

## **DATA DESCRIPTION**

### **a) Use of the Foursquare API**

The Foursquare API will be used to fetch the data about the features (like restaurants, night clubs and parks) that exist in the main neighborhoods of Sao Paulo, in a similar manner to the work done in the previous weeks of the Casptone.

### **b) Other Sources of Data**

Since the geographical data about São Paulo is not well structured as the data about some cities in the USA and Canada, there will be a previous research in the public records about the city to determine the central coordinates of each neighborhood. This kind of data is usually non-structured, so this first step will be more “old school” than the futher clustering analysis.

Later, when the clustering analyis is done, some geographical and social non-structured data will be consulted. These data will be used to provide a brief description of the regions of each cluster, enriching the final product that is delivered to the stakeholders. This process will help to provide a more in-depth contextualization about what the tourists and newcomers can expect in each cluster.