

Coursera Capstone

IBM Applied Data Science Capstone

Opening a New Fitness Club in Moscow, Russia

By: Alexander Konnov
April, 2021

Business Problem

It is necessary to assess the prospects of certain districts of Moscow in terms of placing new fitness clubs there. Using Data Science and clustering methods, it is necessary to prepare answers to the question of where is the most profitable to open a new fitness club.

The report should provide answers to the following questions:

- ▶ The most efficient location for a fitness club aimed at office workers
- ▶ The most efficient location for a family oriented fitness club.

Data

Data expected to be used in preparing this project:

- ▶ Distribution of population density in Moscow, Russia
- ▶ Distribution of offices and colleges in Moscow, Russia
- ▶ Distribution of gyms in Moscow, Russia

Sources of Data:

Open Street Map for boundaries of Moscow

Geocoder package for latitude and longitude coordinates

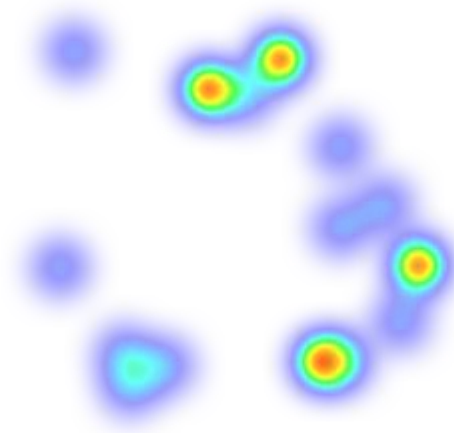
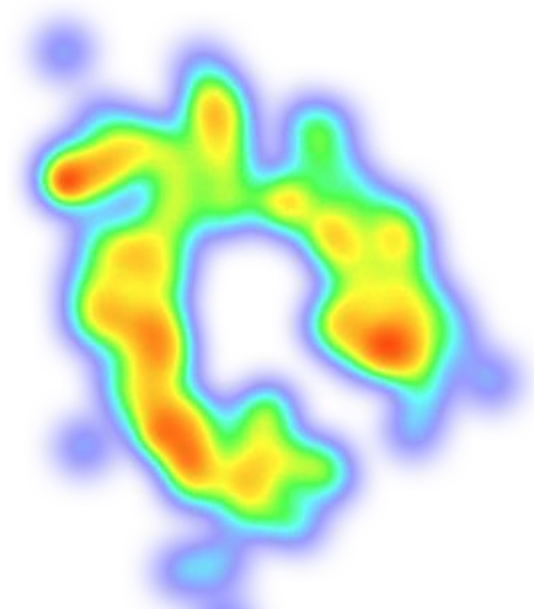
Foursquare API for venue data

Methodology

- ▶ We are going to prepare three sets of data:
- ▶ The first one based on distribution of groceries. It should give me the distribution of population in Moscow
- ▶ The second one based on distribution of colleges and offices. It should give me the distribution of where people spend their working days
- ▶ The third is based on distribution of fitness clubs. It gives the distribution of fitness clubs
- ▶ Then we'll make the DBSCAN clusterisation for each dataset. The goal of clusterisation is to divide each set in two clusters. The first with high density and the second with low density (outliers). To make it we are going to join all clusters gathered by DBSCAN into the first.

Data Analysis

Areas with low density population and low fitness clubs distribution



Results

Here are addresses of places for fitness clubs:

- ▶ - Школа Свиблово, корпус 6, улица Седова, Свиблово, район Свиблово, Москва, Центральный федеральный округ, 129323, Россия
- ▶ - 46 к2, Новочеркасский бульвар, район Марьино, Москва, Центральный федеральный округ, 109369, Россия
- ▶ - 9, Тайнинская улица, Лосиноостровский, Лосиноостровский район, Москва, Центральный федеральный округ, 129345, Россия
- ▶ - 16 к2, Тайнинская улица, Лосиноостровский, Лосиноостровский район, Москва, Центральный федеральный округ, 129345, Россия
- ▶ - 2044, Люблинская улица (дублёр), Марьинский парк, район Марьино, Москва, Центральный федеральный округ, 109369, Россия

Conclusion

- ▶ The goal of this project was to identify areas of Moscow that are promising in terms of gyms location, in order to help stakeholders narrow down their search for the optimal location for a new gym. The distributions of gym density, office density, and population density were calculated based on Foursquare data. These locations were then clustered to create major areas of interest (containing the largest number of potential locations) and the addresses of these area centers were generated to be used as starting points for final stakeholder research.
- ▶ The final decision on the optimal gym location will be made by stakeholders based on the specific characteristics of the neighborhoods and locations in each recommended area, taking into account additional factors such as the attractiveness of each location, prices, etc.

The End

Thanks for attention