Coursera Data Science Project:

The Battle of Neighbourhoods

Measuring the efficient number of pharmacies in Moscow during Covid-19

Cover Sheet

Prepared by Vitaly Okolelov

https://github.com/OkolelovV

Date: 09/05/2020

Introduction

Background

Coronavirus Disease 2019 is an infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1] (see Figure 1). It started in December 2019 and has been continuing to this day. More than 3.5 million cases of the disease and more than 250,000 deaths have been reported worldwide [2]. The vaccine has not yet been invented, so there is still no effective way to fight the virus, except for self-isolation and following WHO recommendations. As a result, the number of active cases is growing, which increases the burden on hospitals and pharmacies around the world. Therefore, reducing the rate at which a population becomes infected would give hospitals more time to resolve the situation and would flatten the COVID-19 curve (see Figure 2).

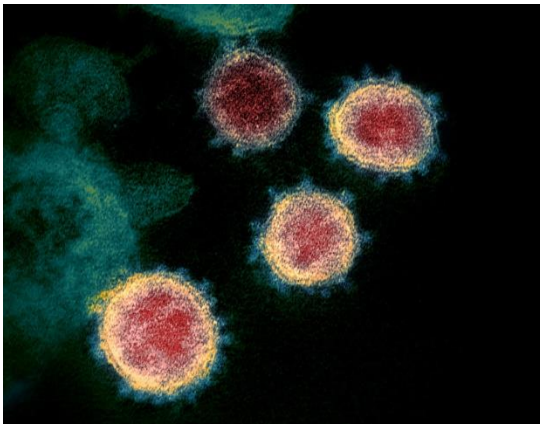

Figure 1. Scanning electron microscope image showing SARS-CoV-2 [3]

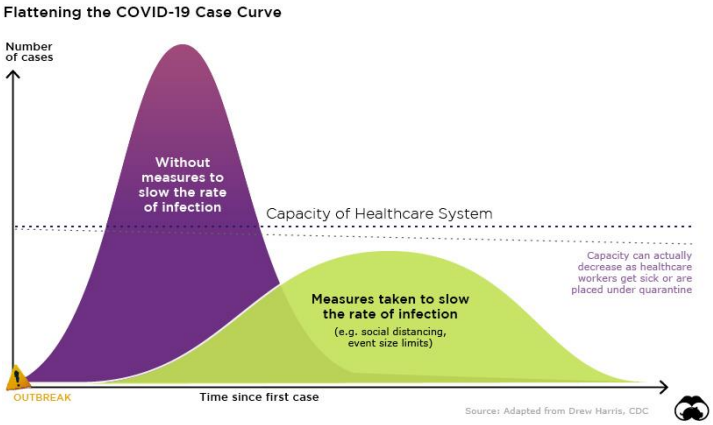

Figure 2. Flattening the COVID-19 case curve [4]

Description of the Problem

One way of the virus spread is between people during close contact, often through the small droplets that may occur during coughing, sneezing, and breathing [5]. The use of disinfectants and face masks helps to curb the spread of the disease, but some countries like Russia have imposed restrictions on the sale of medical products online to stop fraud. For example, in Moscow, the sale of masks is allowed only in pharmacies now.

After 12 May 2020, the wearing of masks is necessary in public places [6]. All Moscow pharmacy institutions must maintain prices and ensure the availability of a mask supply and other means of respiratory protection, which will be controlled by the Office of the Federal Antimonopoly service of Russia. However, Moscow is the second most populated city in Europe with over 12.6 million residents [7], which means that not all the districts may have enough medicine for everyone. Thus, the goal of this project is to identify the neighbourhoods where people might lack the means of protection against the virus.

Target audience

Providing information about pharmacies that might theoretically be in high demand could help suppliers to distribute medicines around the city during the Coronavirus disease. Moreover, if the results recommend the pharmacy infrastructure to be improved in these areas, then this could be considered by the local authorities. From the business perspective, this information could be useful for pharmaceutical companies, as opening new venues in high demand areas usually brings a good benefit for the companies.

Data

The number of pharmacies in every Moscow neighbourhood were estimated in this project. Then, the neighbourhoods were compared by the ratio of population to the pharmacies number and potentially problematic areas were found. To get the data required, the following sources of information were used:

- Wikipedia page about the Moscow neighbourhoods [8] provided the names, types, areas and populations of the neighbourhoods.
- NextGIS [9] provided a GeoJSON file that contains the coordinates of the neighbourhoods' administrative borders.
- OpenStreetMap [10] was used for determining the centre of every neighbourhood and making maps.
- Foursquare [11] was used for searching pharmacies in the neighbourhoods.

See figures below for the examples of the source data.

№	Флаг	Герб	Название района[7] / поселения[3][4]	Название соответствующего внутригородского муниципального образования: муниципального округа / поселения / городского округа[5]	Адми-нистра-тивный округ	Пло-щадь, км²	Насе-ление 2019[6], чел.	Плот-ность насе-ления 2019, чел. / км²	Пло-щадь жилого фонда (01.01. 2019)[7], тыс. м²	Жил-площадь на чело-века (01.01. 2010), м²/чел.
1			Академический	Академический	ЮЗАО	5,83	↗109 387	18762.78	2467,0	22,7
2			Алексеевский	Алексеевский	СВАО	5,29	↗80 534	15223.82	1607,9	20,5
3			Алтуфьевский	Алтуфьевский	СВАО	3,25	↗57 596	17721.85	839,3	15,5
4			Арбат	Арбат	ЦАО	2,11	↗36 125	17120.85	731,0	26,0
5			Аэропорт	Аэропорт	САО	4,58	↗79 486	17355.02	1938,7	25,9
6			Бабушкинский	Бабушкинский	СВАО	5,07	↗88 537	17462.92	1586,3	18,5
7			Басманный	Басманный	ЦАО	8,37	↗110 694	13225.09	1991,8	18,4

Figure 3. The example of Wikipedia data

Figure 4. The example of NextGIS data

Figure 5. The example of OpenStreetMap data

Figure 6. The example of Foursquare data

Methodology

First, the GeoJSON file with the neighbourhoods' administrative borders was used to build a map of Moscow (see Figure 7). After that, it was ensured that the file contains the different types of neighbourhoods and a map indicating different types of neighbourhood with different colours was drawn (see Figure 8). There are 125 municipal districts, 19 settlements, and 2 city districts. Total number of neighbourhoods is 146.

Second, the information about neighbourhoods was extracted from Wikipedia (see Figure 3 for the example of data). This information was processed and resulted in a dataframe with neighbour's name, area, population, and type (see Figure 9).

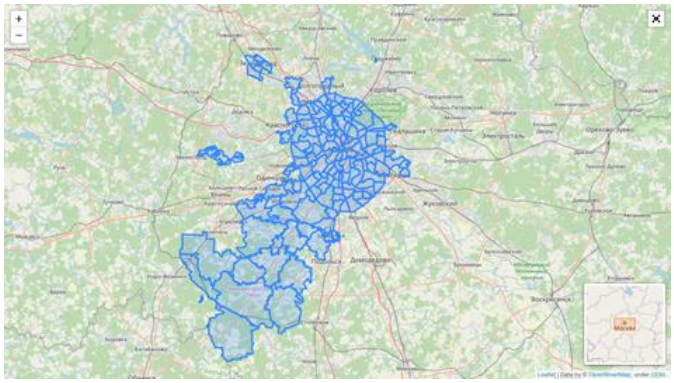

Figure 7. Moscow map

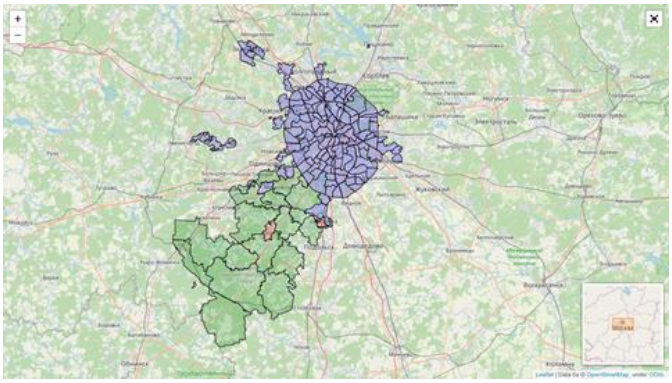

Figure 8. Different types of Moscow neighbourhoods. blue: municipal districts, green: settlements, red: city districts

	Neighborhood	Area	Population	Type
0	Академический	5830000	109387	район
1	Алексеевский	5290000	80534	район
2	Алтуфьевский	3250000	57596	район
3	Арбат	2110000	36125	район
4	Аэропорт	4580000	79486	район

Figure 9. The processed data from Wikipedia

	Neighborhood	Area	Population	Latitude	Longitude
0	Академический	5830000	109387	55.689359	37.577971
1	Алексеевский	5290000	80534	55.811044	37.648999
2	Алтуфьевский	3250000	57596	55.878695	37.586770
3	Арбат	2110000	36125	55.751199	37.589872
4	Аэропорт	4580000	79486	55.800504	37.543864

Figure 10. The dataframe after Nominatim requests

After that, Nominatim was used to get the coordinates of neighbourhoods' centres. The query included the names and the types of neighbourhoods to find the specific location, because the names only were not enough. The latitudes and longitudes were added to the dataframe (see Figure 10). The dataframe was saved to an excel file for the future use. Next, a choropleth map of the Moscow neighbourhoods was created (see Figure 11). Ranging was done by Population to check that the Choropleth method worked, and the centres were assigned correctly. In addition, GeoJsonTooltip was added to interactively show district borders and names when the mouse cursor is hovered over them. Manual adjustment was done for the centres that were not positioned properly.

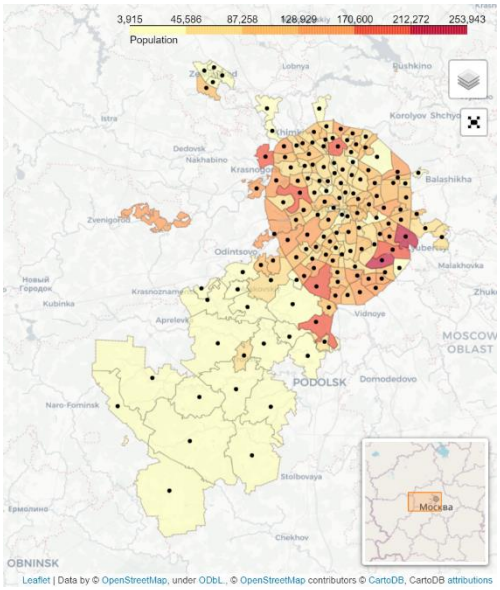

Figure 11. The choropleth Moscow map, ranged by Population with the centres of neighbourhoods

The next step was to calculate the radius for every neighbourhood depending on its area. The formula was

$$R = \frac{1}{2}\sqrt{A}$$

This would allow to roughly adjust the Foursquare queries, which will be used to for pharmacies searching in every neighbourhood. The Area column was dropped (see Figure 12).

	Neighborhood	Population	Latitude	Longitude	Radii
0	Академический	109387	55.689359	37.577971	1207
1	Алексеевский	80534	55.811044	37.648999	1150
2	Алтуфьевский	57596	55.878695	37.586770	901
3	Арбат	36125	55.751199	37.589872	726
4	Аэропорт	79486	55.800504	37.543864	1070

Figure 12. The dataframe after the radii calculation.

A supplementary function which searched the pharmacies in all the neighbourhoods was defined. The following parameters for the Foursquare SEARCH query were provided:

- categoryId = 4bf58dd8d48988d10f951735 (pharmacy ID in Foursquare);
- ll = the centre of a neighbourhood;
- radius = the radius for a neighbourhood calculated before;
- query = аптека (pharmacy in Russian).

After the function was applied to the neighbourhoods, the pharmacies found were saved in an excel file for future uses (see Figure 13). The total number of pharmacies found was 3636.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
0	Академический	55.689359	37.577971	Аптека «Академическая»	55.686728	37.573103
1	Академический	55.689359	37.577971	Семейная Аптека	55.691803	37.575324
2	Академический	55.689359	37.577971	Аптека на Винокурова	55.686727	37.590515
3	Академический	55.689359	37.577971	Аптека 24 «Не болей»	55.684468	37.569058
4	Академический	55.689359	37.577971	Аптека Столички #56	55.688735	37.590941

Figure 13. The dataframe containing Moscow pharmacies

There is one region called Кунцево (Kuntsevo). It consists of three main areas and the search was made only in one. The search operation was repeated for the rest two, and the results were added to the dataframe. However, there were 1392 pharmacies that belonged to several neighbourhoods at the same time due to the way the Foursquare search was used. Therefore, these repeating pharmacies needed to be sorted out, because this may have seriously affected the results. Only the copy that is the closest to the centre of a nearby neighbourhood will remain, and the rest will be discarded.

	Old index	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
0	3539	Краснопахорское	55.417193	37.272232	Аптеки Столицы № 78	55.412298	37.184372
1	3543	Михайлово-Ярцевское	55.402712	37.152744	Аптеки Столицы № 78	55.412298	37.184372
2	3579	Первомайское	55.501905	37.212795	Аптека «Диалог»	55.470707	37.296756
3	3600	Троицк	55.479412	37.298031	Аптека «Диалог»	55.470707	37.296756
4	3583	Первомайское	55.501905	37.212795	Аптека 36.6	55.471532	37.295992
5	3605	Троицк	55.479412	37.298031	Аптека 36.6	55.471532	37.295992
6	3584	Первомайское	55.501905	37.212795	Горздрав	55.471709	37.296524
7	3611	Троицк	55.479412	37.298031	Горздрав	55.471709	37.296524
8	3535	Десёновское	55.501369	37.362555	Горздрав	55.484621	37.305146
9	3610	Троицк	55.479412	37.298031	Горздрав	55.484621	37.305146

Figure 14. The dataframe with duplicates

All duplicates were put into a new dataframe (see Figure 14), where the maximum number of repeats turned out to be 3. Next, while iterating over the duplicates dataframe and comparing every two or three pharmacies, the indices of the unappropriated pharmacies were saved to a new list. This list was then used to remove the found duplicates from the pharmacies dataframe. After this operation, the pharmacies number in every neighbourhood was counted (see Figure 15). The total number of pharmacies was 2934. Additionally, it was decided to show pharmacies distribution on a map (see Figure 16) to check the correctness.

Neighborhood	Venue
Академический	42
Алексеевский	27
Алтуфьевский	8
Арбат	25
Аэропорт	27

Figure 15. The pharmacies number in every neighbourhood

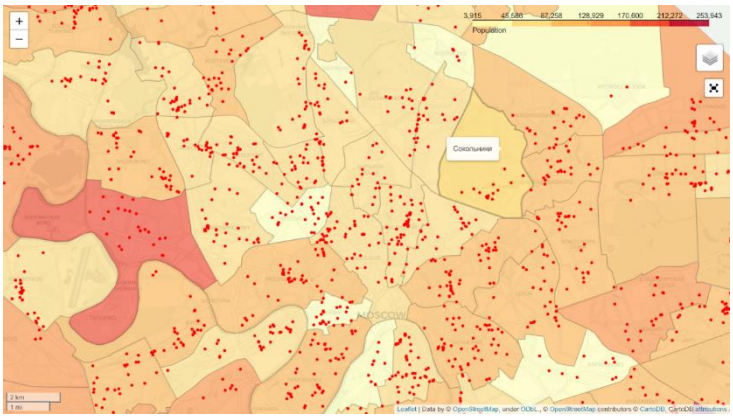

Figure 16. A fragment of pharmacies distribution map

The number of pharmacies were added to a "Pharmacy number" column, then the population per one pharmacy number ratio was calculated (see Figure 17). Finally, a choropleth map was created, where the colours corresponded to the population per one pharmacy (see Figure 18).

	Neighborhood	Population	Latitude	Longitude	Pharmacy number	Population / Pharm num
0	Академический	109387	55.689359	37.577971	42.0	2604.452381
1	Алексеевский	80534	55.811044	37.648999	27.0	2982.740741
2	Алтуфьевский	57596	55.878695	37.586770	8.0	7199.500000
3	Арбат	36125	55.751199	37.589872	25.0	1445.000000
4	Аэропорт	79486	55.800504	37.543864	27.0	2943.925926

Figure 17. The dataframe with the ratio

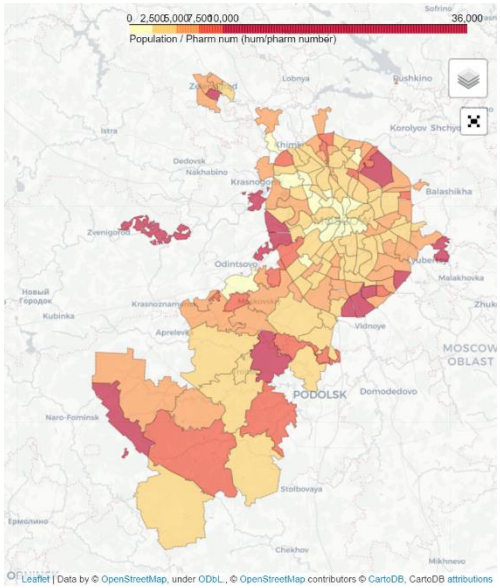

Figure 18. The choropleth map indicating the population per pharmacies number ratio

Neighborhood	Population	Pharmacy number	Population / Pharm num
Некрасовка	71943	2.0	35971.500000
Десёновское	22413	1.0	22413.000000
Можайский	138610	8.0	17326.250000
Бирюлёво Западное	88672	6.0	14778.666667
Киевский	13968	1.0	13968.000000
Кунцево	152364	11.0	13851.272727
Метрогородок	39024	3.0	13008.000000
Капотня	32399	3.0	10799.666667
Бирюлёво Восточное	155863	15.0	10390.866667
Старое Крюково	31053	3.0	10351.000000
Орехово-Борисово Северное	132265	13.0	10174.230769

Figure 19. The list of potentially problematic neighbourhoods

Results

To evaluate the limits, it was assumed to be 5 minutes to serve one client in a pharmacy (12 people per hour). The working day is 8 hours, then around 100 people would be served per day and around 3000 people monthly. According to Pew Research Center [12], the average number of people in one Russian household is 3.2 and, say, every family normally went to a pharmacy once per month. Thus, the pharmacies that serve more than 10000 people were considered as in high demand. These neighbourhoods could be potentially problematic areas and were shown in dark red on the map (see Figure 18). The full list of these neighbourhoods can be seen in Figure 19.

In order to verify the results, the problematic neighbourhoods were checked again. The Foursquare search was made again with the triple radius parameter, and the pharmacies were counted manually from the map. As a result, only eight neighbourhoods left that have population per one pharmacy more than 10000: Десёновское, Можайский, Бирюлёво Восточное, Киевский, Метрогородок, Кунцево, Капотня, Орехово-Борисово Северное (see Figure 20).

Neighborhood	Population	Pharmacy number	Population / Pharm num
Десёновское	22413	1	22413.000000
Можайский	138610	8	17326.250000
Бирюлёво Восточное	155863	10	15586.300000
Киевский	13968	1	13968.000000
Метрогородок	39024	3	13008.000000
Кунцево	152364	13	11720.307692
Капотня	32399	3	10799.666667
Орехово-Борисово Северное	132265	13	10174.230769
Бирюлёво Западное	88672	9	9852.444444
Воскресенское	9822	1	9822.000000
Ярославский	98108	10	9810.800000
Северное Тушино	165762	17	9750.705882
Щаповское	9572	1	9572.000000

Figure 20. The list of neighbourhoods with the high ratio

It is strongly recommended to pay additional attention to these eight neighbourhoods, since their population per pharmacy number exceeds 10000, therefore, their pharmacies may have problems supporting people during the virus. In addition, there are also five neighbourhoods whose ratios exceed 9500, which is close to the estimated maximum pharmacy capacity: Бирюлёво Западное, Воскресенское, Ярославский, Северное Тушино, and Щаповское. These neighbourhoods could take measures to improve the situation as well.

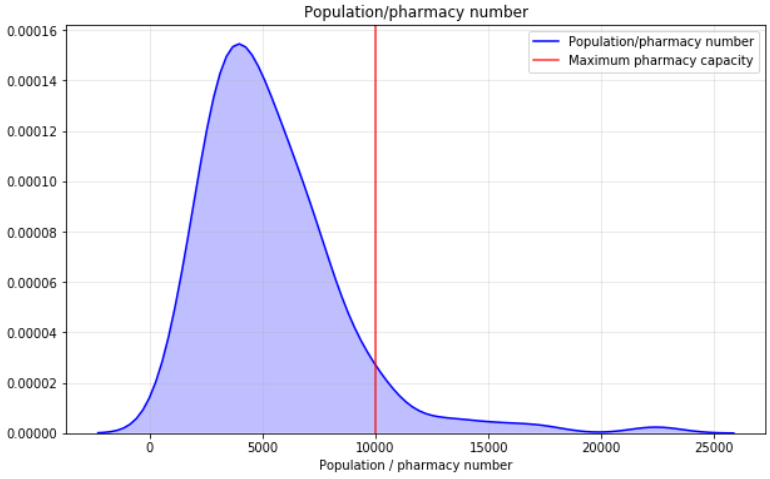

Figure 21. Kernel density estimation of population per pharmacy number

The average population per pharmacy was 5382, which is significantly lower than the maximum pharmacy capacity. From the Kernel density estimation of population per pharmacy number (see Figure 21), most districts have a sufficient number of pharmacies (only 5.48% do not). This means that, in general, Moscow can successfully meet the needs of people during the virus epidemic, provided that there are enough medical supplies.

Discussion

To explore the results further, a scatter graph of population per pharmacy number as a function of population density was plotted (see Figure 22). It would be incorrect to compare the population per pharmacy number with the neighbourhood's area, because some of them contain large areas of non-residential space, such as parks. However, it is possible to compare the ratio with the population density.

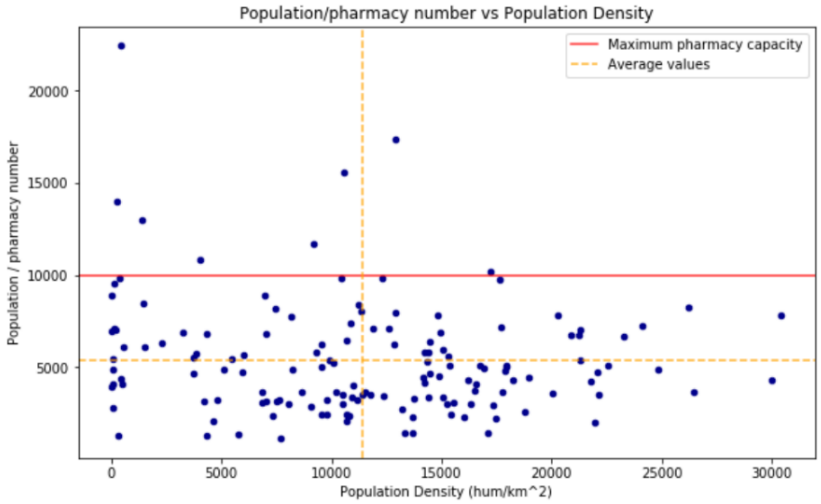

Figure 22. The graph of population per pharmacy number as a function of population density

The graph showed that the problematic neighbourhoods are more common in areas with lower-than-average population densities (less than 11419 people per km^2). However, there are two neighbourhoods that are at risk and have a higher population density: Можайский, Бирюлёво Восточное, and Орехово-Борисово Северное (see Figure 23). These and another two neighbourhoods (Ярославский and Северное Тушино), whose ratios exceed 9500, are recommended paying the greatest attention to the current situation.

Neighborhood	Population	Pharmacy number	Population / Pharm num	Area (km^2)	Population Density (hum/km^2)
Можайский	138610	8	17326.250000	10.73	12917.986952
Орехово-Борисово Северное	132265	13	10174.230769	7.67	17244.458931
Ярославский	98108	10	9810.800000	7.99	12278.848561
Северное Тушино	165762	17	9750.705882	9.40	17634.255319

Figure 23. The neighbourhoods that are at risk and have high population density

The first thing to notice, is that the results obtained are approximate and highly depend on the Foursquare database of pharmacies in Russia. One possible way to improve the results is to use Yandex Maps (like Foursquare, but a Russian service) instead. These maps are the most accurate maps in Russia and are synchronized with other Russian services.

Another way of improvement is to manually or automatically divide the neighbourhood into small subareas. Since a Foursquare search returns no more than 50 venues, this would allow to find more pharmacies. Moreover, if the division is performed well, the whole area of every neighbourhood would be covered, which was not guaranteed in the current method.

Conclusion

In this project, a search for pharmacies in Moscow's neighbourhoods was conducted. As a result, a list of neighbourhoods where pharmacies might potentially be in high demand during the virus was compiled. Additional recommendations were also given to the neighbourhoods in which pharmacies are close to the maximum pharmacy capacity. Moreover, it was found that the problematic neighbourhoods tend to be in areas where population density is lower than 11419 people per km^2, which is the average value. However, there were three neighbourhoods with an exceeded maximum population per pharmacy ratio and a high population density. In overall, the average population per pharmacy was 5382 and only 5.48% of all neighbourhoods do not meet the criteria. Thus, the analysis showed that most Moscow pharmacies should be able to successfully operate during Covid-19, provided that enough medical supplies are available.

References

[1] Mayo Clinic, “Coronavirus disease 2019 (COVID-19),” [Online]. Available: https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963. [Accessed 5 May 2020].

[2] Worldometer, “COVID-19 CORONAVIRUS PANDEMIC,” 5 May 2020. [Online]. Available: https://www.worldometers.info/coronavirus/. [Accessed 5 May 2020].

[3] National Institute of Allergy and Infectious Diseases, “Novel Coronavirus SARS-CoV-2,” 14 February 2020. [Online]. Available: https://www.niaid.nih.gov/. [Accessed 5 May 2020].

[4] N. Routley, “Infection Trajectory: See Which Countries are Flattening Their COVID-19 Curve,” 16 March 2020. [Online]. Available: https://www.visualcapitalist.com/infection-trajectory-flattening-the-covid19-curve/. [Accessed 5 May 2020].

[5] World Health Organisation, “Advice on the use of masks in the context of COVID-19,” 6 April 2020. [Online]. Available: https://www.who.int/publications-detail/advice-on-the-use-of-masks-in-the-community-during-home-care-and-in-healthcare-settings-in-the-context-of-the-novel-coronavirus-(2019-ncov)-outbreak. [Accessed 5 May 2020].

[6] Правительство Московской области, “Масочный режим начали вводить в Подмосковье,” 1 May 2020. [Online]. Available: https://mosreg.ru/sobytiya/novosti/news-submoscow/srochno-v-podmoskove-nachali-vvodit-masochnyy-rezhim. [Accessed 5 May 2020].

[7] Population Stat, “Moscow, Russia Population,” 2020. [Online]. Available: https://populationstat.com/russia/moscow. [Accessed 5 May 2020].

[8] Wikipedia , “Список районов и поселений Москвы,” 2 October 2019. [Online]. Available: https://ru.wikipedia.org/wiki/Список_районов_и_поселений_Москвы. [Accessed 1 May 2020].

[9] NextGIS, “MAP AND DATA,” 2020. [Online]. Available: https://data.nextgis.com/en/region/RU-MOW/. [Accessed 30 April 2020].

[10] OpenStreetMap, “OpenStreetMap,” 2020. [Online]. Available: https://www.openstreetmap.org/#map=9/55.7672/37.4758. [Accessed 5 May 2020].

[11] Foursquare, “Foursquare: City Guide,” 2020. [Online]. Available: https://foursquare.com/city-guide. [Accessed 5 May 2020].

[12] Pew Research Center, “Religion and Living Arrangements Around the World,” 2019.