### **Business Problem**

An internet service company based out of Phoenix is looking to expand its market share in the City. The company was providing services to 2 kinds of customers - individual customers and families; and small and medium size businesses. It was found out that the profit earned from small-size business such as restaurants, Fast Food chains was quite high; however, the company's market share of such customers was quite low (about 5%). From market research, the root cause for such a small market share was due to the customers not aware of the company and its services. Now, the internet service company aims to identify and target all the fast food and restaurant businesses using Foursquare API.

### **Background:**

The given internet service provider is the 5th largest cable service provider and amongst the top 20 internet service providers. In fiscal year 2017, it reported 0.7% customer growth, and this was largely attributed to reduction in broadband promotions. The company slowly started regionally customizable marketing campaigns while moving away from 'one size fits all' marketing philosophy. At the end of 2017, residential high-speed internet (HSI) broadband users grew at 24.7% while video customers decreased by 11.9%. One year later, over 60% of the company's revenue was from residential broadband and business services. By 2019, a whopping 70% of the company's subscribers preferred the buy only it's broadband services rather than bundling it with video. With a slew of new marketing strategies in place and launched new product plans, the company looks to further reduce its customer churn through a data-driven model; and target small-sized businesses which show higher retention and lower customer churn.

### Data

1. Phoenix zip codes were simply taken out from google search as in below link:

https://www.google.com/searchq=phoenix+zip+codes&rlz=1C1CHZL\_enIN746IN746&oq=phoenix+zip+codes&aqs=chrome..69i57j0l7.3844j1j4&sourceid=chrome&ie=UTF-8

2. The zip code data exported from the above link is also uploaded in the link below:

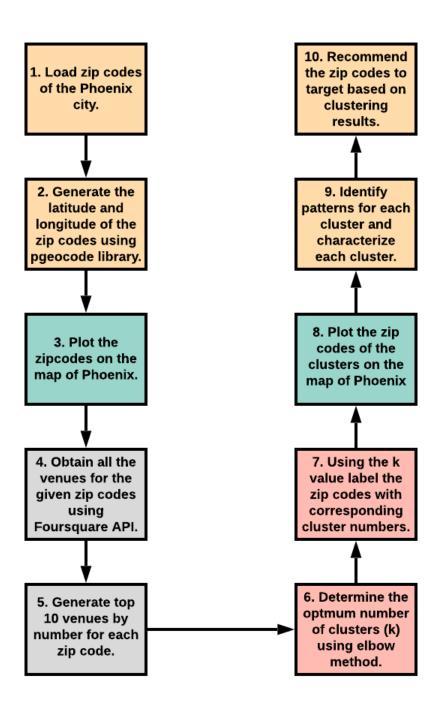
https://github.com/mbsuraj/Clustering Neighborhoods Using FoursquareAPI/blob/master/Phoenix Zip.csv

3. The code to cluster the city and determine the zip codes for marketing campaign is in the link below:

https://github.com/mbsuraj/Clustering Neighborhoods Using FoursquareAPI/blob/master/Neighborhood of Phoenix.ipynb

## Methodology

The approach to the case is illustrated in the flowchart below:



The zip codes in Phoenix city are obtained using google search. And the corresponding latitudes and longitudes are generated using "pgeocode" library in Python as below:

	Zip Code	Latitude	Longitude
0	85001	33.45	-112.07
	85003	33.45	-112.08
2	85004	33.46	-112.07
3	85006	33.47	-112.05

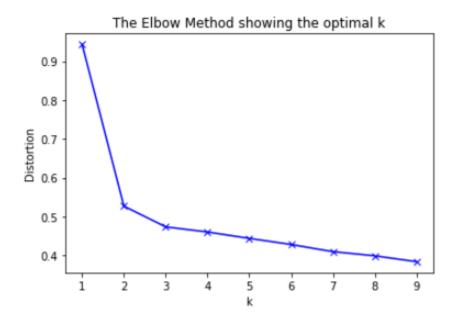
The zip-codes are then visualized on a map using folium package as below:



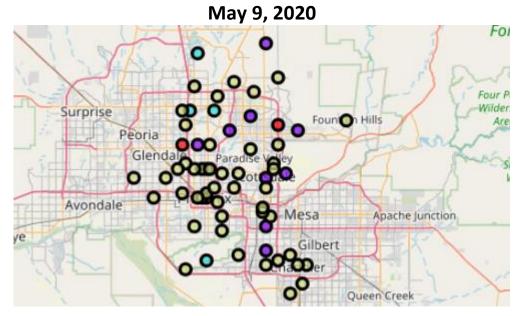
Many zip codes appear to be overlapping due to similar longitude and latitude values and hence all such duplicate zip codes are dropped; the analysis is then continued using around 69 zip codes. Once all the required data is ready, I called Foursquare API to obtain more than 200 different venues in the 500m radius of the given set of latitudes and longitudes. The purpose of obtaining venues for each zip code was to characterize each zip code based on the businesses that exist there. Since there are several venues in each zip code, the better way was to generate the top ten venues for each zip code as below:

	ZipCode	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	85001	Coffee Shop	American Restaurant	Pizza Place	Hotel	Lounge	Mexican Restaurant	Salon / Barbershop	Basketball Stadium	Sandwich Place	Sushi Restaurant
1	85003	Music Venue	Coffee Shop	Southern / Soul Food Restaurant	Brewery	Building	Rock Club	Sushi Restaurant	Theater	Mattress Store	History Museum
2	85004	Art Gallery	Coffee Shop	American Restaurant	Dessert Shop	Bar	Cocktail Bar	Sandwich Place	Taco Place	Performing Arts Venue	Bike Shop
3	85006	Coffee Shop	Taco Place	Café	Clothing Store	Art Gallery	Liquor Store	Bakery	Bar	Restaurant	Seafood Restaurant

Once the top 10 venues for each zip code was clear, I clustered all the zip codes. But before clustering I had to decide the number of clusters (K) to obtain. Hence, I used elbow method to decide K as below:



As seen in the chart above the chart seems to elbow at k = 4. So, I used k = 4 for clustering the zip codes; KMeans algorithm was used here. Then all the clusters are plotted in the map as below:



As it can be seen, the map doesn't seem to help much in terms of preparing any strategy. However, when I looked at each cluster separately, it was easy to characterize as below:

## Cluster 0:

	ZipCode	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	Cluster Labels
0	85260	Farm	Intersection	Rest Area	Yoga Studio	Dive Bar	Fishing Spot	Fast Food Restaurant	Farmers Market	Falafel Restaurant	Fabric Shop	0
1	85051	Home Service	Concert Hall	Seafood Restaurant	Moving Target	Doctor's Office	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fabric Shop	0

## Cluster 1:

	ZipCode	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	Cluster Labels
2	85021	Park	Playground	American Restaurant	Dive Bar	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fabric Shop	1
3	85284	Soccer Field	Spa	Park	Doctor's Office	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm	Falafel Restaurant	Fabric Shop	1

### Cluster 2:

	ZipCode	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	Cluster Labels
11	85085	Scenic Lookout	Trail	Mountain	Auto Workshop	Yoga Studio	Donut Shop	Fishing Spot	Fast Food Restaurant	Farmers Market	Farm	2
12	85023	Fast Food Restaurant	Trail	Yoga Studio	Doctor's Office	Flower Shop	Fishing Spot	Farmers Market	Farm	Falafel Restaurant	Fabric Shop	2

### Cluster 3:

	ZipCode	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	Cluster Labels
15	85255	Pharmacy	Steakhouse	Pizza Place	Indian Restaurant	Salon / Barbershop	Shipping Store	Bank	Miscellaneous Shop	Supermarket	Gas Station	3
16	85053	Hookah Bar	Greek Restaurant	Park	Cosmetics Shop	Deli / Bodega	Disc Golf	Discount Store	Mexican Restaurant	Pizza Place	Bakery	3

### **Results**

Logically each cluster could be characterized as below:

Cluster 0: Places to go for peaceful activities like Yoga, Rest areas etc.

Cluster 1: Places for any outdoor activity like soccer field, park, fishing spot etc.

Cluster 2: Fast food and restaurants

Cluster 3: Wide array of utilities and healthcare venues are present here.

### **Discussion and Conclusion**

As noted above, all the zip codes in cluster 2 seem to be of fast food and restaurant type businesses. The internet provider now has the zip codes it can target for its initial marketing campaign and be assured of greater market penetration. However, I also believe the marketing campaign would not be as much straightforward. We would be learning several things during the campaign itself – such as customer feedback, customer sentiment and the company's brand perception.