# **2. Data**

The data is extracted from Wikipedia. For this study, we will use the list of the 60 busiest airports per boarding passenger on page [*List\_of\_the\_busiest\_airports\_in\_the\_United\_States*](https://en.wikipedia.org/wiki/List_of_the_busiest_airports_in_the_United_States#See_also), see Figure 1, that features, for instance, the names of the airports, IATA codes, cities, and passenger traffic from year 2014 to 2020. We will only use the most recent of data, that is, from 2020. Each airport name is linked to another Wikipedia page that fully describes the airport, from which we can get the geographic location, see example in Figure 2. In order to determine which amenities are available, we call foursquare with a radius of 10km from each airport location.



Figure 1: First ten lines of the airport data table on the wikipedia page



Figure 2: Example of latitude/longitude information at a given airport from its wikipedia page