The Battle of Neighborhoods (Week 2) by Dr K.R.S Dulanjalee

Business Recommendations for Surrey, BC, Canada.

1. Introduction.

- Who would be interested?

Anyone who is interested in starting new business but who has no idea regarding which is the field he should step into and whose convenient geolocation is near to surrey will benefit from this.

-Why Surrey?

Surrey is a city in the province of British Columbia, Canada, located south of the Fraser River and north of the Canada–United States border. It is a member municipality of the Metro Vancouver regional district and metropolitan area. Mainly a suburban city, Surrey is the province's second largest by population after Vancouver and the third largest by area after Abbotsford and Prince George. Seven neighborhoods in Surrey are designated town centers: Cloverdale, Fleetwood, Guildford, Newton, South Surrey, and City Centre encompassed by Whalley.

In 2016 the population was recorded at 517,887, an increase of 10.6% from 2011. This made it the 12th largest city in Canada, while also being the fifth-largest city in Western Canada (after Calgary, Edmonton, Winnipeg, and Vancouver). Surrey forms an integral part of Metro Vancouver as it is the second largest city in the region, albeit while also serving as the secondary economic core of the metropolitan area. When combined with the City of Vancouver, both cities account for nearly 50% of the region's population. In recent years, a rapidly expanding urban core in Downtown Surrey, located in Whalley has transformed the area into the secondary downtown core in Metro Vancouver.

Canada 2016 Census	Population	% of total population
South Asian	168,040	32.4%
Chinese	39,890	7.7%
Filipino	31,865	6.2%
Southeast Asian	13,010	2.5%
Black	9,455	1.8%
Korean	9,285	1.8%
Latin American	7,065	1.4%
Arab	5,765	1.1%

Canada 2016 Census	Population	% of total population
West Asian	3,720	0.7%
Japanese	2,850	0.6%
Another visible minority	1,995	0.4%
Mixed visible minority	6,320	1.2%
Total visible minority population	299,245	57.8%
Total Aboriginal population	13,460	2.6%
Total population	517,887	100%

The average wage in Surrey is bolstered by the high proportion of residents who work in financial services. Surrey has more organization and company headquarters than any other county in the UK. Electronics manufacturers Whirlpool, Canon, Toshiba, Samsung, and Philips are housed here, as are distributors Bur lodge, Future Electronics, Kia Motors and Toyota UK, the medico-pharma companies Pfizer and Sanofi-Aventis and oil giant Esso. Some of the largest fast-moving consumer goods multinationals in the world have their UK and/or European headquarters here, including Unilever, Procter & Gamble, Superdrug, Nestlé, SC Johnson, Kimberly-Clark, and Colgate-Palmolive. NGOs including WWF UK & Compassion in World Farming are also based here. Government Quangos such as SEEDA, SEERA and GOSE are headquartered in Guildford.

2. Data

Data is mainly extracted from Geocoder - https://geocoder.ca/?services=1&showupdates=1 and Foursquare - https://foursquare.com/.

-how data will be used to solve the problem

   a. **Step One**

   Preprocessing the data set of British Columbia, Canada.

-import relevant libraries like numpy, pandas, matplotlib, pylab, seaborn, scipy, geopy, sklear, folium.

1. Import relevant libraries

```
In [367]:  import numpy as np

           import pandas as pd
           pd.set_option('display.max_columns', None)
           pd.set_option('display.max_rows', None)

           import matplotlib as mpl
           import matplotlib.pyplot as plt
           import matplotlib.cm as cm
           import matplotlib.colors as colors

           import pylab as pl
           import seaborn as sns
           import scipy.optimize as opt
           %matplotlib notebook
           %matplotlib inline


           import json

           !pip install geopy
           from geopy.geocoders import Nominatim

           import requests
           from pandas.io.json import json_normalize

           from sklearn.cluster import KMeans

           !pip install folium
           import folium
```

-Load csv data file from personal computer which was prepared in association with Geocoder.

2. Load Data and make a Data Frame

```
In [368]:  import types
           import pandas as pd
           from 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 the notebook.
           client_ccaad49133d14e40883f2bb130ee6f83 = ibm_boto3.client(service_name='s3',
               ibm_api_key_id='t-JdKrw2gybzjwlklHKZ23W6f1coLpq7TIKWeuDZ1FiO',
               ibm_auth_endpoint="https://iam.cloud.ibm.com/oidc/token",
               config=Config(signature_version='oauth'),
               endpoint_url='https://s3-api.us-geo.objectstorage.service.networklayer.com')

           body = client_ccaad49133d14e40883f2bb130ee6f83.get_object(Bucket='courseracapstoneproject-donotdelete-pr-nrnp2be065vtc8',Key='Co
           ursera.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 )

           BC_df = pd.read_csv(body)
           BC_df.head()
```

Out[368]:

	Index	Postal Codes	Latitudes	Longitudes	Borough
0	1	V0A1K1	50.339434	-115.858121	Fairmont Hot Springs
1	2	V0B0V0	49.218994	-122.949596	Burnaby
2	3	V0C2B0	56.247431	-120.879999	Fort St. John
3	4	V0E2Z7	50.711182	-119.261332	Salmon Arm
4	6	V0H0H0	49.160556	-123.114890	Richmond

-Preprocessing

```
In [369]:  BC_df.shape
Out[369]:  (278, 5)
```

```
In [370]:  BC_df.columns
Out[370]:  Index(['Index', 'Postal Codes', 'Latitudes', 'Longitudes', 'Borough'], dtype='object')
```

```
In [371]:  BC_df = BC_df.drop(['Index'], axis = 1)
           BC_df.head()
```

Out[371]:

	Postal Codes	Latitudes	Longitudes	Borough
0	V0A1K1	50.339434	-115.858121	Fairmont Hot Springs
1	V0B0V0	49.218994	-122.949596	Burnaby
2	V0C2B0	56.247431	-120.879999	Fort St. John
3	V0E2Z7	50.711182	-119.261332	Salmon Arm
4	V0H0H0	49.160556	-123.114890	Richmond

```
In [372]:  A = BC_df[BC_df['Borough'] == 'Campbell River']
           A
```

Out[372]:

	Postal Codes	Latitudes	Longitudes	Borough
263	V9H0E6	49.960339	-125.236511	Campbell River
264	V9H0E7	49.961222	-125.238884	Campbell River
270	V9W0E9	45.730295	-76.603025	Campbell River
271	V9W3S4	49.906397	-125.195014	Campbell River

```
In [372]: A = BC_df[BC_df['Borough'] == 'Campbell River']
          A
```

Out[372]:

	Postal Codes	Latitudes	Longitudes	Borough
263	V9H0E6	49.960339	-125.236511	Campbell River
264	V9H0E7	49.961222	-125.238884	Campbell River
270	V9W0E9	45.730295	-76.603025	Campbell River
271	V9W3S4	49.906397	-125.195014	Campbell River

```
In [373]: BC_df = BC_df.drop([263, 264, 270,271])
          BC_df.tail(30)
```

Out[373]:

	Postal Codes	Latitudes	Longitudes	Borough
244	V7X0P4	49.186800	-123.132399	Richmond
245	V7X1J1	49.286137	-123.119745	Vancouver
246	V7X1L7	49.286136	-123.119745	Vancouver
247	V8A0C7	49.886060	-124.544084	Powell River
248	V8B0E3	49.732770	-123.141706	Squamish
249	V8E0Y0	50.120881	-122.936620	Whistler
250	V8J4N7	54.310921	-130.304363	Prince Rupert
251	V8L0H9	48.411411	-123.495567	Colwood
252	V8S0C3	49.699478	-123.154697	Squamish
253	V8T0E1	48.448495	-123.336286	Victoria
254	V8T1E4	48.430694	-123.363411	Victoria

```
In [374]: B = BC_df[BC_df['Borough'] == 'Campbell River']
          B
```

Out[374]:

	Postal Codes	Latitudes	Longitudes	Borough

```
In [375]: BC_df.dtypes
```

Out[375]:
```
Postal Codes      object
Latitudes         float64
Longitudes        float64
Borough           object
dtype: object
```

b. Step Two

Visualization of the data set of British Columbia, Canada.

3. Visualization

```
In [376]: address = 'British Columbia'

          geolocator = Nominatim(user_agent="British Columbia_explorer")
          location = geolocator.geocode(address)
          latitude = location.latitude
          longitude = location.longitude
          print('The geograpical coordinate of British Columbia are {}, {}.'.format(latitude, longitude))

          The geograpical coordinate of British Columbia are 55.001251, -125.002441.
```
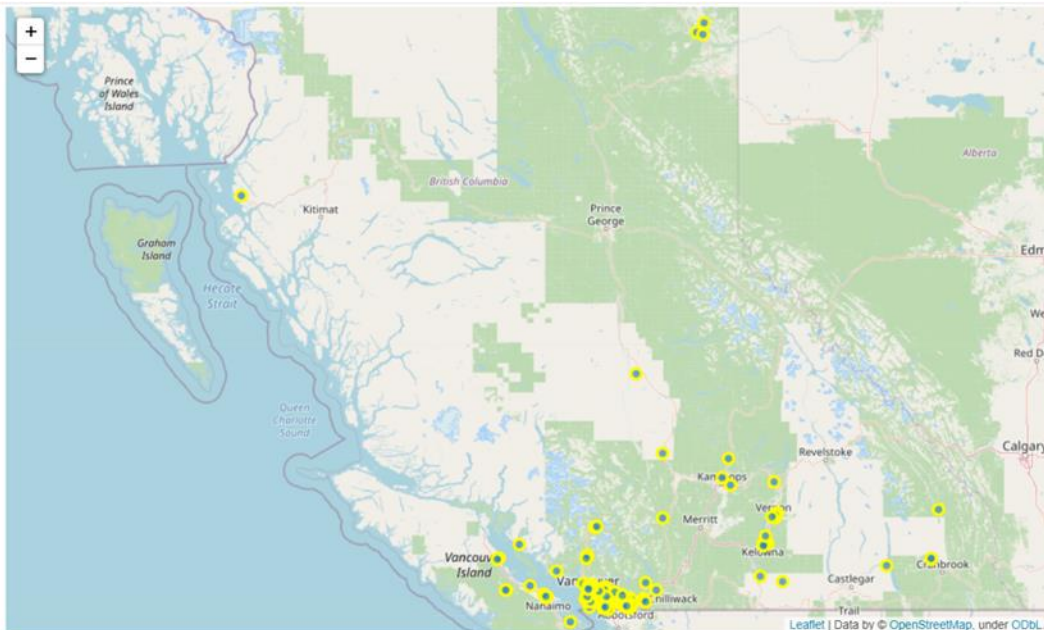
```
In [377]: map_BritishColumbia = folium.Map(location=[latitude, longitude], zoom_start=5)

          for lat, lng, borough in zip(BC_df['Latitudes'], BC_df['Longitudes'], BC_df['Borough']):
              label = '{}'.format(borough)
              label = folium.Popup(label, parse_html=True)
              folium.CircleMarker(
                  [lat, lng],
                  radius=5,
                  popup=label,
                  color='yellow',
                  fill=True,
                  fill_color='#3186cc',
                  fill_opacity=0.7,
                  parse_html=False).add_to(map_BritishColumbia)

          map_BritishColumbia
```

Out[377]:

c. Step Two

Exploring the top 10 frequently visiting venues in Surrey, BC, Canada.

-Neighborhood in Surrey

3. Visualization

```
In [376]: address = 'British Columbia'

          geolocator = Nominatim(user_agent="British Columbia_explorer")
          location = geolocator.geocode(address)
          latitude = location.latitude
          longitude = location.longitude
          print('The geograpical coordinate of British Columbia are {}, {}.'.format(latitude, longitude))

          The geograpical coordinate of British Columbia are 55.001251, -125.002441.
```
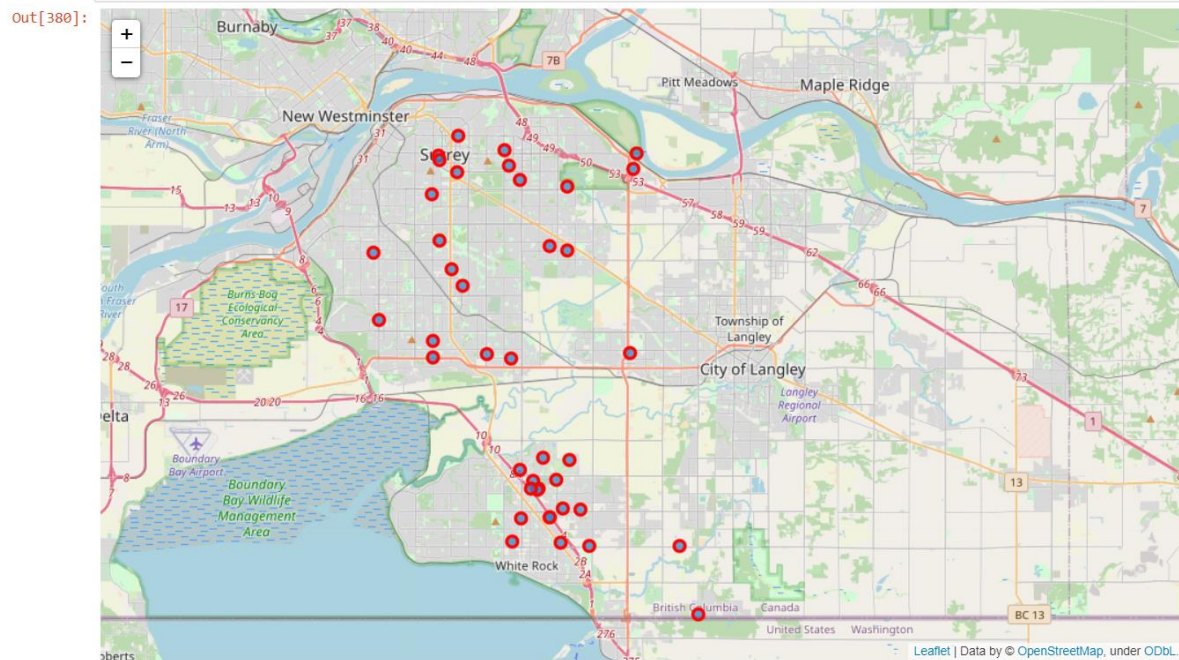
```
In [377]: map_BritishColumbia = folium.Map(location=[latitude, longitude], zoom_start=5)

          for lat, lng, borough in zip(BC_df['Latitudes'], BC_df['Longitudes'], BC_df['Borough']):
              label = '{}'.format(borough)
              label = folium.Popup(label, parse_html=True)
              folium.CircleMarker(
                  [lat, lng],
                  radius=5,
                  popup=label,
                  color='yellow',
                  fill=True,
                  fill_color='#3186cc',
                  fill_opacity=0.7,
                  parse_html=False).add_to(map_BritishColumbia)

          map_BritishColumbia
```

Out[380]:

-Connecting to Foursquare

```
In [384]: LIMIT = 100
          radius = 1000
          url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
              CLIENT_ID,
              CLIENT_SECRET,
              VERSION,
              V3S0S9_latitude,
              V3S0S9_longitude,
              radius,
              LIMIT)
          url
```

-Table with nearest venues

	Neighborhood Postal Code	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	V2S0Z5	49.186697	-122.809079	Disney store	49.188895	-122.803649	Toy / Game Store
1	V2S0Z5	49.186697	-122.809079	Apple Guildford Town Centre	49.189299	-122.803620	Electronics Store
2	V2S0Z5	49.186697	-122.809079	Hanako	49.186283	-122.802545	Sushi Restaurant
3	V2S0Z5	49.186697	-122.809079	Landmark Cinemas 12 Guildford	49.186949	-122.804454	Multiplex
4	V2S0Z5	49.186697	-122.809079	Akasaka	49.185884	-122.803315	Japanese Restaurant
5	V2S0Z5	49.186697	-122.809079	The Lego Store	49.189312	-122.803782	Toy / Game Store
6	V2S0Z5	49.186697	-122.809079	H&M	49.189090	-122.804715	Clothing Store
7	V2S0Z5	49.186697	-122.809079	Purdys Chocolatier	49.189510	-122.804908	Gift Shop
8	V2S0Z5	49.186697	-122.809079	London Drugs	49.189902	-122.805193	Pharmacy
9	V2S0Z5	49.186697	-122.809079	McDonald's	49.188579	-122.805013	Fast Food Restaurant
10	V2S0Z5	49.186697	-122.809079	New York Fries	49.189885	-122.804334	Fast Food Restaurant
11	V2S0Z5	49.186697	-122.809079	EB Games	49.189377	-122.804423	Video Game Store
12	V2S0Z5	49.186697	-122.809079	Boston Pizza	49.184491	-122.803380	Pizza Place
13	V2S0Z5	49.186697	-122.809079	Walmart	49.188312	-122.804918	Big Box Store
14	V2S0Z5	49.186697	-122.809079	The Pantry	49.190215	-122.804895	Diner
15	V2S9J2	49.107889	-122.807934	YMCA Surrey	49.105313	-122.806590	Gym

-Table with venues and its type

	Neighborhood Postal Code	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	V2S0Z5	49.186697	-122.809079	Disney store	49.188895	-122.803649	Toy / Game Store
1	V2S0Z5	49.186697	-122.809079	Apple Guildford Town Centre	49.189299	-122.803620	Electronics Store
2	V2S0Z5	49.186697	-122.809079	Hanako	49.186283	-122.802545	Sushi Restaurant
3	V2S0Z5	49.186697	-122.809079	Landmark Cinemas 12 Guildford	49.186949	-122.804454	Multiplex
4	V2S0Z5	49.186697	-122.809079	Akasaka	49.185884	-122.803315	Japanese Restaurant

d. Step Three

Visualize the frequency of top 10 visiting venues in a sunburst chart.

In [402]:

	venue	freq	Neighborhood Postal Code
0	Toy / Game Store	0.13	V2S0Z5
1	Fast Food Restaurant	0.13	V2S0Z5
2	Pharmacy	0.07	V2S0Z5
3	Japanese Restaurant	0.07	V2S0Z5
4	Multiplex	0.07	V2S0Z5
5	Pizza Place	0.07	V2S0Z5
6	Clothing Store	0.07	V2S0Z5
7	Electronics Store	0.07	V2S0Z5
8	Gift Shop	0.07	V2S0Z5
9	Diner	0.07	V2S0Z5
10	Gym	0.29	V2S9J2
11	Supermarket	0.14	V2S9J2
12	Pharmacy	0.14	V2S9J2
13	Sandwich Place	0.14	V2S9J2
14	Japanese Restaurant	0.14	V2S9J2
15	Park	0.14	V2S9J2
16	Ice Cream Shop	0.00	V2S9J2
17	Indian Restaurant	0.00	V2S9J2
18	Pakistani Restaurant	0.00	V2S9J2
19	Health Food Store	0.00	V2S9J2

In [403]: `temp_df_append.shape`

Out[403]: (340, 3)

In [404]:
```python
fig=px.sunburst(temp_df_append, path=['Neighborhood Postal Code', 'venue'], values = 'freq')
fig.show()
```

e. Step Four

Suggest clusters of neighborhoods who has same business tastes in Surrey, BC, Canada

[400]:

	Postal Codes	Latitudes	Longitudes	Borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th M Comm Ver
0	V2S0Z5	49.186697	-122.809079	Surrey	1.0	Toy / Game Store	Fast Food Restaurant	Electronics Store	Diner	Pizza Place	Clothing Store	Pharmacy
1	V2S9J2	49.107889	-122.807934	Surrey	0.0	Gym	Japanese Restaurant	Park	Pharmacy	Supermarket	Sandwich Place	Donut Sho
2	V2W5Y4	49.144302	-122.845344	Surrey	1.0	Electronics Store	Gas Station	Convenience Store	Hotel	Donut Shop	Wine Shop	Construct & Landscap
3	V3A8X9	49.043329	-122.784200	Surrey	1.0	Bank	Burger Joint	Coffee Shop	Restaurant	Grocery Store	Sandwich Place	Big Box Store
4	V3E6Y6	49.156208	-122.852509	Surrey	1.0	Pizza Place	Carpet Store	Storage Facility	Wine Shop	Donut Shop	Construction & Landscaping	Convenie Store
5	V3N0N6	49.058459	-122.779531	Surrey	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6	V3R0L3	49.184033	-122.841478	Surrey	1.0	Coffee Shop	Restaurant	Greek Restaurant	Breakfast Spot	Discount Store	Park	Paper / Office Supplies Store
7	V3R9M9	49.180639	-122.802458	Surrey	4.0	Convenience Store	Pizza Place	Restaurant	Wine Shop	Donut Shop	Construction & Landscaping	Deli / Bodega
8	V3S0H5	49.066744	-122.771797	Surrey	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	V3S0M6	49.046999	-122.775315	Surrey	1.0	Coffee Shop	Grocery Store	Restaurant	Bank	Bookstore	Breakfast Spot	Clothing Store

12]:
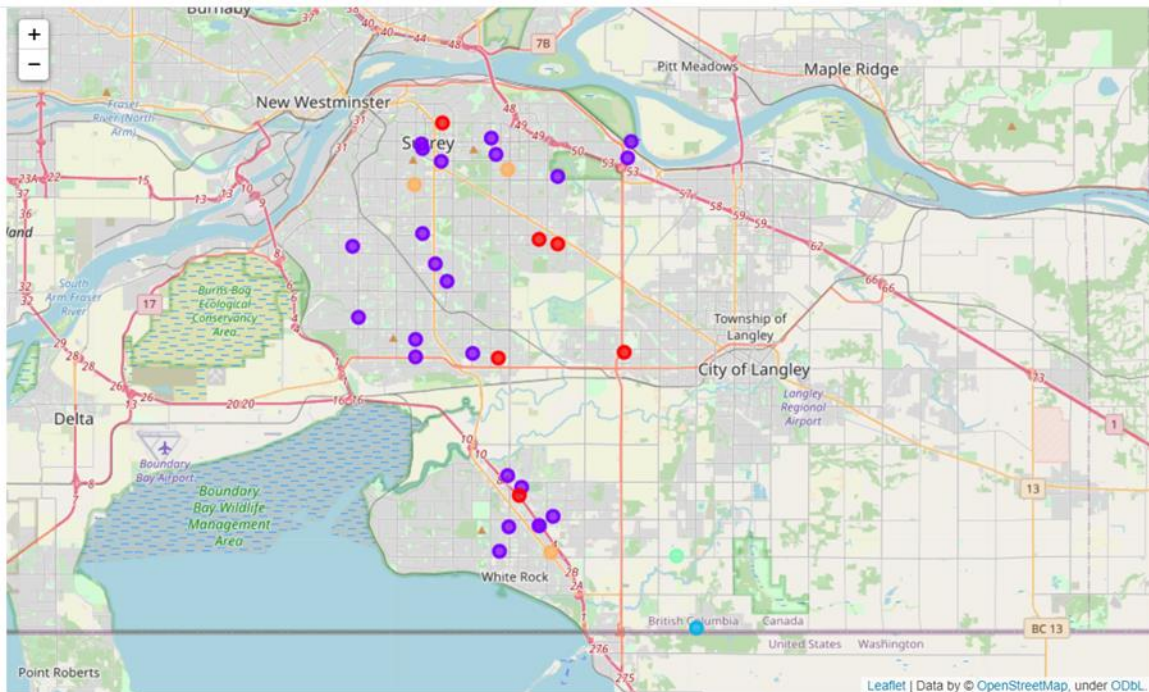
Cluster 1

```
[13]: surrey_merged.loc[surrey_merged['Cluster Labels'] == 0, surrey_merged.columns[[0] + list(range(5, surrey_merged.shape[1]))]]
```

[13]:

	Postal Codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	V2S9J2	Gym	Japanese Restaurant	Park	Pharmacy	Supermarket	Sandwich Place	Donut Shop	Construction & Landscaping	Convenience Store	Deli / Bodega
15	V3S8T5	Wine Shop	Pharmacy	Indian Restaurant	Park	Athletics & Sports	Coffee Shop	Convenience Store	Deli / Bodega	Department Store	Dim Sum Restaurant
16	V3S8W2	Bank	Pakistani Restaurant	Sandwich Place	Gym / Fitness Center	Korean Restaurant	Market	Chinese Restaurant	Pharmacy	Salon / Barbershop	Gastropub
20	V3T2K7	Korean Restaurant	Sandwich Place	Mexican Restaurant	Malay Restaurant	Bus Stop	Construction & Landscaping	Food & Drink Shop	Flea Market	Convenience Store	Deli / Bodega
31	V4B4Z7	Cheese Shop	Bakery	Bank	Portuguese Restaurant	Japanese Restaurant	Wine Shop	Electronics Store	Deli / Bodega	Department Store	Dim Sum Restaurant
33	V4M4V8	American Restaurant	Flea Market	Casino	Lounge	Bowling Alley	Convenience Store	Athletics & Sports	Gastropub	Gas Station	Doctor's Office

Cluster 2

```
[4]: surrey_merged.loc[surrey_merged['Cluster Labels'] == 1, surrey_merged.columns[[0] + list(range(5, surrey_merged.shape[1]))]]
```

[4]:

	Postal Codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	V2S0Z5	Toy / Game Store	Fast Food Restaurant	Electronics Store	Diner	Pizza Place	Clothing Store	Pharmacy	Big Box Store	Japanese Restaurant	Sushi Restaurant
2	V2W5Y4	Electronics Store	Gas Station	Convenience Store	Hotel	Donut Shop	Wine Shop	Construction & Landscaping	Deli / Bodega	Department Store	Dim Sum Restaurant
3	V3A8X9	Bank	Burger Joint	Coffee Shop	Restaurant	Grocery Store	Sandwich Place	Big Box Store	Café	Electronics Store	Fast Food Restaurant
4	V3E6Y6	Pizza Place	Carpet Store	Storage Facility	Wine Shop	Donut Shop	Construction & Landscaping	Convenience Store	Deli / Bodega	Department Store	Dim Sum Restaurant

Cluster 3

```
[415]: surrey_merged.loc[surrey_merged['Cluster Labels'] == 2, surrey_merged.columns[[0] + list(range(5, surrey_merged.shape[1]))]]
```

[415]:

	Postal Codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
21	V3T4B6	Auto Garage	Wine Shop	Farmers Market	Convenience Store	Deli / Bodega	Department Store	Dim Sum Restaurant	Diner	Discount Store	Doctor's Office

Cluster 4

```
[416]: surrey_merged.loc[surrey_merged['Cluster Labels'] == 3, surrey_merged.columns[[0] + list(range(5, surrey_merged.shape[1]))]]
```

[416]:

	Postal Codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
17	V3S9W7	Stables	Wine Shop	Electronics Store	Construction & Landscaping	Convenience Store	Deli / Bodega	Department Store	Dim Sum Restaurant	Diner	Discount Store

Cluster 5

```
In [417]: surrey_merged.loc[surrey_merged['Cluster Labels'] == 4, surrey_merged.columns[[0] + list(range(5, surrey_merged.shape[1]))]]
Out[417]:
```

	Postal Codes	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	V3R9M9	Convenience Store	Pizza Place	Restaurant	Wine Shop	Donut Shop	Construction & Landscaping	Deli / Bodega	Department Store	Dim Sum Restaurant	Diner
25	V3W5R2	Convenience Store	Wine Shop	Vietnamese Restaurant	Construction & Landscaping	Deli / Bodega	Department Store	Dim Sum Restaurant	Diner	Discount Store	Doctor's Office
32	V4B4Z8	American Restaurant	Convenience Store	Deli / Bodega	Bakery	Farmers Market	Department Store	Dim Sum Restaurant	Diner	Discount Store	Doctor's Office

2. Methodology

data analysis – Sunburst chart

machine learnings – K mean clustering.

3. Results

Anyone who is interested in starting new business but who has no idea regarding which is the field he should step into and whose convenient geolocation is near to surrey will benefit from this.

Thank You.