

## Finding New Business Clients for Wi-Fi Services

### Data

As mentioned in the Introduction, the aim of this project is to gather data about businesses to determine which **types** of businesses will benefit the most from providing Wi-Fi to their customers as well as to determine which **specific** businesses should be targeted. For this project, I am going to focus on three types of businesses: coffee shops, sandwich places and donut shops. In addition, to keep the scope of this project manageable, I am going to focus on businesses in the greater Los Angeles area (though the conclusions drawn should be applicable to other major metropolitan regions).

To obtain this data, I will rely upon the Foursquare database; specifically, I will use a Python program to make calls to Foursquare using their various application programming interfaces (APIs); each call that I make to Foursquare will contain the specifics of the information that I need (such as venue type and location); once I obtain the results from each call, the Python program will parse the results and create records for each business that can be used for my analysis.

The first step will be to obtain lists of 1) coffee shops, 2) sandwich places and 3) donut shops in the Los Angeles area. For each type of venue, a call will be made to Foursquare (using their API) to obtain a list of venues specified by type and location. Once the results of each call are received, the list of venues returned in the query will be appended to a master list of venues stored in my Python program. (Based on my experience using the Foursquare API, I should be able to obtain about 150 examples of each type of venue, which should provide a sufficient number of records for this project.)

The records received from Foursquare (in this type of query) contain many fields, but I will only store those values which are relevant for my analysis, including:

- The ID number assigned to the venue by Foursquare
- The name of the venue
- The street address of the venue
- The city where the venue is located
- The state where the venue is located
- The ZIP Code where the state is located
- The latitude and longitude of the venue
- The phone number of the venue
- The type of the venue (i.e. "Coffee Shop", "Sandwich Place", "Donut Shop")

For my analysis, I'll also need the current rating of the venue (a value on a scale of 0.0 to 10.0) and a field which indicates whether or not the venue has Wi-Fi. However, these fields are not present on records received in the type of query described above, so in order to obtain that information, I will make an additional call (using a different Foursquare API) for each individual venue in order to fetch detailed information for that venue. When the results of each individual call are received, I will extract the rating and "has wi-fi" values and will update each record in my dataset accordingly.

Once I have obtained all of the information described above, I will have a sufficient amount of data to perform my analysis.