

The following data sets will be used for this report:

1. A dataset containing the districts and neighborhoods of Madrid will be used. This dataset will include the geographic coordinates of each district for subsequent location on maps and clusters. Madrid has 21 districts that, in addition, group a total of 131 administrative neighborhoods. This information will be obtained from Wikipedia (https://en.wikipedia.org/wiki/Districts_of_Madrid) and from official organizations such as (<http://portalestadistico.com/munbiernoencifras/?pn=madrid&pc=ZTV21&idp=34&idpl=1328&idoma=>)
2. We will use Foursquare.com to obtain leisure venues that may be of interest to tourists, from each district and existing hotels. Crossing this information with the previous one of the districts we will be able to group the information obtained in clusters and we will be able to represent them on maps for a better understanding. In this way we will be able to determine the best location for the new hotel.