

Data Description

Introduction

I got a small circle of close friends, relatives and old colleagues, some of them emigrated to Melbourne of Australia in the past years and some of them intend to emigrate there in the coming few years. We all decided to pool money and work together for a possibility of setting up HK-styled restaurant as a sustainable business there and have close contact via WhatsApp and/or face-to-face meetings to discuss all related issues.

Data

We are going to use 2 data. The first one is number of pedestrians in different places. This can make sure there are enough customers so there is more profit.

The second data will be restaurant data of Melbourne. We will consider to open the restaurant there if there are not many restaurants there.

Methodology

First, we need to find a suitable location for the restaurant. This data is a research done by Melbourne government, counting number of pedestrians in different places from May 2009 to Dec 2018. We will choose the place with the most number of pedestrians to open the restaurant.

Source of the data:

<https://data.melbourne.vic.gov.au/Transport-Movement/Pedestrian-volume-updated-monthly-/b2ak-trbp/data>

In [8]:

```
import types
import pandas as pd
from ibm_botocore.client import Config
import ibm_boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage.
# It includes your credentials.
# You might want to remove those credentials before you share your note
# book.
client_393220bac62a487e84a1c851cb195d98 = ibm_boto3.client(service_name
='s3',
    ibm_api_key_id='HAQDeilACvSu2oFQmXGxNhV17TAaYQt7WOH6hCfYrk-5',
    ibm_auth_endpoint="https://iam.bluemix.net/oidc/token",
    config=Config(signature_version='oauth'),
    endpoint_url='https://s3-api.us-geo.objectstorage.service.networkla
yer.com')

body = client_393220bac62a487e84a1c851cb195d98.get_object(Bucket='myjup
yternotebookonibmdatascience-donotdelete-pr-nyx3lciuxuzpyj',Key='Data.c
sv')['Body']
# add missing __iter__ method, so pandas accepts body as file-like obje
ct
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __i
ter__, body )

df_data_2 = pd.read_csv(body)
df_data_3 = df_data_2.drop(columns=['ID', 'Date_Time', 'Year', 'Month', 'Md
ate', 'Day', 'Time', 'Sensor_ID'])
df_data_4 = df_data_3.groupby('Sensor_Name').sum()
df_data_5 = df_data_4.sort_values(by='Hourly_Counts', ascending=False)
df_data_5
```

Out [8] :

| | Hourly_Counts |
|-----------------------------------|---------------|
| Sensor_Name | |
| Town Hall (West) | 118271826 |
| Flinders Street Station Underpass | 99111597 |
| Melbourne Central | 98607295 |
| Bourke Street Mall (South) | 93420083 |
| Princes Bridge | 86990885 |
| Bourke Street Mall (North) | 83574898 |
| State Library | 62256715 |
| Spencer St-Collins St (North) | 53603225 |
| Flinders St-Elizabeth St (East) | 52804132 |
| Flagstaff Station | 52090258 |
| The Arts Centre | 44230514 |
| Flinders St-Swanston St (West) | 43250367 |
| Southern Cross Station | 41877673 |
| Southbank | 40805120 |
| Collins Place (South) | 36599252 |
| Australia on Collins | 32371408 |
| Sandridge Bridge | 31406243 |
| Collins Place (North) | 29491356 |
| Bourke St-Russell St (West) | 28216029 |
| City Square | 27950685 |
| Chinatown-Swanston St (North) | 26087902 |
| | |

| | |
|--|----------|
| Birrarung Marr | 25478985 |
| Hourly_Counts | |
| QV Market-Elizabeth St (West) | 25138344 |
| Sensor_Name | |
| Melbourne Convention Exhibition Centre | 24105150 |
| Lonsdale St (South) | 21822833 |
| Flinders La-Swanston St (West) | 21178783 |
| Collins St (North) | 19871607 |
| St Kilda Rd-Alexandra Gardens | 19576407 |
| Chinatown-Lt Bourke St (South) | 17011095 |
| New Quay | 16255300 |
| Victoria Point | 14888977 |
| Spencer St-Collins St (South) | 14530693 |
| Lygon St (West) | 13498876 |
| Little Collins St-Swanston St (East) | 12254750 |
| Webb Bridge | 12212799 |
| Melbourne Central-Elizabeth St (East) | 11921446 |
| Lonsdale St-Spring St (West) | 9876491 |
| Queen St (West) | 9419949 |
| Waterfront City | 8665354 |
| Grattan St-Swanston St (West) | 7994561 |
| Alfred Place | 7721046 |
| Monash Rd-Swanston St (West) | 6969185 |
| Flinders St-Spring St (West) | 6367118 |
| QV Market-Peel St | 5892701 |
| Flinders St-Spark La | 5856992 |
| Lygon St (East) | 5791013 |

| Sensor_Name | Hourly_Counts |
|------------------------------------|----------------------|
| Elizabeth St-Lonsdale St (South) | 4997319 |
| Tin Alley-Swanston St (West) | 3434026 |
| Faraday St-Lygon St (West) | 3248277 |
| QVM-Queen St (East) | 2489319 |
| Bourke St Bridge | 2000411 |
| QVM-Therry St (South) | 1971698 |
| Bourke St - Spencer St (North) | 1733013 |
| Lonsdale St - Elizabeth St (North) | 1199017 |
| QVM-Franklin St (North) | 1166600 |
| Pelham St (South) | 1105929 |
| Elizabeth St-La Trobe St (East) | 491681 |
| Lincoln-Swanston (West) | 421414 |
| Lincoln-Swanston(West) | 300194 |
| Pelham St (S) | 224089 |

From the above data, we found out that *Town Hall (West)* has the most number of pedestrians. The location of Town Hall is (-37.815135, 144.966439).

We are going to open the restaurant near there. Next, we are going to make sure there are not more than 15 restaurants in 200 metres nearby so there is less competition. We are going to use data provided by Melbourne government.

Data link:

<https://data.melbourne.vic.gov.au/Economy/Cafes-and-restaurants-with-seating-capacity/xt2y-tnn9/data>

In [40]:

```
body = client_393220bac62a487e84a1c851cb195d98.get_object(Bucket='myjupyternotebookonibmdatascience-donotdelete-pr-nyx3lciuxuzpyj',Key='Data2.csv')['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )

df_data_6 = pd.read_csv(body)
df_data_7 = df_data_6[['x coordinate','y coordinate']]
df_data_8 = df_data_7.loc[(df_data_7['x coordinate'] >= 144.965) & (df_data_7['x coordinate'] <= 144.967)]
df_data_9 = df_data_8.loc[(df_data_8['y coordinate'] >= -37.815) & (df_data_8['y coordinate'] <= -37.814)]
results = df_data_9.drop_duplicates()
df = results.reset_index(drop=True)
count = df['x coordinate'].count()
print("There are ", count, " restaurants in the area.")
```

There are 13 restaurants in the area.

Results

Based on the 2 creteria of within 200m of CBD with less than 15 pre-existing restaurants, we are going to open the restaurant near the Town Hall.

Discussion

Within 200m means walking distance for most customers in the CBD. Nevertheless, many customers would still prefer nearer restaurants. Therefore, opening a restaurant within 200m of CBD will be preduent to balance operating cost and viable business. Below 15 pre-existing restaurants would also be prudent to allow viable business.

Conclusion

In order to win over a competitive environment and groom a sustainable business in the CBD in Melbourne, careful planning and deliberation of suitable location for the intended restaurant is required in order to secure a viable business. It is concluded that we are going to open a restaurant near the Town Hall based on the criteria of within walking distance of 200m from the CBD and still reasonable market situation of below 15 pre-existing restaurants.