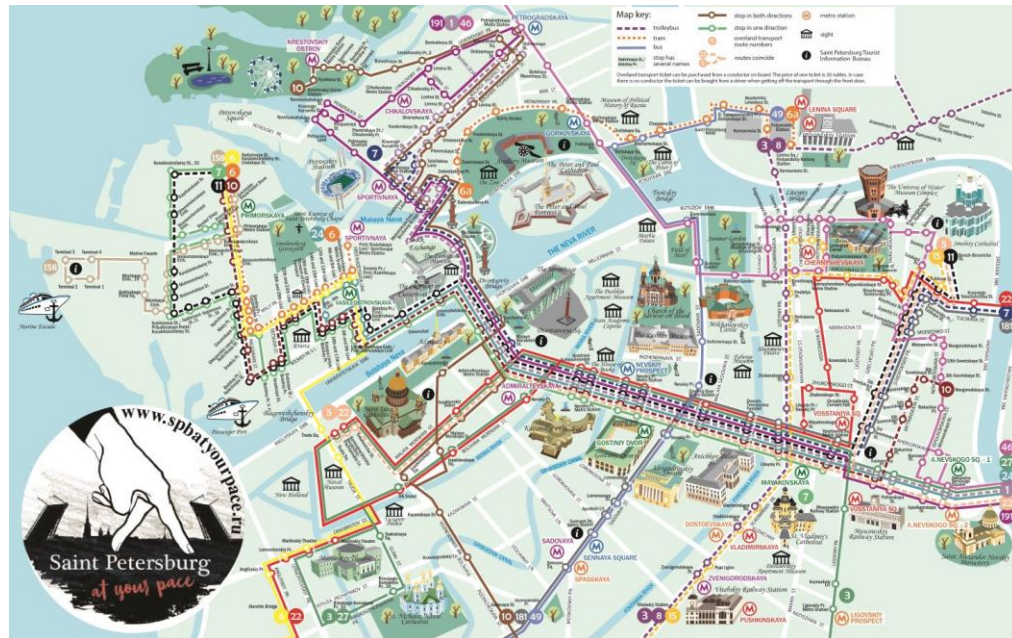# Choosing the district, where to open an Italian Restaurant in Saint-Petersburg, Russia
Aleksandr Kan
July 2021

1. Introduction

1.1. Background

Starting your own business is always a challenge and big risks, something can always go wrong. You may not have enough money, your product may not find a response from buyers, you may be wrong with the chosen marketing strategy, or choose the wrong place to open your business. Among other things, opening your own restaurant is one of the most difficult beginnings in my opinion. Choosing the right area to open is a difficult choice and a lot depends on it. Will there be a sufficient number of people around, how strong is the competition and if there is none at all, then perhaps for a reason.

1.2.Problem

In this study, we will consider one of the factors for the successful opening of a restaurant - the population density in a particular area of St. Petersburg and the number of Italian restaurants open around. We will be based on the fact that if there are no Italian restaurants in a certain radius in the area, then this area is not suitable for us, since most likely other restaurateurs conducted research and the absence of an Italian restaurant in it for a reason.

1.3.Interest

This research will be of interest to those who want to open their own restaurant and choose a suitable area in St. Petersburg. Also, the study will be interesting to anyone who wants to open a business, to see the competition around, since it will be enough just to change a few parameters. For example, by entering the coordinates of an existing grocery store, changing the search radius and the name of the place found in Foursquare, you will see the number of competitors around.

2. Data acquisition and cleaning

2.1.Data description

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.

2.2.Data sources

- Foursquare API
- Coordinates of Saint-Petersburg and all districts (Wikipedia)
- Square area and population of districts (Wikipedia)
- Python libraries:
    1. Pandas: to create and manipulate Data Frames
    2. JSON: to handle with JSON files
    3. Folium: to show the map of the city
    4. Geopy: to get the coordinates
    5. BeautifulSoup: to parse necessary data from websites