

IBM Capstone Project: The Battles of Neighborhood – Final Project

By Cristiane Foust

Opening an Upscale Brazilian Steakhouse in Toronto, CA



1. Introduction/Business Problem

In this project, we will explore ideal neighborhood locations to open a Brazilian steakhouse restaurant. Stakeholders interested in opening an upscale all-you-can-eat Brazilian steakhouse restaurant in Toronto, Canada will use this report. Stakeholders already have 6 successful upscale all-you-can-eat Brazilian steakhouses in Brazil and 10 in United States. They decided to expand their business to Canada starting in Toronto since Toronto is the largest city with population more than 5.4 million. We will look for existing upscale restaurants in the vicinity of Toronto in each neighborhood and if there are any upscale Brazilian steakhouses. We will then focus our preliminary analysis to give the best possible neighborhood options for stakeholders in order to open an upscale all-you-can-eat Brazilian steakhouse restaurant in Toronto.

2. Data

In this project, we will use the following data sources:

1. Neighborhood data of Toronto City from the Wikipedia page https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. (From this data, we will extract 'Postal code', 'Borough' and 'Neighborhood' for Toronto)
2. Download location data using the link http://cocl.us/Gespatial_data and extract the 'Latitude' and 'Longitude' for Toronto neighborhoods.
3. Foursquare API to retrieve geo-location information for existing upscale restaurants in each neighborhood in Toronto and then verify if there are any upscale Brazilian steakhouses in Toronto.

Finally we will analyze the above data sources to identify ideal neighborhood locations in Toronto to open a Brazilian steakhouse restaurant. We just need to be aware that this report relies on the accuracy of Foursquare API for Toronto, CA.

2.1 Import Libraries

```
import pandas as pd
import numpy as np
import json
import requests
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)
from pandas.io.json import json_normalize # transform JSON file into a pandas dataframe
# Use geopy library to get the latitude and longitude values
!conda install -c conda-forge geopy --yes
from geopy.geocoders import Nominatim
!pip install beautifulsoup4
from bs4 import BeautifulSoup
!pip install html5lib

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

# !conda install -c conda-forge folium=0.5.0 --yes
import folium # map rendering library

print('Libraries imported.')
Libraries imported.
```

2.2 Transform neighborhood data of Toronto City from the Wikipedia page using the BeautifulSoup package into a pandas dataframe.

```
CA_website_url = requests.get('https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M').text

# import lxml
soup = BeautifulSoup(CA_website_url, "html5lib")
htmltable = soup.find('table', { 'class' : 'wikitable sortable' })
```

```

def tableDataText(table):
    rows = []
    trs = table.find_all('tr')
    headerow = [td.get_text(strip=True) for td in trs[0].find_all('th')] # header row
    if headerow: # if there is a header row include first
        rows.append(headerow)
    trs = trs[1:]
    for tr in trs: # for every table row
        rows.append([td.get_text(strip=True) for td in tr.find_all('td')]) # data row
    return rows

```

```
list_table = tableDataText(htmhtable)
```

```

import pandas as pd
dftable = pd.DataFrame(list_table[1:], columns=list_table[0])
dftable.head(10)

```

	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
5	M5A	Downtown Toronto	Regent Park
6	M6A	North York	Lawrence Heights
7	M6A	North York	Lawrence Manor
8	M7A	Queen's Park	Not assigned
9	M8A	Not assigned	Not assigned

```

indexNames = dftable[ dftable['Borough'] =='Not assigned'].index
dftable.drop(indexNames , inplace=True)

```

```
dftable.loc[dftable[ 'Neighbourhood'] =='Not assigned' , 'Neighbourhood'] = dftable[ 'Borough']
```

```

result = dftable.groupby(['Postcode','Borough'], sort=False).agg( ' , '.join)
df_new=result.reset_index()
df_new.head(15)

```

	Postcode	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Harbourfront, Regent Park
3	M6A	North York	Lawrence Heights, Lawrence Manor
4	M7A	Queen's Park	Queen's Park
5	M9A	Etobicoke	Islington Avenue
6	M1B	Scarborough	Rouge, Malvern
7	M3B	North York	Don Mills North
8	M4B	East York	Woodbine Gardens, Parkview Hill
9	M5B	Downtown Toronto	Ryerson, Garden District
10	M6B	North York	Glencairn
11	M9B	Etobicoke	Cloverdale, Islington, Martin Grove, Princess ...
12	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union
13	M3C	North York	Flemington Park, Don Mills South
14	M4C	East York	Woodbine Heights

2.3 Use Toronto location data csv file to create a dataframe with latitude and longitude values

```
!wget -q -O 'Toronto_long_lat_data.csv' http://cocl.us/Geospatial_data
df_long_lat = pd.read_csv('Toronto_long_lat_data.csv')
df_long_lat.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

```
df_long_lat.columns=['Postalcode','Latitude','Longitude']
df_long_lat.head()
```

	Postalcode	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

```
df_pc_long_lat = df_long_lat.rename(columns={'Postalcode': 'Postcode'})
df_pc_long_lat.set_index("Postcode")
df_new.set_index("Postcode")
toronto_data=pd.merge(df_new, df_pc_long_lat)
toronto_data
```

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront, Regent Park	43.654260	-79.360636
3	M6A	North York	Lawrence Heights, Lawrence Manor	43.718518	-79.464763
4	M7A	Queen's Park	Queen's Park	43.662301	-79.389494
5	M9A	Etobicoke	Islington Avenue	43.667856	-79.532242
6	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
7	M3B	North York	Don Mills North	43.745906	-79.352188
8	M4B	East York	Woodbine Gardens, Parkview Hill	43.706397	-79.309937
9	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937

Note: Top 10 results are listed above.

2.4 Get geographical coordinates of Toronto

```
address = 'Toronto, ON'
geolocator = Nominatim(user_agent="Toronto")
location = geolocator.geocode(address)
latitude_toronto = location.latitude
longitude_toronto = location.longitude
print('The geographical coordinates of Toronto are {}, {}'.format(latitude_toronto, longitude_toronto))
```

The geographical coordinates of Toronto are 43.653963, -79.387207.

2.5 Use Foursquare API to retrieve geo-location information from all upscale restaurants in each neighborhood in Toronto limiting to 500 meters of radius and 100 venues for each neighborhood

```
CLIENT_ID = 'MHSM3SCFN51MFFPY0IWYRVORSOWWLMONRIEGDU2YYCRDGDX3' # your Foursquare ID
CLIENT_SECRET = 'DJA15GQSOHK3TRZZT0XP1X4U5FEI5BZRLH2HQWNZNRAE1Z1' # your Foursquare Secret
VERSION = '20180605' # Foursquare API version

print('Your credentials:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET: ' + CLIENT_SECRET)
```

```
# defining radius and limit of venues to get
radius=500
LIMIT=100
```

```
import json # library to handle JSON files
def getUpscaleRestaurantVenues(names, latitudes, longitudes, borough, radius=500):

    venues_list=[]
    for name, lat, lng, borough in zip(names, latitudes, longitudes, borough):

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]["groups"][0]["items"]

        # return only relevant information for each nearby venue
        venues_list.append([
            name,
            lat,
            lng,
            borough,
            v["venue"]["name"],
            v["venue"]["categories"][0]["name"] for v in results]
        )

    return(venues_list)
```

```
nearby_food_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_list])
nearby_food_venues.columns = ['Neighbourhood',
                             'Neighbourhood Latitude',
                             'Neighbourhood Longitude',
                             'Borough',
                             'Venue',
                             'Venue Category']

return(nearby_food_venues)
```

```
: toronto_venues = getUpscaleRestaurantVenues(names=toronto_data['Neighbourhood'],
                                              latitudes=toronto_data['Latitude'],
                                              longitudes=toronto_data['Longitude'],
                                              borough=toronto_data['Borough']
                                              )
```

```
: toronto_venues
```

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Borough	Venue	Venue Category
0	Harbourfront, Regent Park	43.654260	-79.360636	Downtown Toronto	Cluny Bistro & Boulangerie	French Restaurant
1	Harbourfront, Regent Park	43.654260	-79.360636	Downtown Toronto	Pure Spirits Oyster House & Grill	Seafood Restaurant
2	Don Mills North	43.745906	-79.352188	North York	Gonoé Sushi	Japanese Restaurant
3	Ryerson, Garden District	43.657162	-79.378937	Downtown Toronto	Barberian's Steak House	Steakhouse
4	St. James Town	43.651494	-79.375418	Downtown Toronto	GEORGE Restaurant	Restaurant
5	St. James Town	43.651494	-79.375418	Downtown Toronto	Carisma	Italian Restaurant

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Borough	Venue	Venue Category
6	St. James Town	43.651494	-79.375418	Downtown Toronto	Wildfire Steakhouse Cosmopolitan	Steakhouse
7	Berczy Park	43.644771	-79.373306	Downtown Toronto	Harbour 60 Toronto	Steakhouse
8	Leaside	43.709060	-79.363452	East York	GRILLTIME	Steakhouse
9	Central Bay Street	43.657952	-79.387383	Downtown Toronto	Barberian's Steak House	Steakhouse
10	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Bosk at Shangri-La	Asian Restaurant
11	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Estiatorio Volos	Greek Restaurant
12	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Hy's Steakhouse	Steakhouse
13	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Kojin	Colombian Restaurant
14	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Michael's on Simcoe	Steakhouse
15	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Nota Bene	American Restaurant
16	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Bymark	Restaurant
17	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	The Fifth & Terrace	Modern European Restaurant
18	Adelaide, King, Richmond	43.650571	-79.384568	Downtown Toronto	Los Colibris	Mexican Restaurant
19	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	Downtown Toronto	Harbour 60 Toronto	Steakhouse
20	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	Downtown Toronto	Aria Ristorante	Italian Restaurant
21	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	Downtown Toronto	Streak's Bar and Grill	Steakhouse
22	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	Downtown Toronto	Hot Stove Club	Steakhouse
23	Little Portugal, Trinity	43.647927	-79.419750	West Toronto	Böhmer	New American Restaurant
24	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Bymark	Restaurant
25	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Hy's Steakhouse	Steakhouse
26	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Bosk at Shangri-La	Asian Restaurant
27	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Estiatorio Volos	Greek Restaurant
28	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	TOCA	Italian Restaurant
29	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Michael's on Simcoe	Steakhouse
30	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Kojin	Colombian Restaurant
31	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Little Anthony's	Italian Restaurant
32	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Wildfire Steakhouse Cosmopolitan	Steakhouse
33	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	Hot Stove Club	Steakhouse
34	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576	Downtown Toronto	LA's Italian + Bar	Italian Restaurant
35	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Bymark	Restaurant

	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Borough	Venue	Venue Category
36	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Hy's Steakhouse	Steakhouse
37	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Carisma	Italian Restaurant
38	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Estiatorio Volos	Greek Restaurant
39	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Bosk at Shangri-La	Asian Restaurant
40	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Little Anthony's	Italian Restaurant
41	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	BATCH	Gastropub
42	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Wildfire Steakhouse Cosmopolitan	Steakhouse
43	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	LA's Italian + Bar	Italian Restaurant
44	Commerce Court, Victoria Hotel	43.648198	-79.379817	Downtown Toronto	Hot Stove Club	Steakhouse
45	Studio District	43.659526	-79.340923	East Toronto	M'Eat	Steakhouse
46	Bedford Park, Lawrence Manor East	43.733283	-79.419750	North York	Francobollo	Italian Restaurant
47	Bedford Park, Lawrence Manor East	43.733283	-79.419750	North York	Nara Sushi	Japanese Restaurant
48	Willowdale South	43.770120	-79.408493	North York	The Keg	Steakhouse
49	Willowdale South	43.770120	-79.408493	North York	Symposium Cafe Restaurant & Lounge	Restaurant
50	High Park, The Junction South	43.661608	-79.464763	West Toronto	Aquila	Steakhouse
51	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846	Downtown Toronto	Carisma	Italian Restaurant
52	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846	Downtown Toronto	Wildfire Steakhouse Cosmopolitan	Steakhouse
53	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846	Downtown Toronto	Biagio Ristorante	Italian Restaurant
54	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Bymark	Restaurant
55	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Hy's Steakhouse	Steakhouse
56	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Bosk at Shangri-La	Asian Restaurant
57	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Estiatorio Volos	Greek Restaurant
58	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Michael's on Simcoe	Steakhouse
59	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Kojin	Colombian Restaurant
60	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Little Anthony's	Italian Restaurant
61	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Los Colibris	Mexican Restaurant
62	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	Wildfire Steakhouse Cosmopolitan	Steakhouse
63	First Canadian Place, Underground city	43.648429	-79.382280	Downtown Toronto	LA's Italian + Bar	Italian Restaurant
64	Church and Wellesley	43.665860	-79.383160	Downtown Toronto	Loaded Pierogi	Polish Restaurant

Methodology

This is the most important component of the report. This is where we discuss and explain any exploratory data analysis completed, any inferential statistical testing performed, and what kind of machine learning technique was used and why. From the above data, we used content-based recommendation technique to solve the problem.

After analyzing the Foursquare API data, we noticed that there aren't any upscale all-you-can-eat Brazilian steakhouse restaurants (Brazilian Churrascarias) in Toronto. This means that stakeholders would not have a direct competition if they open their new restaurant in any of the neighborhoods in Toronto. Since stakeholders mentioned that their other restaurants are very popular with businesses and, at the same time, they are very family-friendly, it is essential for them to be close to other popular upscale restaurants. A Brazilian steakhouse is not exactly the same as other steakhouses. This type of steakhouse is called "rodizio" which means that many Brazilian cowboys ("gaúchos") walk table to table with different cuts of meat, chicken, sausages and even seafood. There is always a very large salad bar at the center of the restaurant. We also noticed that most of the upscale restaurants are located in the Borough of Toronto Downtown.

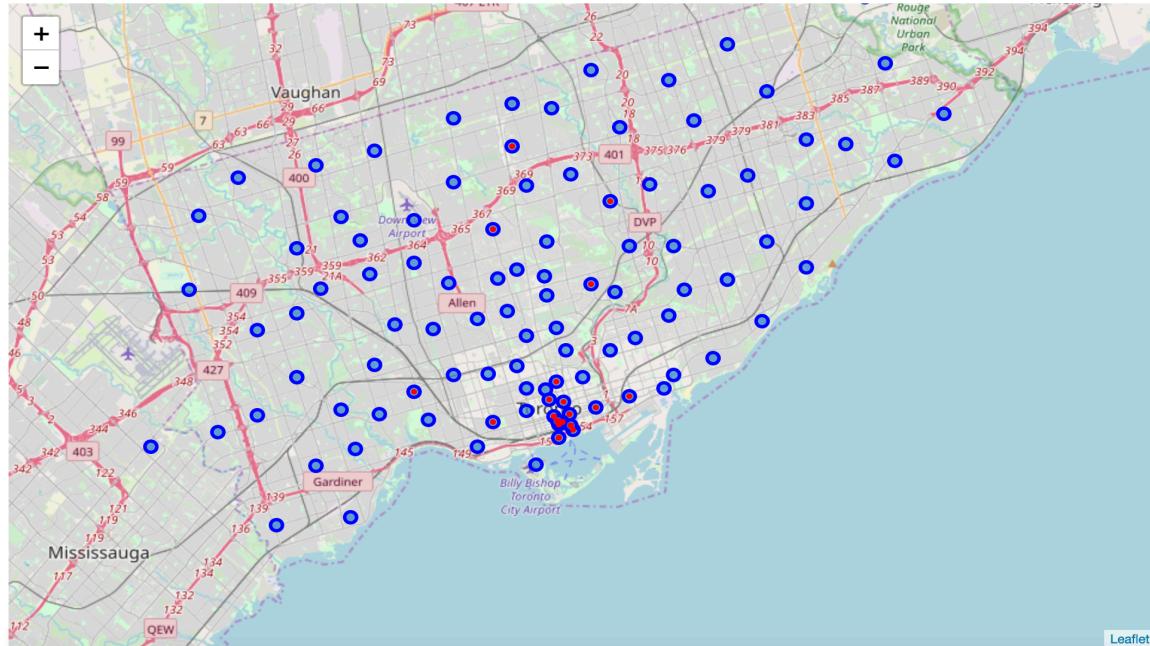
Let's now plot all upscale restaurant locations from above into a folium map to see their exact locations (in red) and then compare them with each neighborhood (in blue).

```
map_toronto = folium.Map(location=[latitude_toronto, longitude_toronto], zoom_start=10)

# add neighborhood markers to map
for lat, lng, borough, Neighbourhood in zip(toronto_data['Latitude'], toronto_data['Longitude'], toronto_data['Borough'],
label = '{}, {}'.format(Neighbourhood, borough)
label = folium.Popup(label, parse_html=True)
folium.CircleMarker(
    [lat, lng],
    radius=5,
    popup=label,
    color='blue',
    fill=True,
    fill_color='#3186cc',
    fill_opacity=0.7,
    parse_html=False).add_to(map_toronto)

# add upscale restaurant markers to map
for lat, lng, Neighbourhood in zip(toronto_venues['Neighbourhood Latitude'], toronto_venues['Neighbourhood Longitude'], toronto_venues['Neighbourhood']):
label = '{}'.format(Neighbourhood)
label = folium.Popup(label, parse_html=True)
folium.CircleMarker(
    [lat, lng],
    radius=1,
    popup=label,
    color='red',
    fill=False,
    parse_html=False).add_to(map_toronto)

map_toronto
```



Let's find how many upscale restaurants are there per neighborhood.

```
toronto_venues['Neighbourhood'].value_counts().to_frame(name='Count')
```

	Count
Design Exchange, Toronto Dominion Centre	11
First Canadian Place, Underground city	10
Commerce Court, Victoria Hotel	10
Adelaide, King, Richmond	9
Harbourfront East, Toronto Islands, Union Station	4
St. James Town	3
Stn A PO Boxes 25 The Esplanade	3
Bedford Park, Lawrence Manor East	2
Harbourfront, Regent Park	2
Willowdale South	2
Little Portugal, Trinity	1
Berczy Park	1
Ryerson, Garden District	1
Don Mills North	1
Leaside	1
High Park, The Junction South	1
Church and Wellesley	1
Studio District	1
Central Bay Street	1

Discussion

The top 4 neighborhoods from the list above that have the highest number of upscale restaurants per postal code are all located in Downtown Toronto:

- Design Exchange, Toronto Dominion Centre (M5K): 11 upscale restaurants
- First Canadian Place, Underground city (M5X): 10 upscale restaurants
- Commerce Court, Victoria Hotel (M5L): 10 upscale restaurants
- Adelaide, King, Richmond (M5H): 9 upscale restaurants

A new upscale Brazilian restaurant could be opened at any of these neighborhoods.

Conclusion

Stakeholders were able to find 4 (four) best neighborhood locations per postal code for them to open their new Brazilian restaurant in Toronto Downtown.

Further analysis to gather profitability, popularity and property prices for these areas should take place to create a weighting matrix. After this matrix is developed, we could select an optimal location for a new upscale Brazilian steakhouse in Toronto Downtown.