Project

Data Visualization

Benjámin Olivér Maday

2023-11-06

0.1 Introduction

In the spring of 2021 I've spent one month in the beautiful city of Paris. The whole city was under a major lockdown due to the pandemic, but Parisians being Parisians did not really follow the regulations, nor did they obey the curfew. This was fairly strange to me. In my home country, Hungary, the enforcement of such regulations is quite harsh, thus not a lot of people would disobey and risk a hefty fine. A few months later I visited a sociology lecture at a different university from mine. Though I was mainly interested in one of the audience members (why would anyone go to university lectures in their free time if not for a girl), I did pay enough attention to catch one of the lecturer's more debatable sentences: "Each countries' current pandemic case and death numbers can easily be explained by their societies values.". I found this quite intriguing, especially the "easily" part. The following year, I wrote my thesis about this topic. I was interested in whether the infamously entitled French are affected worse due to their hubris, or is it only a sociologists mumbling, thinking their theories can be finally tested in a crisis. The findings? By creating a panel regression model I showed that the social values of a country only have marginal effects on their pandemic statistics, but they do have some.

I did not use any visualization in my thesis (honestly speaking, I only knew R from Youtube tutorials and I did not have the capacity nor the willingness to learn visualization, but it did encourage me to pursue a career in data science). In this project I would like to correct that and retroactively insert graphs into my paper which can illustrate a few of my findings. However this essay will not follow the logic of the thesis. I am not aiming on recreating any findings from it, but will occasionally note whether the original model arrived at the same conclusions or not.

Making objective judgements about a countries' culture is hard in itself, but creating a framework according to which we can compare them could be considered near impossible, yet a lot of scientist did try to do exactly that. In my comparison I will be using Hofstede's and Schwartz's cultural value indices. Both of them use a survey with which they try to quantify given societies cultural leanings, what do they value more or less in different "cultural dimensions". I will describe the dimensions necessary for our discussion in the part about variables.

The original research question was: Do cultural values have an effect on a given countries' Covid-19 statistics? In this project I would like to illustrate that we can see weak connections between them, but it is more constructive to focus on more

obvious predictors. Thus I would like to change the question to what are the main predictors of the number of cases, death rate and vaccination rate?

0.2 The dataset

I am trying to recreate the dataset that I used more than one and a half year ago. The statistics are from Our World in Data's coronavirus page (https://ourworldindata.org/coronavirus). Since I wrote my thesis in April 2022, I am only using information available up until 2022-04-21. For the visualiyations I will be using two datasets: one which contains the daily Covid-19 statistics of every country and another one I created which summarizes their most important statistics at 2022-04-21. The two sets of data contain observations about 48 countries (these are the countries for which we have both Hofstede's and Schwartz's indices), and 39 variables. The variables that I will use in this report are the following:

Table 1: Description of variables

Variable name	Variable description	Data type
Total cases per million	The total number of Covid-19 cases in a country per a million people. Countries use different methods to determine what constitutes as a confirmed case.	integer
Total deaths per million	The total number of Covid-19 related deaths per a million people	integer
Fully vaccinated per hundred	The number of fully vaccinated individuals per a hundred people. Full vaccination means that they received the necessery amount of vaccines in order to gain immunity against the virus (for example two Pfizer shots)	integer
GDP per capita	Whether a country is in the highest, lowest, or middle third of GDP per capita from these 48 countries	categorical
Median age	The populations median age	continous
Life expectancy	The countrie's life expectancy	continous
Hospital beds per thousand	The number of available hospital beds per a thousand people	continous

Variable name	Variable description	Data type
Individualism	A countries score on Hofstede's identity dimension. A high score means a very individualistic, while a low score means a very altruistic/collectivist society	continous
Power distance	A countries score on Hofstede's power dimension. A high score means a society accepts large differences between individuals and respects hierarchies. A low score suggests a more egalitarian society.	continous
Affective autonomy	Schwartz differentiates between two kinds of individualisms/autonomies. Affective autonomy is the independent pursuit of pleasure, seeking enjoyment by any means without censure. In many societies there are limits when affective autonomy leads to taking banned substances or acting in ways that distresses or harms others. A high score suggests a very open and accepting society	continous
Intelectual autonomy	Intellectual Autonomy is the independent pursuit of ideas and thought, whether it is theoretical, political or else. A high score means a society that is open to new ideas and debates, even if these ideas can be harmful, while a low score suggests an Orwellian thought police from the world of 1984.	continous
Hierarchy	A country's score in Schwartz's hierarchy dimension. In hierarchical cultures, there is a clear social order, with some people in superior positions while others are in inferior positions. This slightly differs from Hofstedes' power distance, since it is more about the states or other communaly controled entities legitimacy.	continous

If the reader can still remember, the winter periods were the harshest just like with the common flu. This can be seen on Figure 1. The smoothed lines mostly follows countries in the northern hemisphere, thus it is safe to say that the two winter periods caused the biggest spike in both infections and Covid related deaths.
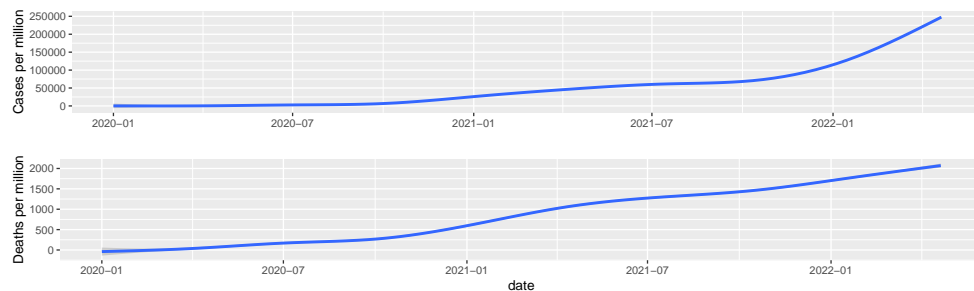

Figure 1: The number of Covid-19 cases and related deaths per million, a worldwide average.

Before beginning working on the visualizations I have three a priori hypothesis:

1. Individualism, affective autonomy or intelectual autonomy will have visible correlation with pandemic cases, deaths and vaccination rates. More individualistic, younger individuals will be less likely to take the pandemic seriously, thus spreading the virus more rapidly. Since they do not fear the disease they are less likely to get vaccinated as well, since it's individual utility is small.

3

2. A higher power distance or hierarchy score will result in higher vacination rates. If hiearchies are accepted in a given society they will believe in experts and their governments more, leading to a smoother voluntary vaccination program.

3. Developed countries will have higher infection rates, but lower death rates. This is due to the combination of them being more individualistic, but having a better healthcare system.

0.3 Number of cases

If we take a look at the number of Covid-19 cases per country it is quite hard to see any correlation to any of the cultural indexes. Indeed during my research I didn't find any significant effect of neither individualism nor hierarchy on the occurence of the virus.
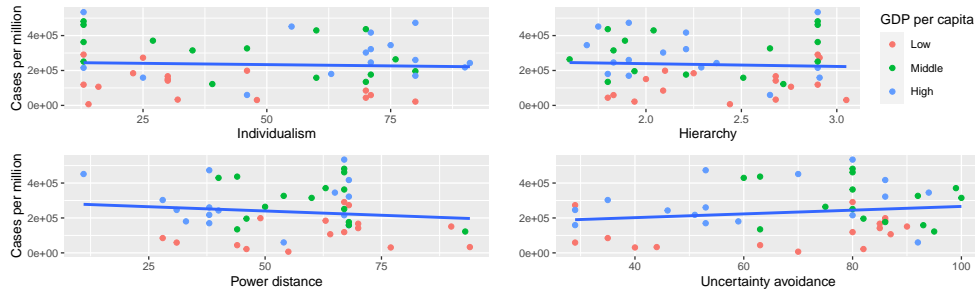

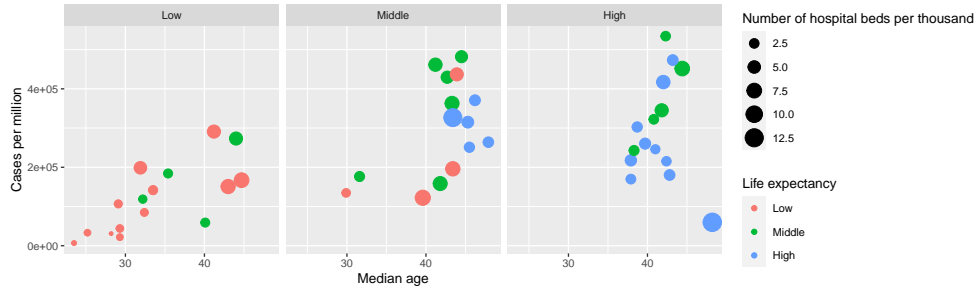Figure 2: Number of cases per a million people plotted against the aforementioned cultural indices

Figure 3: Number of cases per a million people in relation to the median age, faceted by the countries GDP per capita ranking, coloured by life expectancy

On the other hand, more traditional metrics did prove to be useful. On Figure 3 we can see that countries with the lowest median age and life expectancy experienced fewer cases of Covid-19. Countries with a higher per capita income had more cases, but it seems that countries with a life expectancy between 77-82 had the most amount of cases. This can be attributed to a lot of different factors (just as an obvious example more business hubs and higher spendable income means more mobility which leads to an easier spread of the virus, better infrastructure could potentially means higher report rate), but it also gives us an answer to why it is easy to think that individualism has a significant effect on the infection numbers. Well-off countries are

usually considered to be more individualistic (even tho it is not the case in a lot of instances, take for example Sweden) and less likely to accept hierarchies thus giving the illusion that individualism causes the easier spread of the virus.

0.4 Death rate

Examining the relationship between cultural indices and death rates yielded more results, tho they were not the expected ones. Both intelectual and affective autonomy seem to have a negative correlation with the number of deaths. This contradicts the theoretical background, since both of them are "parts" of the overall individualism in a society. Power distance however seems to increase the death rate, which again refutes what one would think according to Hofstede's index. The more acceptance of power structures would suggest that people are less likely to violate curfews or other government regulations, resulting in less cases and thus, smaller mortality.
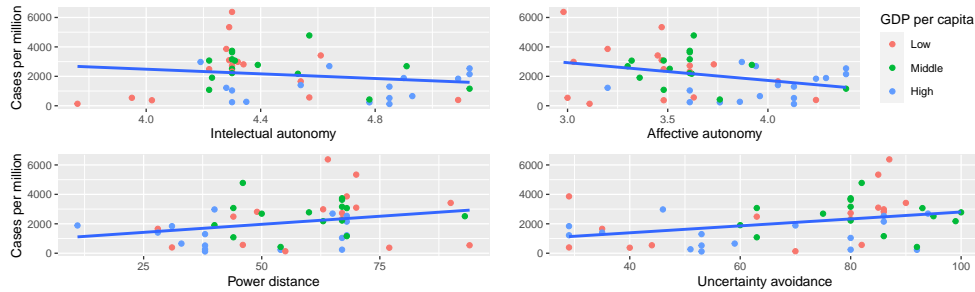

Figure 4: Number of cases per a million people plotted against the aforementioned cultural indices

Looking at our standard metrics, we can discover an interesting phenomena. Even though countries with a higher per capita earning, longer life expectancy and older population had a bigger quantity of cases their death rates are still low. This is most likely thanks to a well developed healthcare, meaning even though more people get the virus, they can be treated well in a hospital. The three countries with the biggest death rate are Peru, Bulgaria and Hungary of course, which is a (sadly) textbook example of what happens when governments systematically neglect the healthcare system.
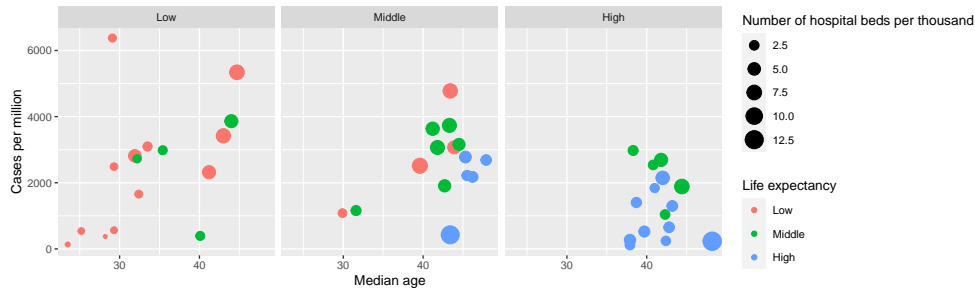

Figure 5: Lower income countries have a higher death rate than high income countries

0.5 Vaccination

Perhaps not so suprisingly for some, the biggest impact of societal customs were on the roll-out of vaccines. The number of fully vaccinated people is mostly dependent on the availability of the vaccine. This can be clearly seen on the animated plot (uploaded seperately), as nearly all countries have a hike when they get their hands on a bigger collection of them regardless of their cultural values. It is quite safe to say (at least in Europe), that by the spring of 2022 everyone was able to get vaccinated if so they wished, thus we can also conclude that Individualism had a significant effect.
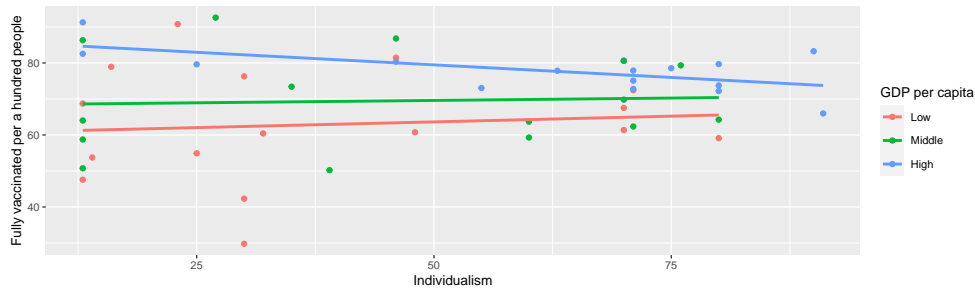

Figure 6: The vaccination rate in relation to a countrie's individualism index. There is a negative correlation in higher income countries

If we take a look at Figure 6 we can see that as individualism increases so does the number of fully vaccinated people. This wouldn't be constistent with what we think individualism would do, but if we only look at high income countries we will meet our expectations. In high income countries, the more collectivist a society is the more likely they will be to get vaccinated. This seems in line with what common sense would dictate.
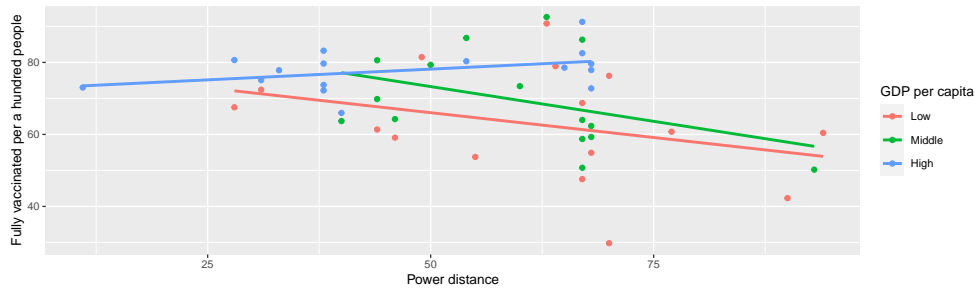

Figure 7: Power distance seems to negativelly correlate with the vaccination rate

Having said that, Hofstede's power distance does not seem to behave in the way one would expect. The larger the acceptance of hierarchical structures the lower the vaccination rate is. The lack of vaccines could somewhat explain this in lower income countries but not in middle ones. This contradicts my theory that governments are capable of doing large scale vaccination programs if their authority is seen more legitimate. The only category where this hypothesis seems to be correct is in countries with a high GDP per capita. The findings are identical if we look at Schwartz's hierarchy dimension as well.

0.6 Conclusion

The findings presented in this project are self contradictory. I was really interested whether such an abstract, and frankly, theoretical thing as a "cultural value" can have an effect on our everyday life, and the differences between cultures could explain their performance in tackling a worldwide pandemic. The answer: maybe. It is very hard to draw any clear conclusions from the data. They do not seem to affect the spread of the virus at all, even though that would have been the most obvious connection between them. They do seem to correlate with death rates, but I would suggest that there is a confounding variable (like more advanced nations tend to have a more individualistic society) behind this relation. Vaccination however seems to be partially explainable by individualism. I do believe, that there is some kind of connection, between cultures and their reactions to such worldwide catastrophies, but I was unable to find any conclusive evidence in this report.