

This project sources and integrates data from the Wikipedia:

[https://en.wikipedia.org/wiki/Category:Suburbs\\_of\\_Algers](https://en.wikipedia.org/wiki/Category:Suburbs_of_Algers). Another data source is Foursquare API. This section describes each of these data sources and provides examples of the data.

The Wikipedia page provided suburbs in Algiers as well as populated areas in the city. The data included provinces in an alphabetical order and grouped as such.

As the data was obtained from a Wikipedia page, it was scraped and the Beautiful soup package was used to parse the data. Then the data was appended into a list and a pandas dataframe created.

Foursquare API, a social networking service which provides a mobile app that allows each user to search for venues close by and see information and reviews from other users was used to get location data. Users feed information to Foursquare passively as the app tracks users' locations. Users also do so actively whenever they enter venue names, locations, and reviews. This project will access Foursquare venue data for the selected Suburbs in Algiers to obtain neighbourhoods, specifically popular venues. The Foursquare venue data will particularly seek to identify popular areas that have unique categories. These data will then be used for subsequent categorization of neighborhoods to provide insight to places where shopping malls are present or totally absent.