**An Exercise in Open Data: Mapping Istanbul’s Health Services**

**A store inside of a building

Description automatically generated**

The world has gone through tough times in the past four months. The coronavirus has imposed a “new normal” upon us. Both the virus and the new normal that it brought along has been very visible in my hometown Istanbul.

EMPTY STREETS HERE

Istanbul is a giant metropolis with 16 million inhabitants packed in a not – so – large landmass. Controlling the virus here has required severe lockdowns as to prevent the local intensive care units from being flooded with patients. I have done my civic duty and practiced self - quarantine to prevent the images of overwhelmed ICU units making rounds on the TV channels and on the social media. After all, the situation of health institutions, the burden on the shoulders of healthcare workers and the reports of both tragedy and hope had been making frequent appearances on all media platforms. Health and the fear of losing it was the main topic for a long while.

All this talk reminded me of an important part of the infrastructure that made the functioning of my city possible: its healthcare services. Like my previous attempt at gaining an insight about my city’s green spaces(LINK), I wanted to learn more about the distribution of healthcare services across Istanbul. I wanted to try and reveal existing patterns and possible inequalities that characterized the health services in my city.

**A variety of services**

The first question that I had in my mind was the extent of the healthcare services provided in the city. Luckily, I only had to interact with two or three types of healthcare institutions so far in my life. However, there surely were more than just the three types that I encountered with. What were they?

A close up of a logo

Description automatically generated

As it can be seen above, **there are a total of 3522 healthcare institutions in Istanbul distributed across 33 categories.** These categories include not only the familiar hospital and the general family practitioner, but also some very niche health institutions such as the tuberculosis dispensary.

It is impossible to make any comments about the distribution of healthcare institutions across Istanbul by looking at these big numbers alone. **To see the spatial distribution of the healthcare institutions across the greater Istanbul area, I enlisted the help of a choropleth map combined with a bar chart:**

A picture containing map, text

Description automatically generated

This is more informative. The map shows at first glance that there is a vast difference in between different districts when it comes to the total number of healthcare institutions.

**Is this map enough evidence to say that there exists an unfair distribution in healthcare services?** My first instinct was to say yes. However, taking a second or two to think again revealed to me some obvious flaws that this map has:

* My choropleth map does not account for the population difference in between districts. Districts to the north of Istanbul and districts that are on the peripheries have far few people living in them when compared to more central districts. **Comparing numbers alone without normalizing for population in such a situation will most probably lead to a faulty judgment.**
* Secondly, health institutions differ on their level and capacity of care. The dataset that I am basing my analysis on considers a broad range of medical institutions, as it can be seen in the first bar plot. Considering this, it is very normal for most health institutions to center around well – connected districts that offer fast access to healthcare service seekers coming from other districts.

I know from experience that the districts that rank in the top five of the total health institution ranking are districts that are well connected to nearby ones. Kadıköy, the district with most health institutions, serves as a major commerce center in the Anatolian side whereas Şişli, Beşiktaş, Bahçelievler and Fatih are also all very populated and very central districts.

**We can conclude that the total number of health institutions in a district is not a good metric for measuring any potential inequality.**

Now that we have established that the above choropleth map is not of much use, we must base our analysis on other parameters.

**Different Levels of Care**

As we have seen above, different health institutions offer healthcare services at different levels and they have different capacities. We must account for this reality in our analysis.

A close up of a device

Description automatically generated

I decided to split Istanbul’s health institutions into three different categories:

* Those who provide general hospital – level healthcare (hospitals, research hospitals, polyclinics etc.)
* Those who offer specialized care (dental clinics, gynecology and obstetrics clinics, physical therapy centers etc.)
* Those who offer individual or community level care (general practitioner, community health centers, home – visit care centers etc.)

**Healthcare institutions that provide hospital – level and specialized care serve a much larger geography at a larger capacity. Therefore, their analysis requires a few key metrics** (total staff, medical services present, ease of access etc.) **that we cannot extract from the dataset that we have.** However, healthcare services that provide individual or community level care are most often bound to the districts that they operate in. We can reach some insights about their distribution from the data that we have.

**Distribution of First Step Healthcare Services**

**A picture containing text, map

Description automatically generated**

**In order to analyze the distribution of first step healthcare institutions at the district level, we are once again resorting to the use of a choropleth map.** However, this time we are considering two additional metrics that can offer us more information about whether or not there is an inequality in the distribution of first step healthcare institutions: the Pearson’s correlation coefficient of population – number of first step healthcare institutions and of yearly average household income – number of first step healthcare institutions.

The choropleth map shows that there is a discrepancy in between different districts. However, we can claim that this occurrence is expected when we look at the strong linear correlation in between the population of a district and the number of first step healthcare institutions that same district has. This would mean that such a spatial distribution is not the signifier of any underlying inequalities at the district level. **First step healthcare institutions provide preventative and first step treatment to families and individuals in a local area. The more population a district has, the more first step healthcare institutions it houses.**

Our claim is further strengthened when we consider that the linear correlation in between the yearly average household income of a district and the number of first step healthcare institutions that same district has is nearly zero. **The number of first step healthcare institutions in a district does not increase the more money the residents of that district has.**

Remember that we talked about normalization before? This choropleth map is not normalized. There is a high linear correlation in between population and health institution count. **I believe that, in such a situation, normalizing for population will not offer any additional insight over a normal choropleth map.**

**Slicing the Dataset Further: Private and Public Healthcare Institutions**

Looking at the distribution of first step healthcare institutions at district level did not reveal any signs of inequality.

The seemingly fair distribution can be explained by looking at the providers of first step healthcare services: **Nearly all first step healthcare institutions are state – owned.** There might be a central planning force at work that distributed first step healthcare institutions according to population.

This possible explanation leads to another way of searching for inequalities in the distribution of healthcare services**. Perhaps classifying institutions as private or public could lead to a deeper understanding?**

**A screenshot of a social media post

Description automatically generated**The side – by – side bar chart above shows the private – public split in the eight difference healthcare institutions that the private sector has a presence in.

**Right away we can see the enormous difference in between some types of health institutions.** For example, there are only 33 state – owned dental health centers in Istanbul while there are nearly thirteen times more private counterparts. **There are less public dental health centers than there are districts in Istanbul.** A similar pattern can be seen when we look at physical therapy centers and dialysis centers.

We can now look at how private health institutions are distributed on the district – level:

**A picture containing text, map

Description automatically generated**

It is now obvious that some districts have far more private health institutions than others. Districts like Kadıköy, Şişli and Beşiktaş are once again at the top of the private health institutions ranking just as they were for the overall health institution ranking. We can also infer from the scatterplot below that there is a mild linear correlation in between the average yearly household income of a district and the number of private health institutions in the same district.

Perhaps looking at the distribution of some specific healthcare institutions can lead to more interesting findings. Let’s now look at the distribution of dental health centers and physical therapy centers more closely:

A close up of a map

Description automatically generated

By looking at the side – by – side bar charts, we can say that the nearly every district except three (Şişli, Maltepe, Kağıthane) has at least one state – owned dental health center. However, we can also see that the number of private dental health centers vary wildly in between different districts. This can be interpreted as a sign of inequality: **No matter which district you are in, you have more access to dental health services if you can pay for them. Judging by the amount of public dental health centers in some districts, not having enough money might hamper your right to dental health service access since you will have to deal with an overcrowded center.**

The scatterplot at the lower left corner shows us that the mild linear correlation that we have noted for the overall count of private health institutions also seems to hold for dental health centers. However, we can see that the Pearson’s correlation coefficient here is weaker than what we have seen before. **This implies that there exists one or more private healthcare institution whose distribution over districts is not as strongly correlated with yearly average household income as the overall number of healthcare institutions is with the same metric.**

**Let’s look at one more private healthcare institution to confirm the statement above:**

A screenshot of a cell phone

Description automatically generated

By looking at how physical therapy centers are distributed across various Istanbul districts, we can once again see that some districts have far more private physical therapy centers than others. **Considering that there are more districts that do not even have one state – owned physical therapy center when compared to the distribution of dental health centers, this might once again be the signal of an underlying inequality.**

**However, the more important finding that we can talk about is the fact that the number of private physical therapy centers in a district is not correlated with average yearly household income of that same district at all.** The same holds true for population. Different districts have the same number of private physical therapy centers even though they differ wildly in both population and average yearly household income.

Now that we have found at least one private healthcare institution that confirmed our statement about lower levels of correlation, it might be useful to look at levels of correlation across all private health institution types:

**A close up of a piece of paper

Description automatically generated**

By looking at this small multiple of scatterplots, we can clearly see that there is no linear relationship at all in between the number of some types of private healthcare institutions and yearly average household income.

**Caveats and Conclusion**

**Caveats**

In the brief analysis above, I tried to uncover some signs of possible inequality that existed in the distribution of healthcare institutions. My method of analysis has one big caveat. This caveat has to do with the way I am trying to quantify “inequality.” First of all, by zooming in only so far as the district level, I am ignoring the fact that different health institutions have different dynamics that allow them to have a broader range of service. **For example, hospitals and polyclinics that are in a well – connected location within the city can serve people farther away than the immediate district they are in.**

**Secondly**, merely counting the number of health institutions in a district is not a great indicator of healthcare access within that district. Without any way to quantify the exact capacity and quality of care of a specific healthcare institution, my analysis can only be considered as a shallow overview.

**Conclusion**

Still, I believe that this analysis that is based on open data only has allowed me to better understand the city that I live in. Perhaps in the future I can re-do such an analysis with a more granular dataset and with a deeper understanding of the factors that might stimulate inequality within a city. Perhaps then I can fix the shortcomings that I have mentioned above.