

1. Data acquisition and cleaning

1.1.Data sources

We will use different types of data in our research. First of all, we will need to determine the coordinates of the city of St. Petersburg, as well as the coordinates of the city districts. We need the coordinates so that using the application programming interface (API) we can see what is around a given point and determine the number of Italian restaurants. We will also need the population size in the city districts, as well as their area to determine the population density. The data sources for our research will be the site, where the coordinates of the city will be indicated. Wikipedia, from there we can take data on the population in each district, as well as the coordinates of the districts and the area of each district. And of course, data from Foursquare, from where we will take the data we need on the location of Italian restaurants.