

Coursera Capstone  
Simulated Business Exercise

Soumya Ranjan Prusty  
April 2020



Little Elm  
TEXAS

---

## ***Business Opportunity and Audience***

---

Hypothetically, I am a commercial account executive at a local bank and I am considering opportunity in my area.

Little Elm, Texas is an up and coming suburb of North Dallas, and is currently developing at an intense rate. There is construction at every corner. This is a prime time to begin entrepreneurship. Little Elm is also landlocked by other cities, so this is the ideal opportunity to act in the business sector as either an investor or business operator, before the point of entry is saturated and opportunity becomes scarce.

I am interested in exploring the type of businesses, in Little Elm, that my company should invest its capital in. This will influence which types of business loans I approve, and what sort of marketing my company will undertake in order to attract clients interested in opening businesses in my area. I could also use this information, to potentially, advise my existing, valued investment clients in the types of businesses to invest their money in. This would allow me to capitalize on an opportunity to improve relationships with my current clients, increasing their loyalty.

I will explore 5-10 options and compile my target list.

---

## ***Data Description, Source, and Extraction Steps***

---

Using Foursquare Data, from foursquare.com, I will explore the business landscape in Little Elm and at least one surrounding city, which is more historically developed. I will compare the volume of businesses in that city to what exists currently in Little Elm and see what will be ideal categories for growth, assuming for the purposes of this project, that Little Elm could support a similar business profile as the comparison city (regardless of actual comparison).

---

## Methodology

---

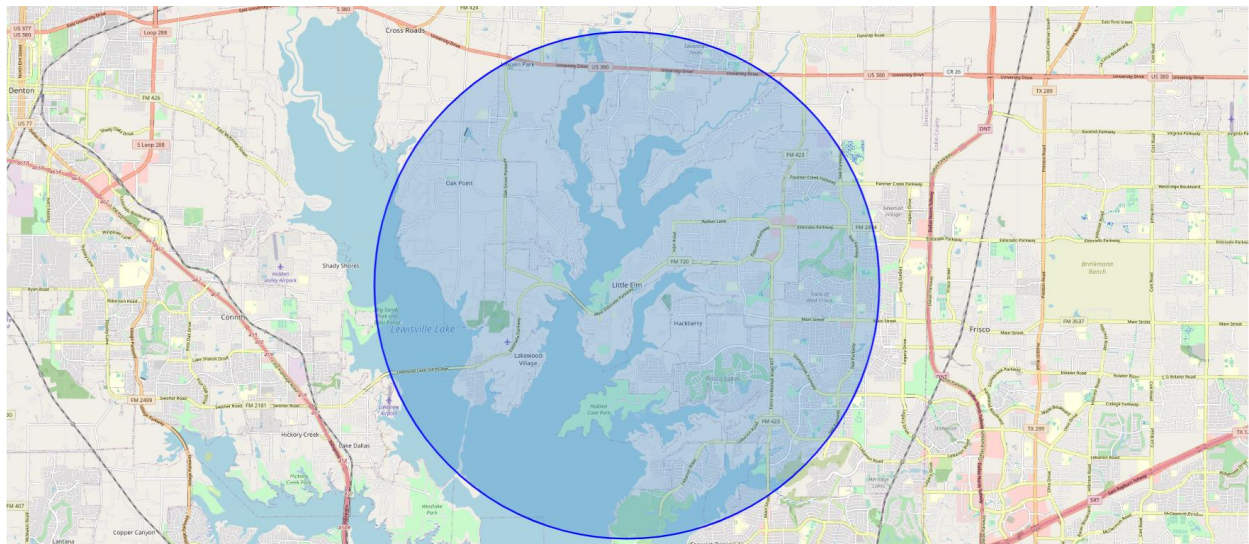
After importing vital libraries that allow me to render maps, read urls, and create dataframes, such as pandas, json, and folium, I was able to move on to creating code to solve my problem.

I began my research by surveying the current business segments that exist in Little Elm. Foresquare.com allows you to use the location of an address, and query user compiled data. I was able to generically set “Little Elm, Tx” as my address and I extracted the latitude and longitude from the geopy.geocoders library. These variables are used multiple times, both in extracting the data from foresquare, and also map rendering. I created a dataframe to house this data.

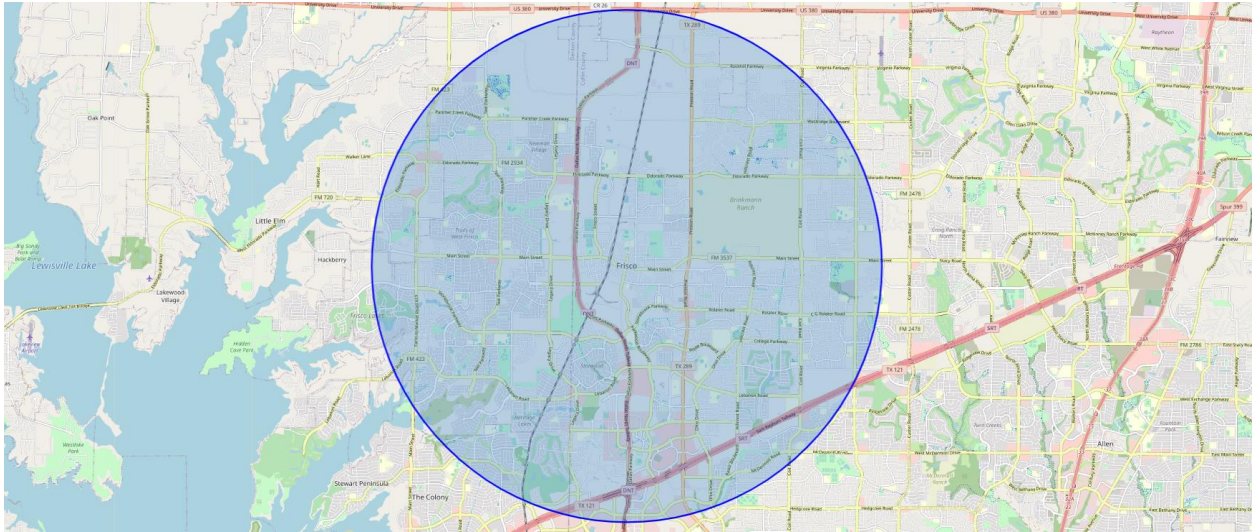
	City	Latitude	Longitude
0	Little Elm	33.162619	-96.937505

The next tool I used was rendering a map of my target location. I used the folium map rendering tools and created a map. I also enhanced the default zoom level so the full city could easily be seen. I increased the size of my CircleMarker function to call much detail to the entire city. This was useful to ensure my location matched my intended targets.

Little Elm Map



## Frisco Map



Next, I created variables for my foresquare credentials so I could create a URL to pass to foresquare to extract my desired data. These credentials were created in advance.

Important considerations were made for the limit of venues I wanted to consider and the radius in which I wanted to search. I chose a limit of 200 venues to get a large variety of businesses. I also chose a radius of 8000 meters, which is approaching 5 miles. After the data was returned using the requests and json fuction/library I saved the data to a dataframe, including only the elements I was interested in evaluating. The transformation of the data is depicted by the following figures.

```
Out[159]: {'meta': {'code': 200, 'requestId': '5e98a110fb34b5408dd9d0e1'},
  'response': {'suggestedFilters': {'header': 'Tap to show:',
    'filters': [{'name': '$-$$$$', 'key': 'price'},
      {'name': 'Open now', 'key': 'openNow'}]},
    'headerLocation': 'Little Elm',
    'headerFullLocation': 'Little Elm',
    'headerLocationGranularity': 'city',
    'totalResults': 226,
    'suggestedBounds': {'ne': {'lat': 33.29761953500013,
      'lng': -96.77653910563417},
      'sw': {'lat': 33.02761926499986, 'lng': -97.09847109436582}}},
```

Out[14]:

	name	categories	lat	lng
0	Little Elm Crafthouse	Bar	33.155467	-96.944999
1	Beard Park	Park	33.163368	-96.932129
2	Little Elm Park	Park	33.159002	-96.947125
3	Beach	Beach	33.159321	-96.947473
4	Little Elm Rec Center	Athletics & Sports	33.165091	-96.939876

```
In [15]: print('{} venues were returned by Foursquare.'.format(nearby_venues.shape[0]))
```

100 venues were returned by Foursquare.

100 venues were returned by Foresquare.

My next task was to select a city to use for comparison. I chose Frisco, TX. Frisco is a well-known suburb of North Dallas. Growth projections have been staggering in the past 2 decades. In 2017, Frisco was the fastest-growing city in the nation, and it also held that title from 2000 to 2009 according the Wikipedia. In addition, Frisco is projected to double in size in the next 10 years. It is a prime candidate of comparison to use for an aspiring city.

With this in mind, I repeated each step of obtaining data for Frisco that I did for Little Elm. I found the latitude and longitude, and then I made the call to Foursquare to obtain venue data.

:

	name	categories	lat	lng
0	Babe's Chicken Dinner House	Southern / Soul Food Restaurant	33.149031	-96.830446
1	Mariana's Taco Shop	Mexican Restaurant	33.151367	-96.824001
2	Frisco Public Library	Library	33.150182	-96.834209
3	Frisco Commons	Park	33.153426	-96.813934
4	ILoveKickboxing.com - Frisco, TX	Gym / Fitness Center	33.149256	-96.805984

```
print('{} venues were returned by Foursquare.'.format(Frisco_nearby_venues.shape[0]))
```

100 venues were returned by Foursquare.

100 venues were alos returned for Frisco.

My final data configurations was to take the data obtained for each city and aggregate each venue by the category of business. I used a group by, count, and order by function to obtain a total for each category in the data, and then utilized the order to give me the most and least popular.

I used the head function to obtain the top and bottom 10, simply switching the order by from the default ascending to descending.

Now, it was the big payoff. I reviewed the data for Frisco, looking at the top categories, such as Mexican Restaurants, Coffee Shops and Pizza Places. To my surprise, Little Elm is doing a spectacular job with their development plan! There were only 3 categories in Frisco's top ten that were not in Little Elm's top ten!

To expand my results, I revisited my data collection, expanding my radius for each city from 8000 meters to 15000 meters, which is closer to 10 miles. I hoped to pick up additional variation in business types. I

also expanded my view of the top categories from 10 to 15. After these adjustments I was satisfied with the quality of data to evaluate.

```
]#sorted by the count in decending order to get top 15
LE_grouped = nearby_venues.groupby('categories').count().reset_index()
LE_grouped.sort_values(by=['name'],inplace=True, ascending=False)
LE_grouped.head(15)
```

[29]:

	categories	name	lat	lng
34	Mexican Restaurant	8	8	8
14	Fast Food Restaurant	7	7	7
39	Park	6	6	6
24	Grocery Store	4	4	4
40	Pizza Place	4	4	4
12	Coffee Shop	4	4	4
33	Liquor Store	3	3	3
29	Ice Cream Shop	3	3	3
26	Gym / Fitness Center	3	3	3
21	Golf Course	3	3	3
30	Italian Restaurant	3	3	3
45	Supermarket	3	3	3
46	Sushi Restaurant	3	3	3
2	BBQ Joint	3	3	3
27	Hardware Store	2	2	2

```
LE_grouped = nearby_venues.groupby('categories').count().reset_index()
LE_grouped.sort_values(by=['name'],inplace=True)
LE_grouped.head(10)
```

[0]:

	categories	name	lat	lng
49	Wine Bar	1	1	1
32	Library	1	1	1
35	Motorcycle Shop	1	1	1
48	Warehouse Store	1	1	1
23	Greek Restaurant	1	1	1
22	Golf Driving Range	1	1	1
36	Movie Theater	1	1	1
20	Gas Station	1	1	1
19	Garden Center	1	1	1
18	Furniture / Home Store	1	1	1

```
#sorted by the count in decending order to get top 15
Frisco_grouped = Frisco_nearby_venues.groupby('categories').count().reset_index()
Frisco_grouped.sort_values(by=['name'],inplace=True,ascending=False)
Frisco_grouped.head(15)
```

[5]:

	categories	name	lat	lng
38	Mexican Restaurant	7	7	7
13	Coffee Shop	6	6	6
29	Ice Cream Shop	5	5	5
30	Italian Restaurant	5	5	5
56	Sushi Restaurant	4	4	4
20	Furniture / Home Store	3	3	3
17	Fast Food Restaurant	3	3	3
27	Gym / Fitness Center	3	3	3
45	Pizza Place	3	3	3
55	Supermarket	2	2	2
51	Southern / Soul Food Restaurant	2	2	2
26	Gym	2	2	2
44	Park	2	2	2
24	Grocery Store	2	2	2
2	BBQ Joint	2	2	2

```
Frisco_grouped = Frisco_nearby_venues.groupby('categories').count().reset_index()
Frisco_grouped.sort_values(by=['name'],inplace=True)
Frisco_grouped.head(10)
```

[1]:

	categories	name	lat	lng
0	American Restaurant	1	1	1
32	Juice Bar	1	1	1
33	Korean Restaurant	1	1	1
34	Latin American Restaurant	1	1	1
35	Library	1	1	1
36	Lingerie Store	1	1	1
37	Liquor Store	1	1	1
39	Museum	1	1	1
40	Nail Salon	1	1	1
41	New American Restaurant	1	1	1

## Project Findings and Recommendations

I began by listing each category that was present in the original top 10 Frisco dataframe but not the Little Elm dataframe as my Green Light List. Those included Fried Chicken and Sushi Restaurants and Ice Cream Shops.

Once I expanded my data set, I also added Furniture/ Home Store to my list, especially since it is on the list of the lowest volumes of business that exist in Little Elm currently.

I also felt items like Coffee shops that are popular in Frisco might be good candidates, however Little Elm has almost as many as Frisco, so I would approach this type of business with more caution, and make sure the borrower had a solid business plan, and perhaps had more collateral in our deal.

Finally, I concluded that Little Elm has an abundance of Fast Food Restaurants, Grocery Store/Supermarkets (I visably combined these numbers) and these are businesses in which would not be desirable to consider for business investment. Additional businesses I would hesitate to make contracts with would be the on the bottom 10 on the Frisco list.

---

## *Project Summary*

---

With Python and public data sources, I was able to evaluate a real world business challenge and come up with promising insights. I feel if commercial banking were actually my profession, the data I gleaned from this evaluation would leave me with actionable data in which to make sound business decisions. I might duplicate my efforts with other neighboring cities and compile each list to have a bigger variety. I might also expand the geographical area of Little Elm, or evaluate where each existing business is, and compare these clusters of Little Elm with the Frisco data to see if perhaps certain neighborhoods in Little Elm were ideal for specific businesses. My final list is presented below

Little Elm Targe Business Investement List		
Green	Yellow	Red
Fried Chicken	Coffee	
Shop	Shop	Fast Food
Sushi		Grocery
IceCream		Juice Bar
		Lingerie
		Store
		Museum