

# **Caffe Shops - Investment opportunities in Toronto Canada for Colombian investors.**

Alvaro Quiceno

March 29, 2020

## **1. Introduction**

### **1.1 Background**

Colombia is one of the largest coffee producers in the world, the quality of its product is considered the best for its softness, sweetness and aroma, and it has a strong globally known brand of Caffe from Colombia.

### **1.2 Problem and interest**

A group of coffee producers (coffee growers) called Almacafe wants to take advantage of the advantages described above, so they plan to open coffee shops in the City of Toronto, Canada.

They investors want to know:

- If there is a potential market for coffee shops.
- The market share compared to the different categories of consumption.
- The neighborhood(s) to establish the coffee shop.

## **2. Obtaining and cleaning the data**

The data that will be used to carry out the analysis and deliver the conclusions to investors will be the following:

**2.1 Political division of Toronto Canada**, this information will be used to identify the neighborhood where the coffee shops are located, the information will be obtained from the following route, [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) and it will be stored locally in a file called `Toronto_FSAs.csv`, here is an example of the data:

```
df_Toronto = pd.read_csv("Toronto_FSAs.csv", ";")  
df_Toronto.head()
```

	Postcode	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront

**2.2 Toronto Canada geospatial data** that will be related to the political division and Foursquare will be downloaded from the route [http://coc1.us/Geospatial\\_data](http://coc1.us/Geospatial_data) and will be stored locally in a file called `Geospatial_Coordinates.csv`, at below is an example of the data:

```
dfgeo = pd.read_csv("Geospatial_Coordinates.csv",",")
dfgeo.head()
```

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

**2.3 Know the consumption trends** and places that potential clients visit, the Foursquare API will be used, an example in JSON format of the information obtained is shown below, and the main Features are highlighted (this information will be processed and will become a Panda's DataFrame):

```
a. {'meta': {'code': 200, 'requestId': '5e7f4bce98205d001b004f17'},
b. 'response': {'warning': {'text': "There aren't a lot of r
results near you. Try something more general, reset your fi
lters, or expand the search area."},
c. 'headerLocation': 'Malvern',
d. 'headerFullLocation': 'Malvern, Toronto',
e. 'headerLocationGranularity': 'neighborhood',
f. 'totalResults': 1,
g. 'suggestedBounds': {'ne': {'lat': 43.8111863045, 'lng':
-79.18812958073042},
h. 'sw': {'lat': 43.80218629549999, 'lng': -79.20057721926
96}},
i. 'groups': [{'type': 'Recommended Places',
j. 'name': 'recommended',
k. 'items': [{'reasons': {'count': 0,
l. 'items': [{'summary': 'This spot is popular',
m. 'type': 'general',
n. 'reasonName': 'globalInteractionReason'}}]},
o. 'venue': {'id': '4bb6b9446edc76b0d771311c',
p. 'name': 'Wendy's',
q. 'location': {'crossStreet': 'Morningside & Sheppard
',
```

```

r.      'lat': 43.80744841934756,
s.      'lng': -79.19905558052072,
t.      'labeledLatLngs': [{'label': 'display',
u.          'lat': 43.80744841934756,
v.          'lng': -79.19905558052072}],
w.      'distance': 387,
x.      'cc': 'CA',
y.      'city': 'Toronto',
z.      'state': 'ON',
aa.     'country': 'Canada',
bb.     'formattedAddress': ['Toronto ON', 'Canada']],
cc.     'categories': [{'id': '4bf58dd8d48988d16e941735'
,
dd.     'name': 'Fast Food Restaurant',
ee.     'pluralName': 'Fast Food Restaurants',
ff.     'shortName': 'Fast Food',
gg.     'icon': {'prefix': 'https://ss3.4sqi.net/img/c
categories_v2/food/fastfood_',
hh.     'suffix': '.png'},
ii.     'primary': True}],
jj.     'photos': {'count': 0, 'groups': []}},

```

```
'referralId': 'e-0-4bb6b9446edc76b0d771311c-0'}}]]]]}
```

A detailed explanation of the data cleaning and description of the processes will be provided in the next sections.