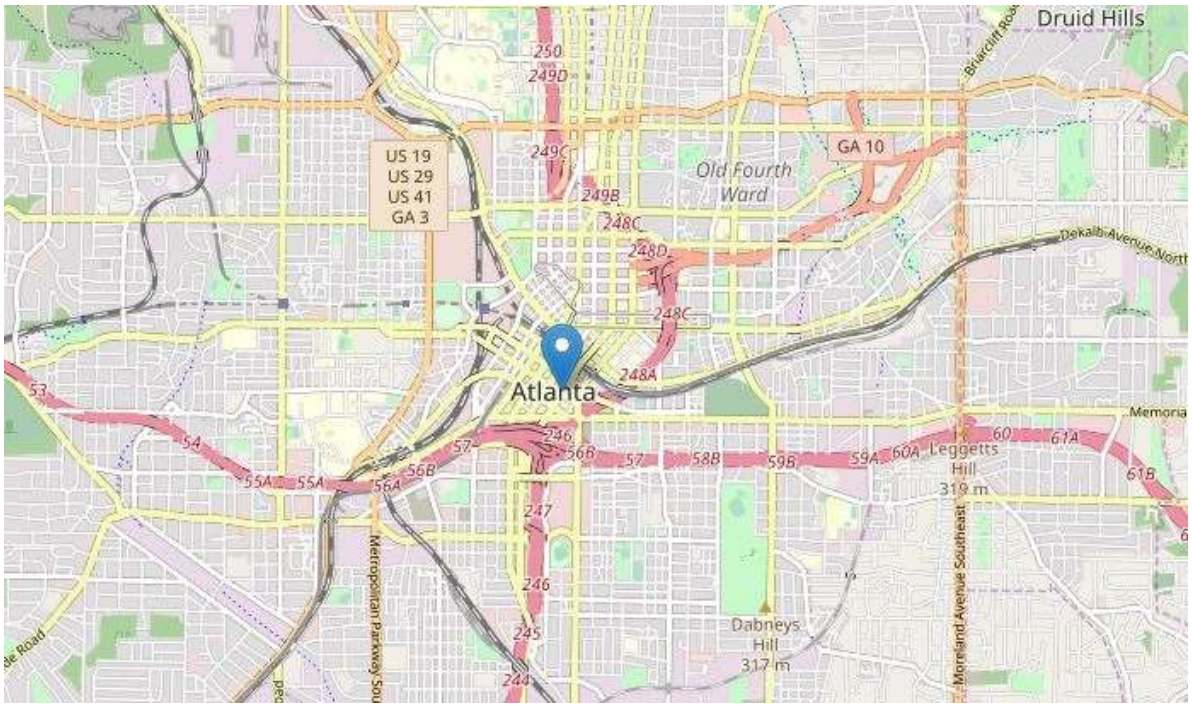
CLUSTER ANALYSIS OF RESTAURANTS TO VISIT, ATLANTA, GA

2. Data Acquisition and Data cleaning

Let's use the folium library to create a complete map zoomed on Atlanta. We'll also plot a marker on the coordinates we just identified above. This would give us a relatively good look at the center point we will be considering. The LATITUDE and LONGITUDE of ATLANTA are 33.7490, -84.3894.



2.1 Data Acquisition

Foursquare API:

We will use Foursquare API to fetch venue details around longitude and latitude of Atlanta. We'll call the

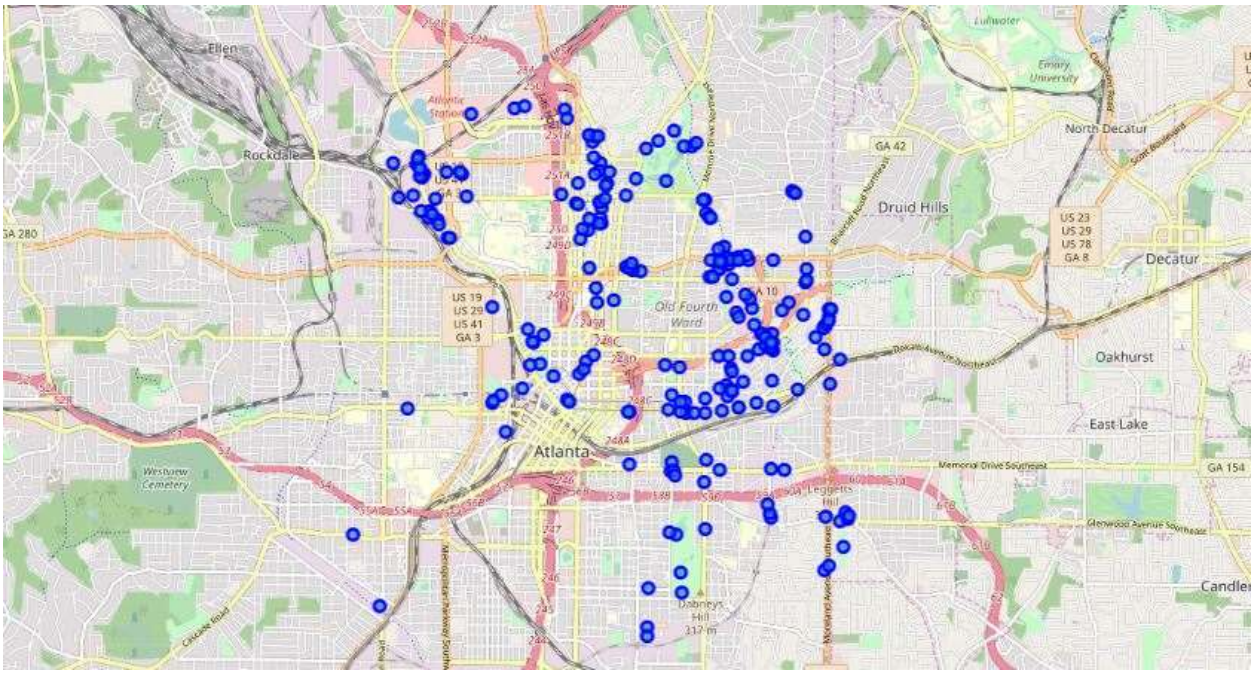
API over and over till we get all venues from the API within the given distance. The maximum venues this API can fetch is 100, so we will fetch all venues by iteratively calling this API and increasing the offset each time.
Foursquare API requires ClientID, and client secret to function which can be accessed after creating a developer account. We will set the radius as 5 Kilometers. The version is a required parameter which defines the date on which we are browsing so that it retrieves the latest data.

Zomato API:

The Zomato API allows using its search API to search for any given venue based on certain search filters such as query, latitude, longitude and more. Zomato also requires a Zomato user key which can be accessed with a developer account. We'll use the name, lat, and lng values of various venues fetched from Foursquare API to use the search API and get more information regarding each venue. The query will be the name of the venue. The start defines from what offset we want to start, so we'll keep
it at 0. The count defines the number of restaurants we want to fetch. As we have the exact location coordinates, we'll fetch only one. We will supply the latitude and longitude values. We will set the sorting criteria as real_distance so each time we get the venue we're searching based on location coordinates.
The data from multiple resources might not always align. Thus, it is important to combine the data retrieved from multiple resources properly.

We'll first plot the two data points on the map. We'll then try to combine data points that have their latitude and longitude values very close to one another. From the remaining selected venues, we will inspect the venues to ensure that any remaining mismatched venues are also removed from the final dataset of venues before we begin any analysis.
We will first plot the Foursquare data on the map and then Zomato data on map.

Foursquare Venues on map:



Zomato Venues on map: