

Is there room in Melbourne's inner suburbs for a new music venue

Melbourne's inner suburbs are teeming with live music venues, whether its a well known venue like 'The Tote' or a small intimate cafe or jazz bar. Is there space for another venue? or is this area too saturated. Using Machine Learning can we find a small pocket in the inner suburbs crying out for a live music venue.

Leveraging location data from Australia Post and the Foursquare API this project aims to explore Melbourne's inner suburbs to investigate whether there is a good location to set up a live music venue or whether they should look elsewhere.

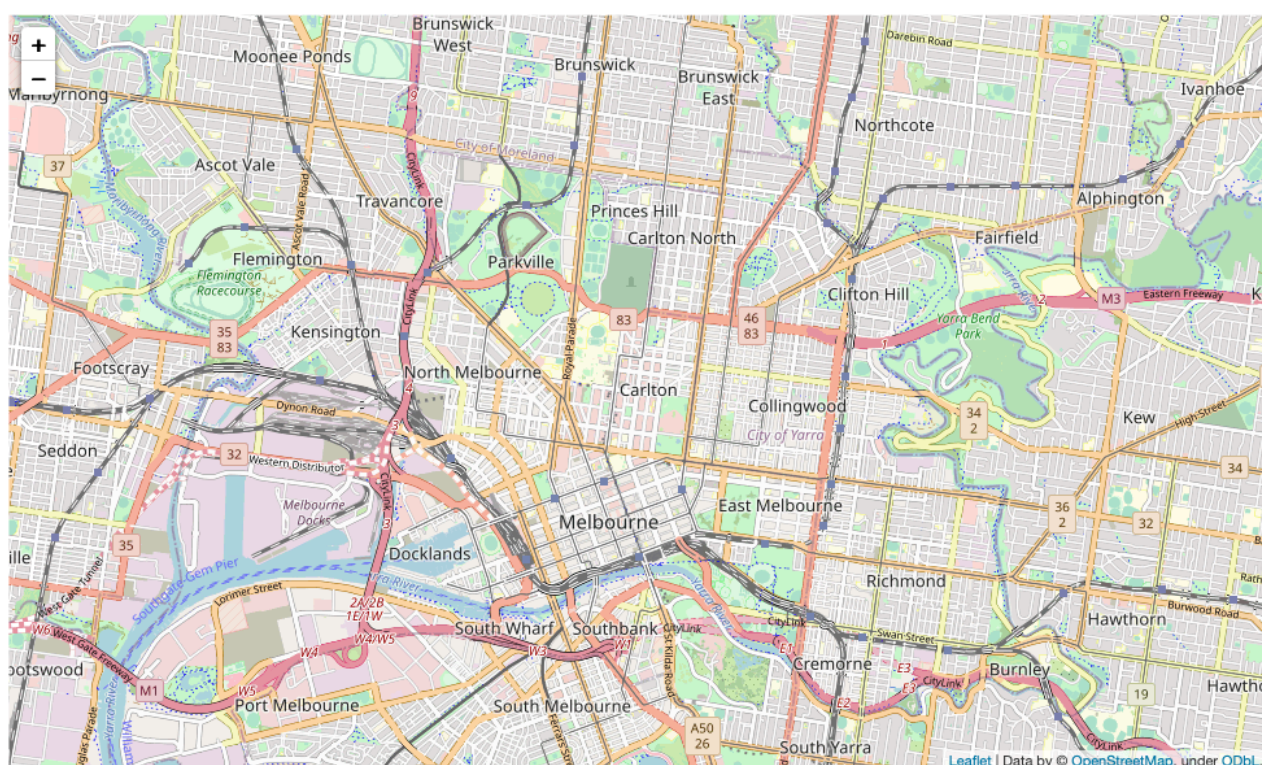
Data

This project draws data from a few different sources. Firstly a list of suburbs and post codes taken from Australia Post. Then using this data and the python geocode package we can then obtain latitude and longitudes for each post code. Finally, using the coordinates we can then obtain information on nearby music venues in each suburb using the Foursquare API.

Methodology

Firstly, I began by collecting Melbourne suburb and post code data from the Australia post website, the site generates a xml file which I filtered and exported to a csv file that contained only the inner suburbs and their corresponding post codes.

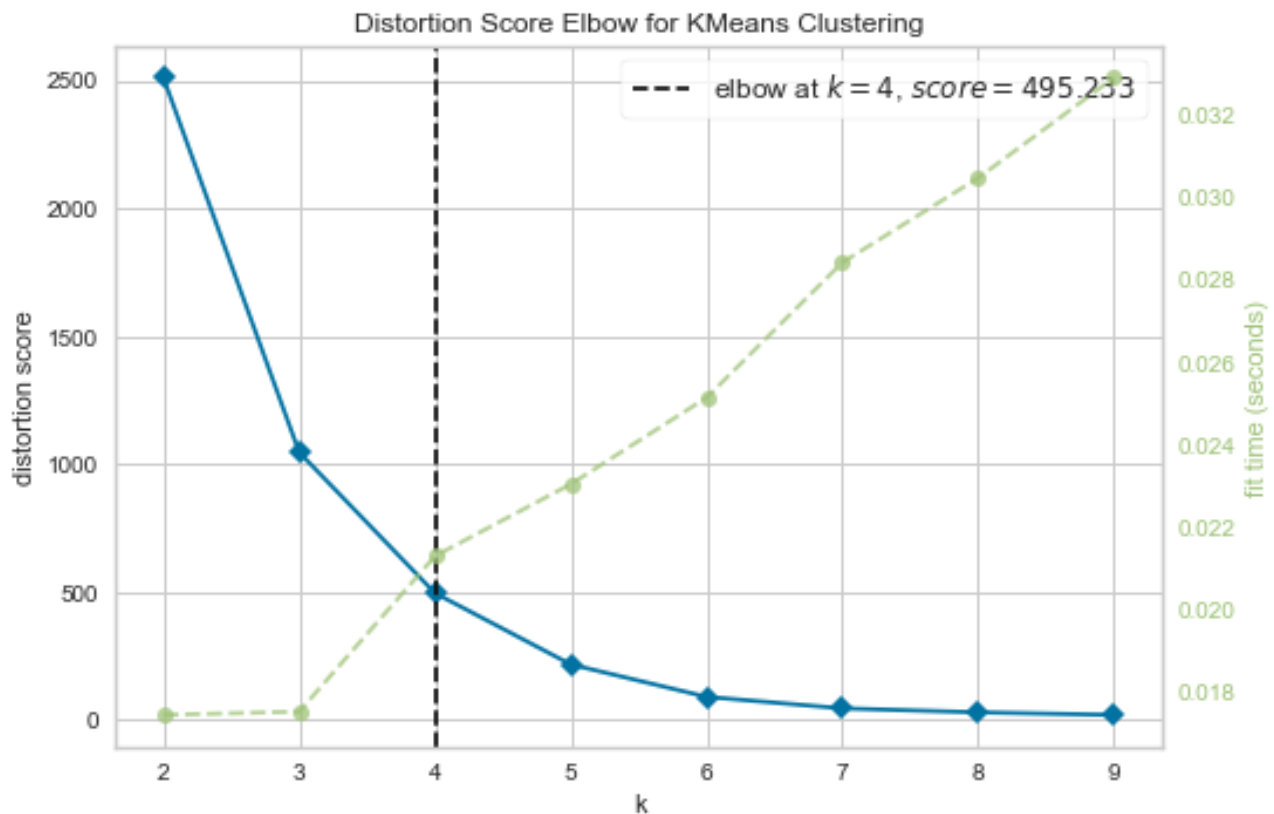
Using these post codes I then used the pgeocode package in python to get their corresponding latitudes and longitudes. Using Folium I then checked a few of the coordinates to make sure they were correct producing the following map.



Using this data, we can then get music venue location data from the Foursquare API. A bit of poking around on the Foursquare website I found the id category code for locations tagged as “music venue”. Using the API, the total amount of music venues in each suburb was found and added to the data frame.

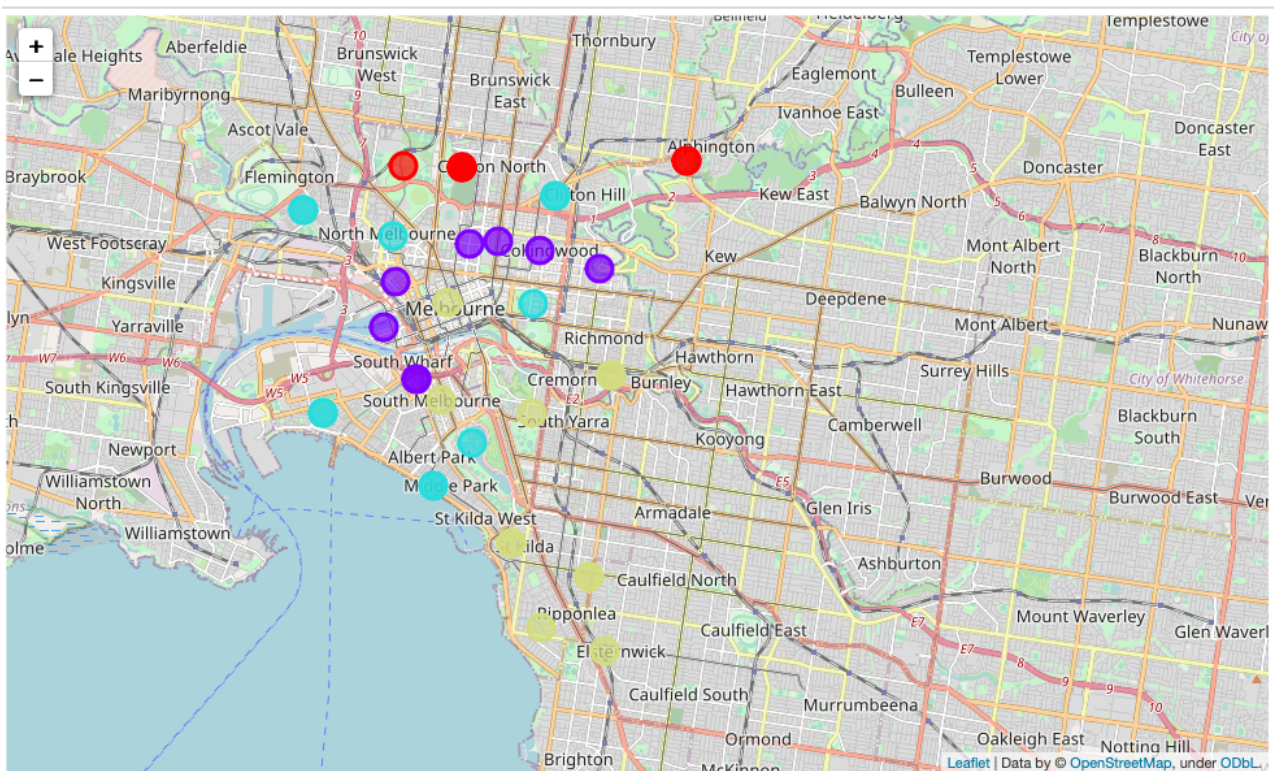
	Suburb	Post code	longitude	latitude	Music venues
0	Carlton	3053	144.96710	-37.8007	34
1	Carlton North	3054	144.96495	-37.7847	69
2	Docklands	3008	144.94450	-37.8183	38
3	East Melbourne	3002	144.98390	-37.8133	45
4	Flemington	3031	144.92295	-37.7937	56

Now we can begin clustering with Kmeans. We start by creating a clustering data frame with just Longitude, Latitude, and amount of music venues. Using the elbow method we then need to find the optimal amount of clusters.

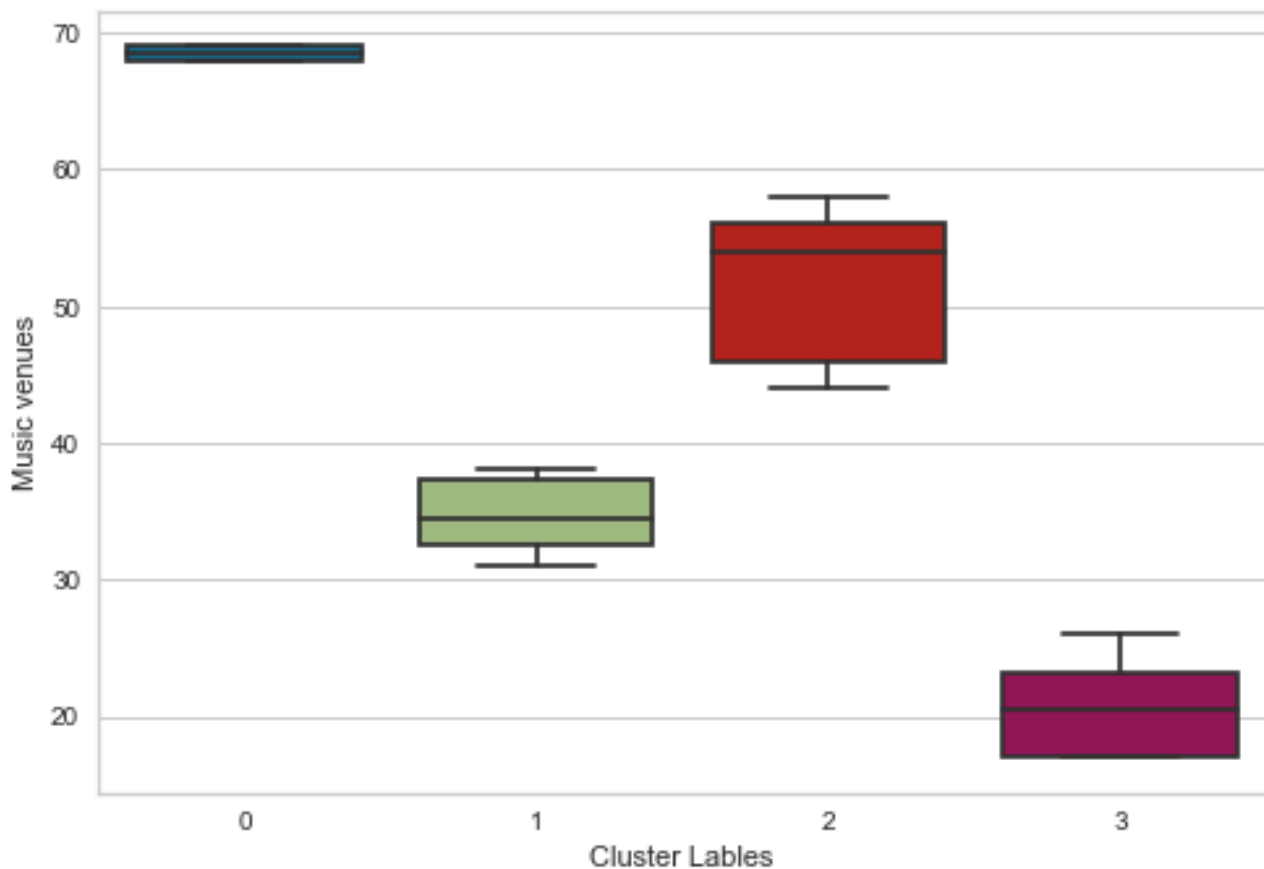


As we can see the optimal amount of clusters is $k=4$. After fitting the model with $k=4$ we then take cluster tables and add them to a new column in the original data frame.

Using folium we can then plot where the clusters are located.



Results



Looking at a boxplot of the clusters we can see that cluster 3 has the fewest amount of surrounding music venues. Looking at the map we can see that cluster 3 is located in the southern area of the inner suburbs, particularly around St Kilda and Richmond.

Conclusion

Looking at our data we can see that the inner north is heavily saturated with lots of surrounding live music venues, but heading south a music venue in a suburb like St Kilda or Richmond could have some success.