Background

Jakarta, as the heart of economy of Indonesia, is an interesting and lucrative place for one to start their self owned business. Starting a business however, is never an easy task as it requires a lot of preparation and research to make sure the business runs well. Knowing current market situation and business map will be beneficial to help guarantee business sustainability. This capstone project will try to identify business that could possibly be beneficial to start in Jakarta.

Problem

Clustering most common venue in Jakarta, to indicate which area belongs to a cluster of typical venue.

Later, the results hopefully can help to give an overview of suitable business to start in Jakarta.

Data Used

1. List of districts in Jakarta and its coordinate.
2. Foursquare API to check common venue around the coordinate.

Data Exploratory

First, we find the list of district in Jakarta area. We can get it from https://id.wikipedia.org/wiki/Daftar_kecamatan_dan_kelurahan_di_Daerah_Khusus_Ibukota_Jakarta

There are 44 districts in Greater Jakarta area, however, in this project we exclude districts in Kepulauan Seribu since they are located in a separate island from Jakarta mainland. Then, by adding the longitude and the latitude of the district, we create a dataframe of it at python and it the result is as follows.

	City	District	Latitude	Longitude
0	South Jakarta	Cilandak	-6.289759	106.772842
1	South Jakarta	Jagakarsa	-6.33033	106.790444
2	South Jakarta	Kebayoran Baru	-6.243176	106.783831
3	South Jakarta	Kebayoran Lama	-6.249235	106.745362
4	South Jakarta	Mampang Prapatan	-6.250242	106.804506
5	South Jakarta	Pancoran	-6.257905	106.826132
6	South Jakarta	Pasar Minggu	-6.289773	106.821566
7	South Jakarta	Pesanggrahan	-6.254211	106.720168
8	South Jakarta	Setiabudi	-6.221734	106.812879
9	South Jakarta	Tebet	-6.22548	106.833254
10	West Jakarta	Cengkareng	-6.154878	106.702954
11	West Jakarta	Grogol Petamburan	-6.163922	106.769051
12	West Jakarta	Taman Sari	-6.146123	106.801865
13	West Jakarta	Tambora	-6.146278	106.783175
14	West Jakarta	Kebon Jeruk	-6.191465	106.730607
15	West Jakarta	Kalideres	-6.13606	106.67388
16	West Jakarta	Palmerah	-6.190951	106.778864

(Longlat dataframe. 42 rows, shown 16)

After getting the longitude and latitude, we connect to foursquare API and get the lists of venues around the coordinates. Then we create a new dataframe containing the data of venue.

```
jkt_venues = getNearbyVenues(names=df_jktlatlong['District'],
                             latitudes=df_jktlatlong['Latitude'],
                             longitudes=df_jktlatlong['Longitude'])
```

```
In [9]: jkt_venues.head()
```

Out[9]:

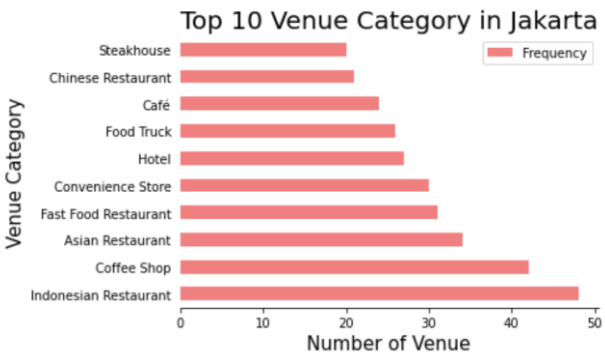
	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Cilandak	-6.289759	106.772842	Sate Padang Talago Biru	-6.288731	106.770937	Food Truck
1	Cilandak	-6.289759	106.772842	Okirobox @ Carrefour Lb. Bulus	-6.287761	106.776271	Japanese Restaurant
2	Cilandak	-6.289759	106.772842	Chatime	-6.287720	106.776384	Bubble Tea Shop
3	Cilandak	-6.289759	106.772842	Imperial Kitchen & Dimsum	-6.287811	106.776385	Asian Restaurant
4	Cilandak	-6.289759	106.772842	Carrefour	-6.287690	106.776346	Supermarket

(Define a new function namely getNearbyVenues, which covers connecting to Foursquare API and using its explore coordinates function)

From the new dataframe, we can then group the data by venue category, and examine what the top 10 Venue Category available in Jakarta are. The result will be as follows.

	Venue Category	Frequency
0	Indonesian Restaurant	48
1	Coffee Shop	42
2	Asian Restaurant	34
3	Fast Food Restaurant	31
4	Convenience Store	30
...	...	...
152	Food Stand	1
153	Udon Restaurant	1
154	Men's Store	1
155	Rental Car Location	1
156	Hookah Bar	1

157 rows × 2 columns



From the Foursquare data, there are 157 unique venue category available in Jakarta, with **culinary business as the most common venue** currently available. The top 3 most frequent available venue is Indonesian Restaurant (48), Coffee Shop (42), and Asian Restaurant (34).

After getting those general overview, we can then start to try creating a cluster of business venue in Jakarta.

Clustering

We start clustering by first doing one hot encoding to the dataset.

```
jkt_onehot = pd.get_dummies(jkt_venues[['Venue Category']], prefix="", prefix_sep="")
jkt_onehot.insert(loc=0, column='District', value=jkt_venues['District'] )
jkt_onehot.shape
```

```
t[20]: (755, 158)
```

```
jkt_onehot
```

t[21]:

	District	Accessories Store	Acehnese Restaurant	Airport Food Court	American Restaurant	Arcade	Art Gallery	Arts & Crafts Store	Asian Restaurant	BBQ Joint	...
0	Cilandak	0	0	0	0	0	0	0	0	0	...
1	Cilandak	0	0	0	0	0	0	0	0	0	...
2	Cilandak	0	0	0	0	0	0	0	0	0	...
3	Cilandak	0	0	0	0	0	0	0	1	0	...
4	Cilandak	0	0	0	0	0	0	0	0	0	...
...	...	...	...	...	...	...	...	...	...	...	...
750	Tanah Abang	0	0	0	0	0	0	0	0	0	...
751	Tanah Abang	0	0	0	0	0	0	0	0	0	...
752	Tanah Abang	0	0	0	0	0	0	0	0	0	...
753	Tanah Abang	0	0	0	0	0	0	0	0	0	...
754	Tanah Abang	0	0	0	0	0	0	0	0	0	...

755 rows × 158 columns

We then average the number of appearance of the venue and then sorting it to get the most common venue showing in each district. The result is as follows.

```
In [24]: num_top_venues = 10

indicators = ['st', 'nd', 'rd']

columns = ['District']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

districts_venues_sorted = pd.DataFrame(columns=columns)
districts_venues_sorted['District'] = jkt_group['District']

for ind in np.arange(jkt_group.shape[0]):
    districts_venues_sorted.iloc[ind, 1:] = return_most_common_venues(jkt_group.iloc[ind, :], num_top_venues)

districts_venues_sorted.head()
```

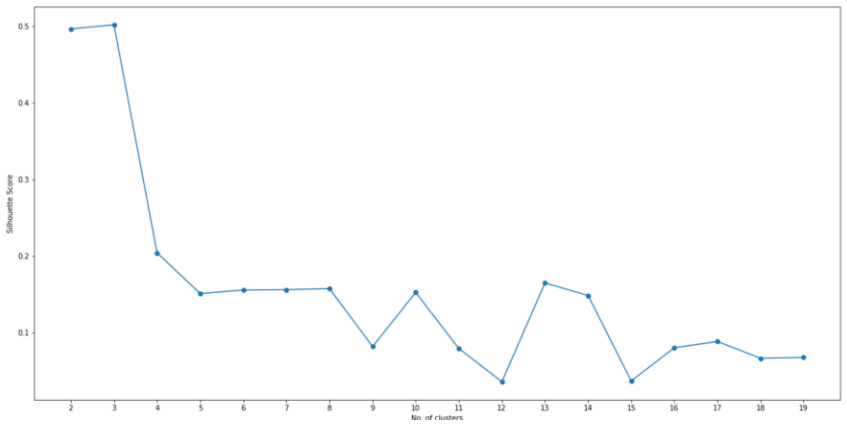
Out[24]:

	District	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	Cakung	Convenience Store	Cosmetics Shop	Soccer Field	Pizza Place	Women's Store	Donut Shop	Farmers Market	Factory	Event Space	Electronics Store
1	Cempaka Putih	Convenience Store	Indonesian Restaurant	Food & Drink Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store	Eastern European Restaurant
2	Cengkareng	Indonesian Restaurant	Noodle House	Asian Restaurant	Café	Women's Store	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store
3	Cilandak	Coffee Shop	Asian Restaurant	Supermarket	Dog Run	Food & Drink Shop	Food Truck	Japanese Restaurant	Donut Shop	Steakhouse	Sundanese Restaurant
4	Cilincing	Indonesian Meatball Place	Fast Food Restaurant	Restaurant	Indonesian Restaurant	Women's Store	Farmers Market	Factory	Event Space	Electronics Store	Eastern European Restaurant

These most common venue later will be the feature of our model to get group them into clusters. To start our model, we use silhouette scores to indicate the optimal k for the cluster from the data. The result is as follows.

```
In [28]: from sklearn.metrics import silhouette_samples, silhouette_score

indices = []
scores = []

for kclusters in range(2, max_range) :

    # Run k-means clustering
    jgc = jakarta_grouped_clustering
    kmeans = KMeans(n_clusters = kclusters, init = 'k-means++', random_state = 0).fit_predict(jgc)

    # Gets the score for the clustering operation performed
    score = silhouette_score(jgc, kmeans)

    # Appending the index and score to the respective lists
    indices.append(kclusters)
    scores.append(score)
```

```
In [38]: plot(max_range, scores, "No. of clusters", "Silhouette Score")
```



Silhouette Scores

From the result, it suggests that 3 will be the optimal k for our cluster model.

Result

By creating 3 cluster, the model produces the following clusters:

Cluster #1:

```
In [34]: jakarta_merged.loc[jakarta_merged['Cluster Labels'] == 0, jakarta_merged.columns[[0] + [1] + list(range(5, jakarta_merged.shape[1]))]]
```

Out[34]:

	City	District	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 Most Common Venue	8th Most Common Venue
0	South Jakarta	Cilandak	0	Coffee Shop	Asian Restaurant	Supermarket	Dog Run	Food & Drink Shop	Food Truck	Japanese Restaurant	Donut Shop
2	South Jakarta	Kebayoran Baru	0	Steakhouse	Coffee Shop	Japanese Restaurant	Café	Dessert Shop	Bakery	BBQ Joint	Korean Restaurant
3	South Jakarta	Kebayoran Lama	0	Café	Food Truck	Food Court	Comfort Food Restaurant	Miscellaneous Shop	Women's Store	Farmers Market	Factory
4	South Jakarta	Mampang Prapatan	0	Coffee Shop	Food Truck	Indonesian Restaurant	Japanese Restaurant	BBQ Joint	Restaurant	Noodle House	Sushi Restaurant
5	South Jakarta	Pancoran	0	Farmers Market	Restaurant	Indonesian Restaurant	American Restaurant	Arcade	Art Gallery	Karaoke Bar	Asian Restaurant
6	South Jakarta	Pasar Minggu	0	Indonesian Restaurant	Asian Restaurant	Convenience Store	Soup Place	Food Court	Food & Drink Shop	Coffee Shop	Restaurant
7	South Jakarta	Pesanggrahan	0	Playground	Donut Shop	Dog Run	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store
8	South Jakarta	Setiabudi	0	Café	Fast Food Restaurant	Hotel	Gym / Fitness Center	Frozen Yogurt Shop	Music Venue	Department Store	Convenience Store
9	South Jakarta	Tebet	0	Coffee Shop	Café	Shopping Mall	Restaurant	Chinese Restaurant	Hotel	Dim Sum Restaurant	Asian Restaurant
10	West Jakarta	Cengkareng	0	Indonesian Restaurant	Noodle House	Asian Restaurant	Café	Women's Store	Fast Food Restaurant	Farmers Market	Factory
11	West Jakarta	Grogol Petamburan	0	Fast Food Restaurant	Coffee Shop	Bed & Breakfast	Chinese Restaurant	Soup Place	Women's Store	Eastern European Restaurant	Fish & Chips Shop
12	West Jakarta	Taman Sari	0	Food Truck	Garden	Convenience Store	Light Rail Station	Department Store	Train Station	Chinese Restaurant	Fast Food Restaurant
13	West Jakarta	Tambora	0	Chinese Restaurant	Vegetarian / Vegan Restaurant	Donut Shop	BBQ Joint	Convenience Store	Seafood Restaurant	Javanese Restaurant	Café
14	West Jakarta	Kebon Jeruk	0	Pizza Place	Sushi Restaurant	Gym / Fitness Center	Coffee Shop	Chinese Restaurant	Café	Eastern European Restaurant	Karaoke Bar	Japanese Restaurant
16	West Jakarta	Palmerah	0	Food Truck	High School	Asian Restaurant	Indonesian Restaurant	Café	Soccer Field	Pizza Place	Arts & Crafts Store	Cosmetics Shop
17	West Jakarta	Kembangan	0	Pharmacy	Donut Shop	Fast Food Restaurant	Steakhouse	Asian Restaurant	Indonesian Restaurant	Farmers Market	Factory	Event Space
18	North Jakarta	Cilincing	0	Indonesian Meatball Place	Fast Food Restaurant	Restaurant	Indonesian Restaurant	Women's Store	Farmers Market	Factory	Event Space	Electronics Store
19	North Jakarta	Kelapa Gading	0	Fast Food Restaurant	Men's Store	Japanese Restaurant	Women's Store	Food & Drink Shop	Farmers Market	Factory	Event Space	Electronics Store
20	North Jakarta	Koja	0	Bakery	Convenience Store	Food Truck	Women's Store	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store
21	North Jakarta	Pademangan	0	Convenience Store	Bookstore	Seafood Restaurant	Badminton Court	Women's Store	Eastern European Restaurant	Fast Food Restaurant	Farmers Market	Factory
22	North Jakarta	Penjaringan	0	Construction & Landscaping	Food Truck	Rental Car Location	Dog Run	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store
23	North Jakarta	Tanjung Priok	0	Theme Park	Sundanese Restaurant	Fast Food Restaurant	Convention Center	Bowling Alley	Park	Arts & Crafts Store	Public Art	Hotel
24	East Jakarta	Cakung	0	Convenience Store	Cosmetics Shop	Soccer Field	Pizza Place	Women's Store	Donut Shop	Farmers Market	Factory	Event Space
25	East Jakarta	Cipayung	0	Fast Food Restaurant	Indonesian Restaurant	Grocery Store	Factory	Stadium	Soccer Stadium	Women's Store	Farmers Market	Event Space
26	East Jakarta	Ciracas	0	Pizza Place	Food Truck	Noodle House	Dessert Shop	Dog Run	Farmers Market	Factory	Event Space	Electronics Store
27	East Jakarta	Duren Sawit	0	Fast Food Restaurant	Snack Place	Pharmacy	Convenience Store	Night Market	Coffee Shop	Park	Soup Place	Smoke Shop
28	East Jakarta	Jatinegara	0	Indonesian Restaurant	Asian Restaurant	Café	Steakhouse	Convenience Store	Grocery Store	Dessert Shop	Restaurant	Karaoke Bar
29	East Jakarta	Kramat Jati	0	Pizza Place	Indonesian Restaurant	Coffee Shop	Bubble Tea Shop	Asian Restaurant	Steakhouse	Department Store	Diner	Sandwich Place
30	East Jakarta	Makasar	0	Cheese Shop	Arcade	Steakhouse	Restaurant	Dog Run	Farmers Market	Factory	Event Space	Electronics Store
31	East Jakarta	Matraman	0	Indonesian Restaurant	Pizza Place	Asian Restaurant	Coffee Shop	Javanese Restaurant	Hotel	Bookstore	Mosque	Park
32	East Jakarta	Pasar Rebo	0	Convenience Store	Acehnese Restaurant	Pharmacy	Park	Department Store	Donut Shop	Fast Food Restaurant	Farmers Market	Factory
33	East Jakarta	Pulo Gadung	0	Pizza Place	Korean Restaurant	Fast Food Restaurant	Multiplex	Convenience Store	Southern / Soul Food Restaurant	Japanese Restaurant	Shopping Mall	Fried Chicken Joint
34	Central Jakarta	Cempaka Putih	0	Convenience Store	Indonesian Restaurant	Food & Drink Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store
35	Central Jakarta	Gambir	0	Fast Food Restaurant	Hotel	Indonesian Restaurant	Convenience Store	Breakfast Spot	Seafood Restaurant	Soup Place	Bookstore	Electronics Store
36	Central Jakarta	Johar Baru	0	Hotel	Soup Place	Donut Shop	Noodle House	Arcade	History Museum	Asian Restaurant	Indonesian Restaurant	Fast Food Restaurant
37	Central Jakarta	Kemayoran	0	Convenience Store	Fast Food Restaurant	Food Court	Train Station	Asian Restaurant	Hotel	Dessert Shop	Electronics Store	Food Truck
38	Central Jakarta	Menteng	0	Indonesian Restaurant	Food Truck	Coffee Shop	Park	Dessert Shop	Donut Shop	Breakfast Spot	Food Court	Electronics Store
39	Central Jakarta	Sawah Besar	0	Chinese Restaurant	Noodle House	Asian Restaurant	Coffee Shop	Fast Food Restaurant	Bakery	Indonesian Restaurant	Steakhouse	Eastern European Restaurant
40	Central Jakarta	Senen	0	Hotel	Indonesian Restaurant	Convenience Store	Japanese Restaurant	Coffee Shop	Padangnese Restaurant	Seafood Restaurant	Café	Asian Restaurant
41	Central Jakarta	Tanah Abang	0	Convenience Store	Women's Store	Restaurant	Food Truck	Fast Food Restaurant	Pizza Place	Deli / Bodega	Cultural Center	Department Store

Cluster #2:

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

Out[35]:

	City	District	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 Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	South Jakarta	Jagakarsa	1	Diner	Dog Run	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store	Eastern European Restaurant	Donut Shop	Women's Store

Cluster #3:

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

Out[36]:

	City	District	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 Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	West Jakarta	Kalideres	2	Airport Food Court	Women's Store	Donut Shop	Fish & Chips Shop	Fast Food Restaurant	Farmers Market	Factory	Event Space	Electronics Store	Eastern European Restaurant

Discussion

From the results and the data in the exploratory, we can expect that f&b provider market has been very concentrated in Jakarta. The clustering model has also suggests in the cluster #1, f&b business has been mushrooming evenly across Jakarta, as it leaves the other remaining clusters only one district each. Cluster #2 and #3 suggests that business in Jagakarsa and Kalideres has a bit more variety than in other area of Jakarta.

Conclusion

From the Foursquare API data, it can be assumed that:

1. Trying to start an f&b business in Jakarta can be a bit tough, due to the competitive industry environment as there are a lot of suppliers.
2. Other kind of business may have a higher chance of success since there are less competition.
3. Jagakarsa and Kalideres are the area that have more variety of business type in Jakarta.

This project can give an overview of business venue distribution in Jakarta and to illustrate a little on what kind of business that may be suitable to start in Jakarta. However, we need to remember that this still doesn't take into account any other important consideration such as avg income of people living in those area, estimaton of demand for related industry, and so on. Thank you.