LAUNCHING OREGANO IN CANADA

1. BUSINESS PROBLEM

1.1. Background

A successful Italian restaurant chain Oregano, with more than 20 restaurants in Italy is planning to expand its business to Canada. The Sales & Marketing division of Oregano conducted extensive market research in Canada and found that there is a high demand for Italian food, especially in Toronto. Oregano has now hired an analytics firm Factual Analytics to help them figure out the best neighbourhood in Toronto, to open their first restaurant.

1.2. Objective

Oregano wants to figure out the best locality in Toronto with high demand for Italian food and lesser number of competitors, to rapidly grow their business, by leveraging the data on the neighbourhoods in Toronto. Additionally, Oregano also wants to ensure that the locality is safe and has less crime rate. The team of data scientists at Factual Analytics would apply cutting edge data science methodologies and machine learning techniques on the data to derive meaningful business insights, which would then be used by Oregano to make informed decisions.

1.3. Target Audience

The key stakeholders from Oregano are the CEO, Board of Directors, Sales & Marketing Team, and Head of Finance. Additionally, the team of Data Analysts, Engineers and Scientists are also involved.

2. DATA DESCRIPTION

For this project, the data engineering team at Factual Analytics would be extracting information from multiple data sources, as listed below. This data would then be collated and analysed by the data analysts and scientists.

I. Wikipedia

The information on neighbourhood boroughs in Toronto and their respective postal codes would be extracted from the Wikipedia page. The data on this page would be scraped using the python library BeautifulSoup. The scraped data would then be loaded to a data frame, which would look like the snapshot below.

	PostalCode	Borough	Neighborhood
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	Regent Park, Harbourfront

II. Public Safety Data Portal

The data on crime rates in Toronto is available in the Public Safety Data Portal, managed by the Toronto Police Service. The information of major crime indicators (MCI) is available here, as a CSV file. A snapshot of the data is given below.

Index	event_unique_id	Division	occurrencedate	reporteddate	premises_type	ucr_code	ucr_ext	offence	r
10182	GO-20142103239	D23	18/05/2014, 05:00	18/05/2014, 05:00	House	1430	100	Assault	201
10298	GO-20142111528	D23	19/05/2014, 05:00	19/05/2014, 05:00	Outside	1430	100	Assault	201
10300	GO-20142111859	D23	19/05/2014, 05:00	19/05/2014, 05:00	Apartment	2120	200	B&E	201
10345	GO-20142116041	D23	20/05/2014, 05:00	20/05/2014, 05:00	Outside	1420	100	Assault With Weapon	201
10346	GO-20142116041	D23	20/05/2014, 05:00	20/05/2014, 05:00	Outside	1420	110	Assault Bodily Harm	201
10347	GO-20142116041	D23	20/05/2014, 05:00	20/05/2014, 05:00	Outside	1420	110	Assault Bodily Harm	201

III. Geocoder

This python package would be used to obtain the geographical co-ordinates (latitude and longitude) for the Toronto neighbourhoods and boroughs, shown below.

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.75245	-79.32991
1	M4A	North York	Victoria Village	43.73057	-79.31306
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.65512	-79.36264
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.72327	-79.45042
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.66253	-79.39188

IV. Foursquare API

The popular and common venues in the Toronto neighbourhoods would be explored through the Foursquare API. The API request returns the data in a JSON format, as shown below.


```
{
"meta": {
  "code": 200,
  "requestId": "5ac51d7e6a607143d811cecb"
 },
 "response": {
  "venues": [
   {
    "id": "5642aef9498e51025cf4a7a5",
    "name": "Mr. Purple",
    "location": {
     "address": "180 Orchard St",
     "crossStreet": "btwn Houston & Stanton St",
     "lat": 40.72173744277209,
     "lng": -73.98800687282996,
     "labeledLatLngs": [
      {
```

```
      "label": "display",
      "lat": 40.72173744277209,
      "lng": -73.98800687282996
    }
  ],
  "distance": 8,
  "postalCode": "10002",
  "cc": "US",
  "city": "New York",
  "state": "NY",
  "country": "United States",
  "formattedAddress": [
    "180 Orchard St (btwn Houston & Stanton St)",
    "New York, NY 10002",
    "United States"
  ]
},
"categories": [
  {
```

```
        "id": "4bf58dd8d48988d1d5941735",
        "name": "Hotel Bar",
        "pluralName": "Hotel Bars",
        "shortName": "Hotel Bar",
        "icon": {
          "prefix": "https://ss3.4sqi.net/img/categories_v2/travel/hotel_bar_",
          "suffix": ".png"
        },
        "primary": true
      }
    ],
    "venuePage": {
      "id": "150747252"
    }
  }
 ]
}
}
```