.

On researching the internet for the above sets of data, we came up with only sets of data satisfying the entire requirement.

1. Singapore towns and their population : The Singapore Government publishes data sets available for public view in **data.gov.sg** and this site has a specific and suitable data set for the project's requirement. The Housing and Development Board (HDB) of Singapore has a data set with the various SG Towns and their population for a 10 year period between 2008 and 2018.

2. Singapore towns' Co-ordinates : For getting the SG Town's coordinates of latitude and longitude we will use the Python Geocoder ArcGIS package

3. Venue Details : We will use Foursquare API top get the venue details of the SG Towns. **https://developer.foursquare.com/developer/** The Foursquare API provides comprehensive location data that can be used to get details of nearby venues, and details of these venues in a given area of SG Towns. Interestingly, it also has natural features like river & Waterfront that will be useful in our specific project requirement.

HOW THIS DATA WILL BE USED TO SOLVE THE PROBLEM

The sequence of steps to be followed in solving the problem for this Project will be as follows

A. First we will install the required packages of Python and import relevant Python libraries that are required for data loading, wrangling, analysis, graphs and maps rendering

B. Next we will get the available data on towns of Singapore from the SG Government's official data website and load that data into the Jupiter notebook

C. Then we will analyse and clean the SG towns' data and shortlist the data based on our requirement.

D. Then we will use the Geocoder package to get the geographical coordinates (Latitude & Longitude) for the SG towns

E. Followed by using the Foursquare API to receive the nearby venues and their details for all the towns that we have shortlisted.

F. Finally we will analyse the venue data based on our initial 3 requirements, and finalise the most appropriate SG Town where Robin can choose to stay

G. As an additional step, we will display the SG map with the shortlisted SG Towns and the selected township will be hardcoded and displayed with a marker.