

Final Report Battle of Neighbourhoods

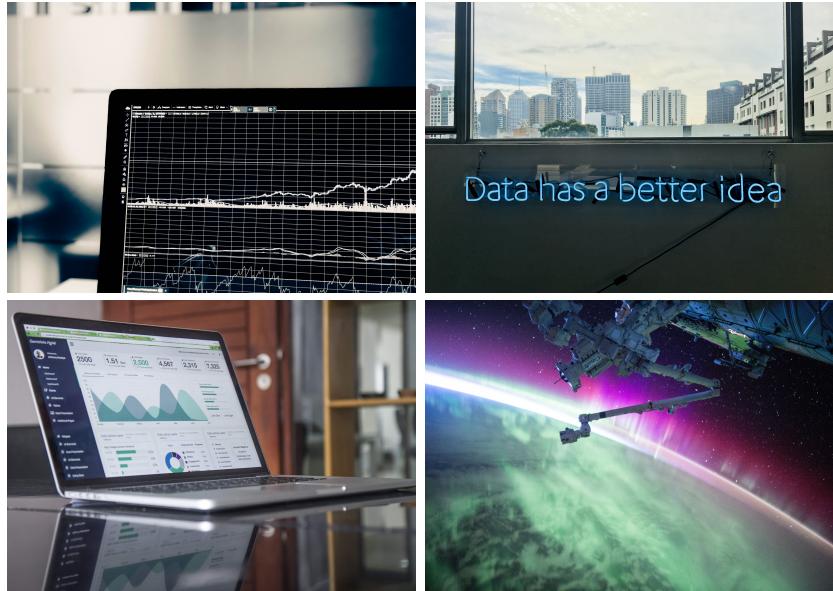


Introduction/Business Problem

As a father, the education of your children is one of the most important matters in the family. In Australia the selection of the school or in this case a secondary school is a very important issue for parents, as the system is competitive among the schools located in different suburbs (neighbourhoods) in the capital cities of Australia. In many cases, families in Australia take the decision of where to live, based on the school's ranking, the amenities and services available in a determined suburb/neighbourhood. For this work, I will develop a model to classify suburbs in the Melbourne Metro area based on the top secondary

**"Decision driven
by data are
better
decisions"**

Victor Alvarado P.



schools as per the website bettereducation.com.au. The whole purpose of this work is to give a tool to parents, in order to facilitate the decision process that will be driven by data analysis and not just for gut feeling.

Data Section Description

The data to be used comprise the use of the Foursquare API, where we will use it to gather relevant information about venues surrounding the top secondary schools in Melbourne. For that end we will be using the data from the website www.bettereducation.com.au, where I will scrap a table with the top secondary schools. On that table there is information related to the address of the school. With that column in the mentioned table I will use a geo-decoder to get the latitude and longitude of each school in the table.

With the geo-location the code will request to Foursquare all the venues around the school. At that point, the data will be cleaned and prepared to be fed in the k-mean clustering to address similarities among the neighbourhoods in Melbourne based on the top secondary schools.