# Problem Description

When new or existing restaurant owners are trying to decide on a new location for their restaurant, there are multiple key factors for them to consider and understand before they decide on a new location.

Based on my research of restaurant location services and restaurant owner magazines, I have narrowed my list down to several key factors owners need to understand before choosing a location.

These factors include visibility, parking, space size, crime rates, surrounding businesses and competitor analysis, accessibility, and safety.

For the purpose of this Capstone project, I have narrowed my choice to one key factor, “Understanding Surrounding Businesses and Competitor analysis for New Restaurant Locations” as the key problem for my project to address.

Background Discussion

Lisa Melbourne is s hypothetical restaurant owner of “Breaking Bread”, a highly successful breakfast-only restaurant operating in Southern California. Lisa is looking to expand her restaurant operations into Northern California, with key interest to find a location within the San Francisco South Bay region. Lisa has enlisted my services and needs my recommendation to identify a city and location with increasing/high cross-business traffic and area having fewest breakfast venues within a couple of miles so she is not launching her restaurant into a competitive atmosphere.

My goal is to provide the research data Lisa requires so she can effect an informative decision for her new restaurant location.

Data Description:

1. San Jose postal zipcode data will be obtained and used via pip install zipcodes
2. Foursquare data API will be used to meet the following requirements:
   1. San Jose regional data showing all business venues will be required and segmented as follows:
      1. All restaurants in San Jose serving breakfast
      2. Surrounding businesses of each breakfast restaurant along with their customer traffic data to assess how well businesses are doing within San Jose region.
      3. Data by zip/suburb to identify specific sub-regions within San Jose our client is interested in.
   2. Business customer traffic data so client can see which business venues and locations their customers are coming from.
3. Competitive Analysis
   1. At a San Jose level, provide Data to show unique grouping of breakfast venues owned by same customer. Data to describe the unique number and percentage of breakfast restaurants operating within San Jose and organized to show level of competition and saturation levels.
   2. Location identification:
      1. For customer to further refine their area of interest, geolocation map is provided for client so client can select and refine their venue selection within 2 mile radius. Popup data is shared showing number of breakfast restaurants in each selected area.
      2. Saturation level of other restaurant venues serving breakfast and their proximity spacing to customer can be seen on map plot area of interest for customer to avoid conflict.
4. I will use Foursquare API for venue analysis, geolocation, and to assess business traffic, as well as to understand how well breakfast venues are doing in the San Jose region as a whole.

1. All San Jose Zipcodes will be converted to sj\_zipcode dataframe
2. Zip data will be used to segment neighborhoods as performed in prior lab.
3. All neighborhoods/zips will have designated latitude and longitudes using geocode information from http://cocl.us/Geospatial\_data
4. Neighborhood data will be merged with geocode data
5. We will use Folium to map San Jose using latitude and longitude values, adding markers above to map and start exploring visuals.
6. Foursquare data formats we will use include: v2/vedirnues/search, v2/venues/explore, v2/venues/nextvenues
7. getNearby Venues function will be used
8. new function and section will be created for specific development of dataframe to show source venue traffic data coming into our restaurant venue (next\_venue)
9. Most common venues dataframe will be utilized for San Jose
10. A new dataframe will be created to include the cluster as well as the top 10 venues for each neighborhood.