IT UNIVERSITY OF COPENHAGEN

Travelling through the pandemic

Communicating how the COVID-19 pandemic influenced the mobility of different socioeconomic groups around the world

Bachelor's Thesis

GitHub repository: https://github.com/deerhjorth/bachelor22

Website (Data Journalism piece): https://deerhjorth.github.io/bachelor22/

BSc in Data Science

Authors: Anna Mejlhede Jensen & Christian Hjorth

Emails: anmj@itu.dk & chhj@itu.dk

December 15th, 2022

Supervisor: Michele Coscia, mcos@itu.dk

ABSTRACT

This paper explores, visualizes, and communicates how the COVID-19 pandemic and the subsequent lockdown measures have had a profound impact on the daily lives of people around the world, including their mobility patterns, through data journalism. In particular, the lockdowns have disproportionately affected individuals from different socioeconomic backgrounds, with some groups facing greater challenges in terms of mobility and access to essential services. This thesis explains the rationale for different design choices made to create visually pleasing and informative data visualizations as well as the considerations made for the implemented website with a scrollytelling format used to convey technical information to a general audience with varying technical and mathematical understanding. Further, this thesis discusses whether this intended goal was achieved and which steps of the process could be altered in the future to achieve this goal and improve our current product. This thesis concludes that while the goal of visualizing complex results to a broader audience was achieved, there were some substantial limitations along the way. These are related to using third-party software as a visualization tool instead of coding from the bottom, as well as information sparsity in low-middle-income countries (LMICs).

Keywords:
`Scrollytelling, data visualization, data journalism, information sparsity, COVID-19 pandemic, LMICs`

CONTENTS

ACKNOWLEDGEMENTS

We would like to extend our gratitude to our supervisor, Michele Coscia, for his support and guidance throughout this project. Although this type of project was new for both ourselves as well as for Michele, he was able to provide valuable insights and direction. We appreciate his mindset, creativity and willingness to support our motivation for this unusual Data Science Bachelor's thesis.

We would also like to thank coffee from Analog for keeping us alive.

1 INTRODUCTION

1.1 THE EFFECT OF COVID-19 MEASURES

In late 2019 and early 2020, the world changed as COVID-19 spread around the world. It was declared a pandemic by WHO in March 2020 Ghebreyesus (2020), and since then, it has affected almost all countries worldwide. In attempt to stop the spread of the virus, most countries introduced several measures including lockdowns, curfews, etc. While these lockdowns were proved to have a positive effect on stopping the spread of COVID-19, there were also several negative impacts. These include massive economic consequences, both for businesses and individuals, as shown by Jackson (2021), direct consequences on the mental health of many individuals, shown by WHO (2022), the loss of jobs for many Nations (2021), and more. While these effects have impacted all parts of the world's population, as well as all layers of societies, some vulnerable societal groups were affected differently, most often harder. Investigating this is the main motivation for this project.

The motivation for this thesis came from the desire to articulate some of the issues that the strict COVID-19 measures have had on different socio-economic groups. The thesis *"The impact of COVID-19 lockdowns on the access to economic opportunities in the metropolitan areas of low and middle-income countries"* by Keerd (2021) investigates exactly this issue. This project aims to use principles of data visualization and data journalism to rework the findings of the aforementioned thesis in such a way that it can be angled towards a non-technical audience. This is done by creating a website containing an article alongside the reworked visualizations.

1.2 PROJECT STATEMENT

This paper explores and analyzes the question: "How did the COVID-19 lockdowns impacts differ between different socio-economic groups?" We intend to communicate this to a general, and even non-technical audience with a data journalism piece using the technique "scrollytelling". We will do so by visualizing the findings of a previous study, which used COVID-19 mobility data, using it to investigate the impact of lockdowns on different socio-economic groups, as well as the overall population of different major cities around the world. The study we investigate and make use of is targeted to a scientific audience, though we intend to visualize this in a way that is engaging and understandable for a general audience.

2 RELATED WORKS

2.1 COVID-19

Living in a time where the COVID-19 pandemic is still present and still taking a toll on many countries, the relevance of analyzing and communicating its effects becomes increasingly more important. Sites such as Our World in Data Mathieu et al. (2020), Financial Times Bernard et al. (2022) and others have all devoted themselves during the pandemic to try and visualize coronavirus-related statistics in a clear and understandable manner. However, these sites primarily focus on presenting their stories through dashboards or 'regular' news articles. This sets our website apart from these as we combine statistical findings with an engaging narrative and storytelling layout, as well as some relatively easy-to-understand visualizations.

Other research, such as Ghosh et al. (2020) has focused on how to scientifically analyze the impact of COVID-19 restrictions on the economy and living conditions of different countries, in this case, India. But what separates our project from this paper, is the level of technical understanding needed for being able to comprehend the key findings and purpose of the research. Similar to this paper is the master thesis by Keerd (2021), which studies the impact of COVID-19 lockdowns on access to economic opportunities in the metropolitan areas of low and middle-income countries. This take on COVID-related research studies societal problems through the lens of a Data Scientist by comparing metrics of mobility and economic accessibility to statistically prove if people of lower incomes tend to be affected harder by strict lockdowns during unique circumstances such as the COVID-19 pandemic. Though, as this research is aimed at a scientific audience, our project instead takes an off-set in the research made and instead investigates how to effectively rework the visualizations made, such that they become intuitive for people with different academic backgrounds.

Another noteworthy inspiration is the work done by Mogaji (2020), which investigates the Impact of COVID-19 on transportation in Lagos, Nigeria. This study consciously chose a developing country to analyze, instead of following the trend of mostly reporting the effects of lockdowns found in developed and high-income countries. A similar motivation was the foundation of our own project, as it also

studies the effect of the implemented lockdown strategies made by low-middle-income countries around the world.

2.2 SCROLLYTELLING

Overall, many research papers may be motivated to study similar problems, but due to their academic level and way of thinking, they only reach a like-minded, academic audience. A typical element that hinders the general public from understanding key findings in academic papers and research are the overly complicated visualizations in these documents that have neglected a proper design process. This problem is discussed in the research article by Moere and Purchase (2011) and was the main motivation of the narrative of our project.

Visually pleasant and aesthetic designs are also important factors in both the fields of UI and UX design, as well as web development. With today's large quantities of news articles, research papers, and stories being published on the world wide web, the need for conveying these in a visually engaging and simple way has grown increasingly important. The New York Times created a new type of news article, "Snow Fall", by John Branch (2012), which has since then revolutionized the way of conveying long stories and sensational news articles to the public, a style that has since then been named "Scrollytelling". Our project uses the same principles, though it challenges the limits of scrollytelling by instead trying to convey the story of a scientific study.

Though "Snow Fall" was named the inventor of scrollytelling, the principle of "Narrative Visualization" was studied a year prior by Segel and Heer (2010). This research paper documents and explains the potential of storytelling through data visualizations. Segel and Heer (2010) covers five different fields with different approaches and functions to utilizing data visualizations in their works. Though this is merely a study of how to use data visualizations effectively and not a project with similar intentions to ours, their work created a solid foundation for the product created and the thought process for this thesis.

3 BACKGROUND

There have been many studies and projects investigating COVID-19 measures, including lockdowns, and specifically, how these measures have impacted societies (see section 2.1). One such study is the one this thesis is drawing its findings and objectives from. More often, Data Science is seen as a programming heavy and technical field, dominated by Machine Learning and programming. A sometimes neglected, though not insignificant part of Data Science, is the field of visualizing and communicating data to people from different backgrounds. The demand for BI-specialists and Data Scientists with communicative skills is growing rapidly, as is the need for conveying information in an easily-consumed manner. This thesis explains and accounts for the attempt made to communicate the findings and topics of a scientific research of how the COVID-19 pandemic influenced the mobility of different socioeconomic groups in a handful of low-middle-income cities from around the world. This study is, naturally, a scientific paper, and thus, the results are communicated with a scientific and technical audience in mind. The aim of this thesis is to successfully communicate and discuss the findings and topics presented in the study, in a way that is more angled towards a general audience. This has been done by writing a data journalism piece in the so-called "scrollytelling" format (see section 2.2), and presenting this on a webpage.

4 DATA AND MATERIAL

This project is a little different than a typical Data Science project, as it builds on the findings of a previous study, communicating those results, rather than producing them from scratch. However, some of data provided for this project still needed to be processed into another format, in order to create some of the final visualisations. This data, as well as other data gathered will be described briefly below. For all other data that has been used to calculate the findings communicated in this paper, as well as more thorough explanations of the provided data, see Keerd (2021).

filename	dtype	# lines	data about	description
bog_combined_raw_mobility.csv	csv	539647	mobility	raw mobility data for Bogotá
del_combined_raw_mobility.csv	csv	1191065	mobility	raw mobility data for Delhi
jak_combined_raw_mobility.csv	csv	1403025	mobility	raw mobility data for Jakarta
lag_combined_raw_mobility.csv	csv	275983	mobility	raw mobility data for Lagos
bog_daily_flow.csv	csv	366	mobility	2020 daily flow data for Bogotá
del_daily_flow.csv	csv	366	mobility	2020 daily flow data for Delhi
jak_daily_flow.csv	csv	366	mobility	2020 daily flow data for Jakarta
lag_daily_flow.csv	csv	366	mobility	2020 daily flow data for Lagos
rwi_bog.csv	csv	842	rwi	rwi data for Bogotá
rwi_del.csv	csv	2485	rwi	rwi data for Delhi
rwi_jak.csv	csv	1522	rwi	rwi data for Jakarta
rwi_lag.csv	csv	839	rwi	rwi data for Lagos
lockdown_col.csv	csv	617	lockdown	lockdown data for Colombia
lockdown_idn.csv	csv	617	lockdown	lockdown data for Indonesia
lockdown_ind.csv	csv	617	lockdown	lockdown data for India
lockdown_nig.csv	csv	617	lockdown	lockdown data for Nigeria

Table 1: Overview of data files provided for the project.

4.1 DATA

Initially, the project our thesis is based on explored the stories of five major low-middle-income countries (LMICs) around the world. Namely, Delhi, Bogotá, Lagos, Jakarta & Mexico City, as well as their surrounding metropolitan areas. Though, as stated by Keerd (2021), the data of Mexico City was suspected to have an issue in its mobility dataset. For this reason we decided to remove this city from our project. The suspected issue came from the observation that after loosening the restrictions in Mexico City, the mobility still kept decreasing, which seems counter-intuitive of how the mobility should behave in such a case. The decision of removing Mexico City was also based on some aesthetic choices that were relevant at the time of the

decision. More specifically, we considered a dashboard-like layout, which clashed with the number of cities, as this would require symmetry in order to look visually pleasing. Additionally, we also wanted to incorporate this symmetry on a higher level. In our case, two of the cities, Delhi and Lagos, showed that the mobility of their low-income socioeconomic groups was affected the most, while Bogotá, Jakarta and Mexico City all had the mobility decrease the most for their high-income socioeconomic groups. With these arguments, we wanted to remove one of the cities from the latter groups, to have an equal distribution of how COVID-19 lockdowns influenced the different socioeconomic groups around the world.

4.1.1 Mobility data

For each city, there is a mobility data set. This section will describe and preview that data. The 'mobility data' consists of two files for each city. When focusing on e.g., Delhi, there is `del_combined_raw_mobility.csv`, the raw mobility data, as described in the thesis paper (Keerd, 2021). The data set for Delhi consists of 1,191,066 lines of data - the other cities have similar quantities of data. The raw mobility data is formatted as follows:

date	start_lon	start_lat	end_lon	end_lat	flow
2020-01-01 00:00:01	77.1569824	28.6616708	77.1130371	28.7194957	-14.8811
2020-01-01 00:00:01	77.2009277	28.5073151	77.244873	28.5459253	-14.6987
2020-01-01 00:00:01	77.7282715	28.9504752	77.2229004	28.5845213	-15.3919

Table 2: First 3 lines of mobility data. This is from the Delhi file.

Asides from the raw mobility data, there is also a daily flow data set for each city - this data consists of one line for each day in 2020, with the average flow from that given day.

date	flow_nonlog
2020-01-01	0.0019386882146199818
2020-01-02	0.0020157917898178234
2020-01-03	0.0021368630928598623

Table 3: First 3 lines of daily flow data. This is from the Delhi file.

4.1.2 RWI data

Each of the cities studied also came with a RWI (Relative Wealth Index) file provided. Each file consists of lines denoting the latitude and longitude of 2.4km$\times$2.4km large micro-regions, and the Relative Wealth Index associated with each of those micro-regions. Furthermore, an error column is included - which depicts the uncer-

taincy associated with the data. However, this measure is beyond the scope of this paper.

latitude	longitude	rwi	error
29.0465652	77.7722168	-0.256	0.553
28.7965458	76.7834473	0.596	0.523
28.8157994	77.5085449	0.082	0.521

Table 4: First 3 lines of RWI data. This is from the Delhi file.

4.1.3 Lockdown data

One of the first to investigate the governmental responses to COVID-19 was Hale et al. (2021), who standardized lockdown measures from around the world by creating the Oxford COVID-19 Government Response Tracker (OxCGRT), a dataset containing information about policy measures of different governments from around the world. This enables the possibility of comparing multiple countries and their responses to the COVID-19 pandemic without having to manually account for differences in lockdown severities between countries as the OxCGRT data uses a new metric called the Stringency Index.

This provided lockdown data (OxCGRT) is a very extensive data set and therefore includes too many columns to be previewed. There is one line for each day in the given file. The different columns contain information about almost all imaginable COVID-19 measures, with columns like "International travel controls", "Income support" and more. However, the column that is most relevant for this paper is the so-called "StringencyIndex", the new metric mentioned earlier. It is a calculated combination of many of the different measures - used as a guideline for how strict a country's measures were.

4.2 DATA PROCESSING

As Flourish had been chosen as the visualization tool, the data had to fit the (relatively 'strict') format Flourish uses for the visualizations. Specifically, for the mobility visualizations, the provided data was structured with 'start_lon', 'start_lat', 'end_lon', and 'end_lat' columns. However, the only format accepted in Flourish is two separate `.csv` files - one 'Locations' file consisting of at least three columns; location_id, latitude, and longitude. The second necessary 'Lines' needs to then have at least two columns, a 'from_id' and a 'to_id' column. While this format is

relatively simple, it meant that some processing of the data had to be done. As the provided data didn't contain IDs, it was necessary to create a script that does the following:

- Loops through every 'start_lon'+'start_lat' and 'end_lon'+'end_lat' pair in every line in the `.csv` file
- Checks if these coordinate pairs are already in locations file
- If not, assign an ID to coordination pair and add to locations file
- Remember ID, and add IDs of 'start_lon'+'start_lat' and 'end_lon'+'end_lat' pairs to lines file

This was done using the `pandas` library in `Python`. As there are more than a million lines in multiple of the cities files, this took a considerable amount time (many hours). After this processing, there are now two files for each city, suitable for Flourish.

LocID	lon	lat
0	77.1569824	28.6616708
1	77.1130371	28.7194957
2	77.2009277	28.5073151

Table 5: First 3 lines of transformed Locations data. This is from the Delhi file.

date	sourceID	targetID	flow
2020-01-01 00:00:01	0	1	-14.8811
2020-01-01 00:00:01	2	3	-14.6987
2020-01-01 00:00:01	4	5	-15.3919

Table 6: First 3 lines of transformed Lines data. This is from the Delhi file.

After this processing, a simple script to limit dates of the Lines file was created. This was simply to make it possible to limit the dates to e.g., a 'before-lockdown' and 'during-lockdown' period. One could argue that limiting these dates before running the relatively slow script on millions of lines of data would be clever. While this is certainly true, a decision was made to transform all of the data at once, and then perform the simple calculations and date limitations on the completed, transformed data set.

It was also needed to combine two data sets - the lockdown data and the daily flow. This was a necessity to create a dual y-axis bar chart in Flourish. This was also done using `pandas` with a simple merge:

```
combined_data = pd.merge(lockdown_data, daily_flow_data,how = 'inner', on = 'date')
```

Lastly, on the same data, an extra column was added - a 7 day rolling average. This is common practice when looking at COVID-19-related data, as well as with

any data that can fluctuate in correlation with weekends. Again, this was done using a simple pandas command:

```
combined_data['rolling_avg'] = combined_data.rolling(window=7)['flow_nonlog'].mean()
```

5 DEVELOPMENT

The development of the website was an iterative process, as is often the case in programming. This required numerous iterations and a significant amount of time, as well as the use of multiple libraries, before the final version of the webpage, article, and visualizations were completed. In the end, it was decided to use Scrollama to implement scrollytelling.

Even when Scrollama had been chosen, there were many issues along the way. As it was very much a learning experience as well, this is also reflected in the amount of iterations necessary to reach a satisfactory product.

While there were several other things that ideally could have been a part of the project, it has been important to limit the scope of the project, and focus on what's feasible in the given time frame. This is also why it was a relatively quick process to choose the scrollytelling library, the visualization tool, etc. As en example, the website was built only containing visualizations and text regarding one city, in order to ensure the desired functionality before expanding the scope to multiple cities.

5.1 VISUALIZATIONS

Visualizing data should be granted a long and detailed design process - as should the process of creating the supporting text for these visual elements. This vital task is often overlooked in scientific and academic domains and leaves many without the possibility of gaining new and exciting knowledge from said domains. For this reason, this thesis explores how to convey academic information in a simplistic, yet informative, manner.

As mentioned in the previous sections, we were presented with a master thesis by Keerd (2021), which analyzes COVID-19 lockdowns effects on mobility, and decided to dissect the project such that we could extract the elements that we found the most neglected and most scientific. We picked the visualizations from the aforementioned thesis based on three qualities: Importance in the presented thesis, relevance to the scope and goal of our project, and how neglected and misinformative they

seemed. As can be found in figures A2, A4, A3, A5, A6, A7 in the appendix, the visualizations all had different color palettes and visualization styles. If they were presented by themselves, one would most likely not even assume that they came from the same research and are all related to each other.

In order to rework the visualizations picked from the master thesis, we looked into "The Truthful Art" by Cairo (2016) and his five qualities of great visualizations. We wished to create informative and captivating visualizations that would successfully support the narrative of our story. Luckily, "What makes a visualization successful?" is the main question asked throughout the book.
Alberto Cairo's five qualities of what makes a visualization successful are as follows:

1. Is it truthful
2. Is it functional
3. Is it beautiful
4. Is it insightful
5. Is it enlighting

Following the third quality - creating a beautiful visualization, we planned to tell the stories of the cities simultaneously by using small multiples. We also hoped that using this technique, it would make our final product more functional, insightful, and enlightening - another three of the qualities listed above. Telling the stories of the four cities simultaneously, meant that our initial sketches played with the concept of a window-like dashboard with a tile for each city, which can be seen illustrated in figure A1. After manually playing with the arrangement of the visualizations, we found that clustering four cities in such a layout ended up not looking simplistic enough. Instead, it seemed rather cluttered and visually heavy for the eyes of the reader. More technically speaking, the data density became too high this way. Having four visualizations grouped together would mean having to either scale them up to a bigger size, such that the x- and y-labels were readable or scale them down to not fill the whole page, thus losing their details. Neither of these solutions were beautiful nor sufficient.

Instead, this realization made us look into the possibilities of creating an individual story for each city, which could then be put together by a website with multiple webpages, as explained further in section 5.3. Following this approach meant that we had to generate a set of similar visualizations for each city and its story and connect them all through articles that all followed the same storytelling and visual layout. In order to create the visualizations we settled on using Flourish, an online visualization tool. We considered other tools such as Tableau, but since Flourish has a "stories" feature that simplifies the process of embedding the created

visualizations with interactive elements and animations directly to a webpage, we discarded the idea of using Tableau.

To ensure that our redesigned visualizations all followed the same style, a shared color palette was used. Due to the overall theme of the stories being "Socioeconomic groups", of which one measure is wealth, the color green was considered but was quickly discarded due to its intensity. Similarly, red was discarded based on the same visually intense appearance. Blue seemed to be closer to the visual effect that we were looking for, but as we wanted our visualizations to have a dark and more serious appearance, the blue began to blend in with the dark grey background color.

We settled on a teal color as it combined the calm appearance of the blue color while adding the association to wealth from the green color. The finalized color palette used a combination of teal and gray hues as can be seen in figure 1. The dark grays were used for the backgrounds of our visualizations, the reason being that the theme of the stories is relatively serious and explains the negative effects of lockdowns on the mobility of different socioeconomic groups. Another reason for utilising darker colors for our visualization backgrounds was to ensure a contrast between our visualizations and the background color of our website which is a very light gray.

Figure 1: Our final color palette consisting of teal and gray hues

The design process of the individual visualizations became a more tedious process than expected. Due to the simplicity of the visualization tool Flourish, many smaller details could not be tweaked or adjusted - in other words, Flourish lacked the ability to customize visualizations beyond the predefined options. These options included choosing color layouts for different visualizations overlayed on a map. Three of our chosen visualizations used a map as a background, namely our empty map of the current city (A6), our RWI visualization, which visualized the relative wealth of each micro-region (A2) and the mobility visualizations showing mobility before and during the lockdowns (A4, A3). Unfortunately, Flourish does not offer the choice of selecting the same map styles for each of these visualizations as they use different visualization types. This means that the backgrounds vary slightly in style, though their overall color scheme remained the same as can be

seen presented in figures A9, A8 & A10.

The visualizations that required the most redesigning were the 'lollipop' chart found in figure A5, and the visualization depicting the effect of the Stringency Index on the daily mobility A7, the so-called "SIDF graph" (Stringency Index & Daily Flow). Using a lollipop chart is an effective way to show differences between multiple bars, but as it is the center of the circle on the top of the bar that marks the value of the lollipop chart, they can be visually harder to read than a bar chart. Additionally, due to their slim body, having multiple lollipop charts of different colors can be harder to distinguish and categorize. The visualization made by Keerd (2021), visualizes three separate plots, and their colors have no meaning, other than to distinguish the three plots, something their borders and positioning already do. This created unnecessary visual clutter as the separation of the three plots already visually separates them by the Gestalt principle of proximity. Instead, we wanted to redesign them in a way that the reader could understand the purpose of each bar based on their color before reading their labels. As we had already decided on a color scheme, and, at this point in time already created the "heatmap"-style inspired visualization that depicts the Relative Wealth Index, we decided to use the same colors as the ones used to depict low-, middle- and high-income. These three hues of teal can be seen in figure A11.

Applying the same colors to the visualization made it significantly easier to understand. However, we still found that having the bars in three separate plots was a waste of space, and could lead to more confusion. Instead, we decided to merge the plot visualizing the mobility of the different socioeconomic groups prior to the COVID-19 lockdowns with the plot visualizing the same but during the lockdown period. This led to a visualization, as seen in figure 2 that is arguably beautiful due to its stylistic simplicity and color choices, is enlightening as it visually shows how different groups were affected by the pandemic, the redesign has made it more insightful by providing a simple visual layout for one single plot without having to navigate between multiple. This visualization was already truthful from the beginning and finally, it accurately depicts the data provided and is now built in a way that leads the reader to meaningful insights based on it. This leads us to the conclusion that the redesign of the first half of the previous lollipop chart has resulted in a successful visualization. Though, due to the simplicity of Flourish, it would not allow for visualizing the y-axis labels. As an alternative, it instead lets the reader hover over every single bar with their mouse, showing the level of daily flow. This is not sufficient in the long run, and would therefore need to be recreated using an-

other visualization tool in the long run, though this is out of the scope of this project.

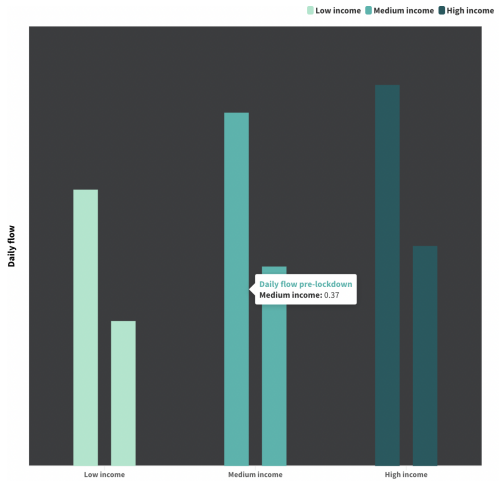

Figure 2: The redesigned version of a plot depicting the mobility before a during lockdown for each socioeconomic group

The same methods of redesign were applied to the last lollipop chart, visualizing the percentage-wise difference for each socioeconomic group. From the beginning of the process, we found this plot confusing to look at as it depicts the negative percentages as well as having upside-down lollipop charts. For the general public, one should avoid making things more technical and confusing than needed, which we believe this visualization did. The general reader, even without a technical background, is still able to grasp the fact that the change in mobility was a decrease. Thus, we decided that could simply visualize the positive percentage and do so with bar charts pointing upwards. Following the same redesign steps with this visualization, we ended up with a similar visualization as seen in figure 3, which clearly and aesthetically communicates how the low-income socioeconomic group of Delhi was affected the most.

The SIDF (renamed by us to Lockdown Severity / Daily Flow) visualization also needed a serious redesign as this was the most information-cluttered visualization (see figure A7). To redesign this visualization, the two main lines visualized in it were given colors that matched our chosen color palette. The SI (Stringency Index) line graph was additionally renamed to Lockdown Severity as our user testing revealed that the interviewees could not understand its original name. As such, it was given a less scientific and more understandable name.

The newly renamed Lockdown Severity graph was changed in appearance as it was given a light gray color in order to be more dimmed in comparison to the Daily Flow graph, which we gave a teal color similar to the ones used in the other visual-

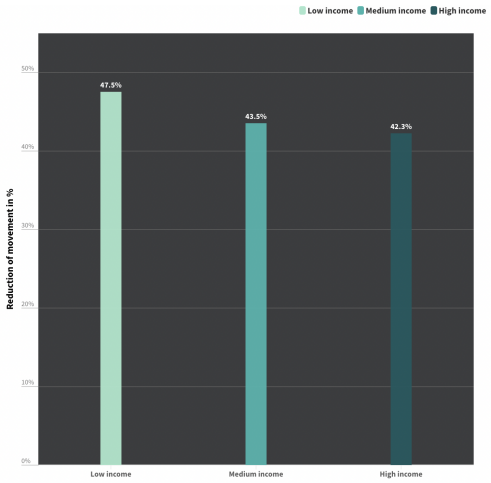

Figure 3: The redesigned version of a plot depicting the mobility before a during lockdown for each socioeconomic group

izations. The line graph of Daily Flow still seemed too insignificant and vague. To change this appearance it was changed to an area graph instead, as this gives it a higher data-ink ratio and makes it the center of the visualization. In practice, area graphs are great at visualizing the change of a variable over time or over the span of another variable. As this graph acts as a timeline, this design choice seemed very appropriate.

Both before and during COVID-19, there has been a trend of commuting less during weekends due to the absence of the need for commuting to work or educational institutions. After nationwide lockdowns were enforced around the globe, even though the overall mobility decreased, a pattern of more travel happening during the weekdays can still be seen. This pattern is also present in the original SIDF graph by Keerd (2021), which makes it harder to follow the trend of the actual mobility as the constantly spiking graph creates noise for the eyes of its readers. To combat this issue, we smoothed the data using a 7-day rolling average, which was described in section 4.2. Lastly, a small detail in the design of this visualization was changing the placement of the ticks on the x-axis and the title of the y-axis. These were moved from the outside of the visualization's border to instead be visualized on the inside. This change made it easier to follow the dates when different SI measures were implemented and when mobility changed drastically.

These steps created the new visualization used for this project and which can be seen in the appendix in figure A12. The redesigned Lockdown Severity and Daily Flow visualization allow the reader to quickly find a correlation between colors from the other visualizations and be presented with a simple plot that shows that the mobility in Delhi dropped drastically as the Lockdown Severity went up. We decided

to not visually implement the y-axis for either of the two graphs, as their measures will not be understood by the average reader without a technical background. It was even considered if keeping the y-axis visual would confuse some of our potential readers, as the Daily Flow measure itself does not explain anything and thus only clutters the visualization. As such, the visualizations had been redesigned and tweaked to fit our story layout and not contain unnecessary information. By having the visualizations with a common theme and similar visual appearance ready for use, the thought process of how to incorporate them in storytelling pieces began.

5.2 ARTICLES

The created visualizations acted as the skeleton for the articles created for this thesis. The process of conveying academic writing and research into a storytelling article was done by following the concepts stated in the paper "Narrative Visualization: Telling Stories with Data" by Segel and Heer (2010). Following the concept of the "martini glass structure", we intend to tell a story layout that engages the reader throughout the whole article. The martini glass structure is a narrative pattern that follows a tight narrative path, the stem of the glass, and then opens up for free exploration, which, metaphorically, is represented by the body of the martini glass. We implemented this structure by creating visualizations that allow for personal interpretation. Though we support the visualizations with text, the stories presented do not analyze the visualizations in all detail. This allows the audience to dive into the visualizations or text bubbles as they prefer and find engaging, rather than controlling them throughout the storytelling process.

Our website structure was made based on our idea to have a single story for each city. Being able to click from one story to another allows for much more freedom than if these articles had all been in the same document or on the same webpage. Narration-wise, we aimed at having a maximum of two text bubbles for each visualization. This was harder than expected, as our articles needed to both tell the story of a country and its people as well as educate and explain the visualizations to a general audience. To successfully convey this objective, we looked into the visual and narrative appearance of other scrollytelling articles on the web. The general tendency in such articles is the short length of the sentences, which ensures that the readers do not lose focus. Additionally, the contrast between the introduction and the body of the article lets the reader dive into a concentrated and informative environment (e.g. passing through the stem of the martini glass). Typically,

the introduction states the general problem that is about to be discussed and why this problem has been investigated. During the introduction, multiple questions are usually asked, in order to plant a seed of curiosity in the mind of the reader.

Once the story goes on to the martini glass stem, the questions asked in the introduction seem to be almost answered after each paragraph or chapter. This keeps the reader on their toes and gives them the knowledge essential to follow along with the next paragraph or visualization. This subconsciously adds to the toolbox of the reader, who after having read such an article might feel enlightened, though not bombarded with new heavy information. These questions invite the reader to be curious and critical which can nudge them in the direction of studying the presented visualizations on a closer level. This motivation is also something we hope to achieve through our stories.

To introduce the theme of our articles, our introduction presents the viewer with an engaging title, "Traveling through the pandemic", as well as some facts related to the theme of the piece, which may to many seem like common knowledge. We decided to add a disclaimer to our introduction to point out that some of our predefined ideas may contradict the reality in other countries. And lastly, before moving on to the body of our stories, we present the reader with a choice of following whichever city they would like, motivating them with a free choice to hopefully be more interested in the content that they are about to be presented with. In case a reader does not have a preference or any knowledge of the pandemic, we have decided on a predetermined order of our cities.

The stories for each of the cities were written using facts and research from various sources on the internet. As mentioned in section 2, multiple news papers and acknowledged websites devoted themselves during the pandemic to research and produce quality statistics about the pandemic. This easy access to proper data and statistics is sadly overshadowed by a lack of said data for LMICs. As this research specifically looks into the handling of the pandemic in low-middle-income countries, looking for relevant research or data was unfortunately often met with uncertified sources, lack of said data or aggregated statistics that neglect the importance of transparency when talking about the economic situation of a country or the number of tested or dead people. Especially Lagos was hard to research, and the development of Lagos' article took a disproportionately longer time than for cities such as Delhi or Bogotá. This has resulted in large quantities of unreported numbers which as a result mean a high uncertainty in the statistics and facts presented

in the four stories, though especially in the case of Lagos.

In order to visually shorten the reading time required to be able to understand the scientific visualizations placed next to the text bubbles, it was considered to create short titles for each text bubble. Instead, we created footers below our visualizations to simply explain them, thus saving people from having to read all of the presented text before moving on to the visualization. This approach seemed to contradict our wish of producing engaging storytelling pieces, as the footers would in come cases almost be able to replace the text completely. Though as we came to learn from the user testing, the less technical interviewees saw a lot of potential in the presented text. To improve the text sections and make them more important while not visually heavy, we decided to highlight important dates, numbers, names, places etc. with bold. Visually this instantly raised the overall appearance of our scrollytelling piece.

Combining all of these aforementioned considerations, we created four stories, ready to be implemented on our website and ready to become a part of our take on how scrollytelling in relation to academic research could look like.

5.3 THE WEBPAGE

5.3.1 Initial ideas

Initially, we started out drawing what we had in mind for the site, as seen in figure 4. Essentially, we wanted our visualizations to evolve with the story, as described in section 5.1. It was important to keep this in mind when looking for the right tools.

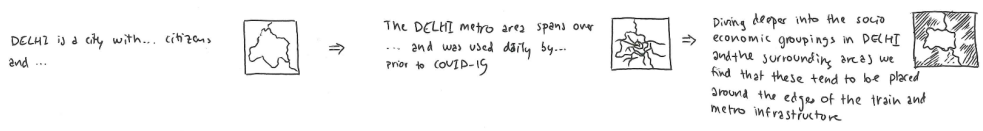

Figure 4: Drawing/storyboard of how we wanted to tell the story

Specifically, we wanted an engaging format, where the text content would have a 'normal' scroll behaviour, while the visualizations remained in the same spot, replacing each other, as different text boxes triggered the next visualization. This can be seen in figure 5.

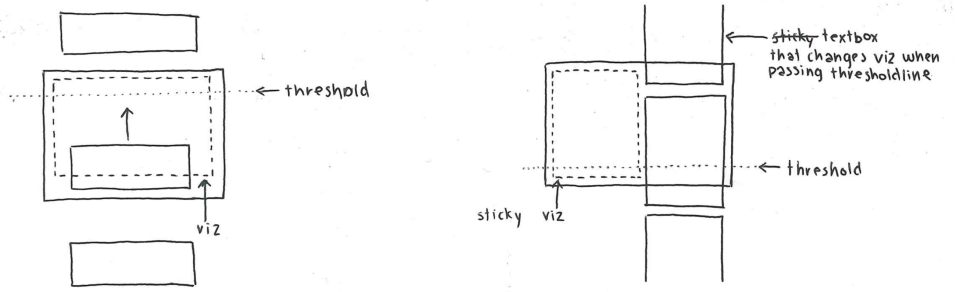

Figure 5: The two initial scrollytelling ideas

As seen above, we had two different approaches in mind when it came to building the scrollytelling site. The approach drawn on the left is a popular approach when working with full-screen visualizations. The issue with this, however, can occur when there are larger text boxes. In order to see the whole text box, it can potentially cover vital parts of the illustration. As we had relatively long paragraphs to be viewed with the visualizations, we decided that this could become an issue. Thus, we decided to go with the second approach, where the visualization stays fixed to one side, not intruded by the text boxes towards the other side. Initially, the idea was for the reader to dive into the visualizations themselves and afterward be met with supporting text to be guided through the stories for each city. After having developed the desired visualizations, it was then clear that even though they had been visually and aesthetically improved, they were still too scientific and information-heavy to stand alone. Our original idea of placing the visualizations to the left and the supporting text boxes to the right was thus discarded. Instead, the idea of presenting the reader with a short and informative paragraph which would give them the tools and background needed to fully comprehend the visualizations became our new aim.

To guide the reader in the most natural way, we considered the visual hierarchy of our website as well as the Gestalt principles explained on UserTesting.com (UserTesting.com), both elements of great importance in the fields of Data Visualization and Web Design. Visual Hierarchy is the principle of arranging the visual elements in such a way that they are natural to follow in an intended order, usually of importance. By this logic, we expect our audience to follow the order of left to right, as this is the reading direction in English as well as in most other languages. This means that simply by placing our visualizations to the right of our text, we will naturally guide our readers in such a way, that they will read about what is visualized to the right, before taking a look at the visualization themselves. In terms of the Gestalt principles, the layout of our scrollytelling piece follows the law of continuity, which lets the users naturally see a pattern in the text boxes, as they all

follow the same path, making them related to one another. By implementing these two simple, yet effective principles, we ensure a better user experience as well as make it possible for us to show them visualizations that otherwise might have been deemed too technical, had they been standing by themselves without their supporting text.

Due to a lack of experience with building websites as well as with the necessary languages (HTML, CSS, and JavaScript), it was not feasible to build it from scratch. Therefore, it was decided to search for a 'template' that could be used with the visualization tools we preferred, while still being relatively accessible in regards to editing it to suit the requirements and needs. While the aim was to utilize someone else's template, it was, naturally, important to only use the scrollytelling functionality of the template, and not the layout, style, etc.

When searching for the right tool to help us build the scrollytelling site, there were a few factors considered when making the decision. First and foremost, functionality. The main objective of using a pre-made template was to obtain a solution where we would not need to program it from the bottom. There were a handful of good solutions that all adhered to this requirement, so to choose between these, we had to look at our next requirement – ease of use. As mentioned, none of us have HTML/CSS/JS experience, and with a time constraint, finding something that was relatively "learnable" was of great importance. Lastly, but importantly, we considered the visual appeal. While the other two factors are, functionally, very critical for the completion of our project, this last requirement is also very important, as it is an essential element of good communication.

As we chose Flourish as our visualization tool relatively early on, we decided to investigate how scrollytelling could be implemented with Flourish. Inspired by an article created by Riley et al. (2019) describing how the British news agency Sky implemented scrollytelling with Flourish, we began investigating Scrollama rusell-samora (2017) – a .js library designed to ease the implementation process of scrollytelling. Early on, Scrollama seemed like a relatively good solution, as there was a lot of documentation online, also describing the use of Scrollama with Flourish. The creator of Scrollama has made tutorials, example sites, and more. Furthermore, other users of Scrollama have shared source code for sites implementing Flourish - specifically, one template sitiaish (2021) implemented a similar approach to what was described in figure 5.

We also considered using Svelte, by Harris (2016). Svelte is a front-end compiler, and not a library per se, but similarly to Scrollama, it would enable handling of the different scrolling "triggers" that decide when to change visualization when scrolling. Additionally, it would enable another layer of customization regarding the visualizations. Unfortunately for this project, the visualizations had to be coded from the bottom, with HTML/JS/CSS. Though this would have given us more artistic freedom and control over every single detail, a shortcoming when using Flourish, the level of experience needed to follow through with Svelte made us stick with Flourish as our visualization tool. This also resulted in us sticking with Scrollama as well.

5.3.2 Code

Firstly, the template was loaded to make sure everything worked properly. Firstly, node.js had to be installed to install the Scrollama.js library. Then, the `index.html` file was tested on both our computers before committing it to our GitHub repo. Using a GitHub repo was an obvious choice, partly for its collaboration features, naturally, but also due to another, very useful feature – GitHub Pages. GitHub Pages hosts our website, which allows us to have a 'normal', functioning website that anyone can access with a link, instead of downloading our repo, having to open a `.html` file in their browser, etc. This is extremely simple – once an `.html` file is committed, it's as simple as turning on a switch in the repo settings. This enables us to test exactly how our code functions as a 'real' website, test it on different devices, etc. In order to make the code more understandable and reusable for all cities' `.html` files, we gathered the JavaScript functions to a separate `script.js` file and all styling to `style.css`.

After making sure the basics functioned properly, we started working on personalizing our article. Firstly, we created an initial, development edition of our Flourish story, containing semi-finished visualizations for Delhi. Then, the Flourish story was embedded in the code and tested to confirm the functionality. Here, there were no real issues, as Scrollama and this specific template handle all Flourish stories the same way.

As Scrollama embeds a Flourish 'story', Flourish handles the build of visualizations, including animations, etc. between them. This makes it easier to work with, but also severely limits the customization options. Scrollama works with Flourish by embedding a so-called "story". This is Flourish's presentation tool – in some

way, it resembles a PowerPoint presentation, just more focused on data visualization aspects, with things like tooltips, animations, etc. To make sure that each city's `.html` incorporated the correct Flourish story ID, code was added to dynamically change the story ID within the embed link.

```
// set the default story ID (Delhi)
let storyId = 1724565;

// Check which HTML file is being loaded and set the story ID accordingly
if (window.location.href.indexOf("bogota.html") >= 0) {
    storyId = 1762941;
} else if (window.location.href.indexOf("jakarta.html") >= 0) {
    storyId = 1764236;
} else if (window.location.href.indexOf("lagos.html") >= 0) {
    storyId = 1764230;
}
```

Code snippet 5.3.1: JavaScript code to update Flourish story ID

Every Flourish story has "slides". The user's position on the page is monitored, i.e., when the user scrolls, Scrollama then triggers a function that changes the embed URL.

```
// Update graphic based on step
const linkHead = 'https://flo.uri.sh/story/${storyId}/embed#slide-';
const slide = response.index

d3.select('.scrolly__chart iframe')
    .attr('src', linkHead + slide);
}
```

Code snippet 5.3.2: JavaScript code to update Flourish story ID

Essentially, Scrollama changes the slide number in the embedded URL every time the user reaches a certain position on the page. Also a part of the template, and, naturally, important for scrollytelling, the visualization must stay in the same position on the user's screen, even when scrolling. However, the text boxes must scroll "naturally". Fixing an element to a specific position is not complicated, but having a template handling this, as well as the text boxes, while still triggering the different steps to change visualizations, was essential for us.

Then, we started adding functionality to our site. The objective was to implement a solution where the user can switch between cities relatively easily. Initially, a simple drop-down menu was created. This is quite simple to code – however, it's not necessarily a very elegant solution for the end user. The user must click to see

the other options, forcing an extra (perhaps unnecessary) click. Hypothetically, if the user should choose between all countries in the world, a dropdown would be more suitable, but as the user must only choose between four cities, being presented with all options is not too much information. Additionally, it saves a click, and thus makes interaction more accessible and "tempting".

Having decided that switching between all four cities must be accessible simultaneously, we again had to weigh our options. As with the decision regarding the scrollytelling approach, similar factors played a part – as they will in all technical considerations and decisions. Ideally, we would have liked to implement either a global navigation bar or a legend for each visualization that would then trigger a JavsScript function to dynamically change the content based on the user's choice – without refreshing the page. While this is technically possible, it was not feasible within the scope of this project. Therefore, we ended up going with the approach that was most balanced when it comes to user-friendliness, feasibility, and functionality. This ended up being a global navigation bar – a trusted approach seen on basically any website.

```
19  <div class="sticky_nav">
20          <a class="active" href="javascript:void(0)">Delhi</a>
21          <a href="bogota.html">Bogota</a>
22          <a href="lagos.html">Lagos</a>
23          <a href="jakarta.html">Jakarta</a>
24        </div>
```

Code snippet 5.3.3: HTML for navigation bar

Technically, this was not incredibly hard to do. The navigation bar, can have its `position` attribute set to "sticky", meaning that it will always stick to the top of the screen, no matter the scroll position of the user. However, we wanted it to only be visible while looking at the city-specific visualizations, and not when reading the intro to the article as well as the other parts that apply to all cities. Thus, a solution was implemented, where the navigation bar is fixed to a certain position (using "fixed" attribute) until the user scrolls to the point where the navigation bar is at the top of the screen – at this point, the bar sticks to the top of the interface. Here, we ran into an issue where the navigation bar was hidden below the scrollytelling content. This was solved using the `z-position` attribute – this is value that can be set for all elements in the code. The higher the number (in relation to the other elements' `z-position`) – e.g. if the `z-position` of the scrollytelling text boxes were set to 1, and the `z-position` of the navigation bar was not set (and then defaulting to 0), the scrollytelling text boxes would cover the navigation bar. Thus, the `z-position`

had to be set to a higher value to be brought to the front.

```
.sticky_nav {
  position: -webkit-sticky;
  position: sticky;
  top: 0;
  background-color: #333;
  display: flex;
  justify-content: center;
  padding: 5px;
  font-size: 20px;
  z-index: 90;
  height: 45px;
}
```

Code snippet 5.3.4: CSS defining the class used for navigation bar

As mentioned, the ideal solution with regards to changing content based on the user's city choice would have been to dynamically change only the content relating to the city. As this was not feasible, we decided to go with the simple solution – an `.html` file for each city. The advantages of this are mainly focused on simplicity. It's a simple as duplicating the `.html` file and changing the text in the corresponding text boxes. Then, in the navigation bar, an `href` can be added linking to `.html` file – GitHub automatically generates a new webpage for each `.html` file. There are a few disadvantages related to this approach. Firstly, and most importantly, when changing between cities, the users scroll position (how far down the user has scrolled) is not remembered across different pages. This is highly critical to the user experience of the site, and hence, fixing this was of upmost importance. It can be solved with a relatively simple JavaScript function, which stores the user's scroll position (Y position) in local storage - a 'web storage' option supported by modern browsers to store various pieces of information locally on the user's device.

```
// Save the user's scroll position when they scroll the page
window.addEventListener('scroll', function() {
    var scrollPosition = window.scrollY;
    localStorage.setItem('scrollPosition', scrollPosition);
});

// Restore the user's scroll position when the page loads
var savedScrollPosition = localStorage.getItem('scrollPosition');
if (savedScrollPosition) {
  window.scrollTo(0, savedScrollPosition);
}
```

Code snippet 5.3.5: JavaScript function for storing the user's scroll position

As seen above, this function stores the user's scroll position in local storage, and when a page loads, it retrieves this stored information, and scrolls to the position.

A second disadvantage of having multiple `.html` files relates to the fact that every time we change anything regarding functionality or design, we must do it in all `.html` files. While this is true, there are some parts of the four files that are completely identical - the header, introduction, conclusion and footer. These were all written in seperate `.html` files, e.g.,`common_header.html`. Then, in each city's `.html` file, it could be referenced `<div>`, like this:

```
48     <div id="common_intro"></div>
```

This only works due to a JavaScript function, which fetches the content from the `.html` file in question, and inserts it into the correct place in the code, where it is referenced.

```
49    // Fetch the common content from the "common_header.html" file
50  fetch("common_header.html")
51    .then(response => response.text())
52    .then(html => {
53      // Insert the common content into the "common_header" element
54      document.getElementById("common_header").innerHTML = html;
55    });
```

Code snippet 5.3.6: JavaScript function for common HTML content

Additionally, a link to our GitHub repo was added – in the form of the GitHub cat resting in the top right corner, waving when hovering the mouse over the logo. The code for this was borrowed from the creator of Scrollama rusellsamora (2017) and was only minimally tinkered with to fit our use.

In addition to adding these usability features, many design changes were considered. This included things like styling the headline, text boxes, colors, navigation bar, etc.

5.4 USER TESTING

User testing is an essential part of the design and development process for any digital product, as it allows us to gather feedback from real users and identify areas for improvement. For our scrollytelling site, user testing was particularly important given the interactive and data-driven nature of the content.

To gather feedback on the site, we conducted user testing with three individuals who represent different levels of technical knowledge. Specifically, we interviewed a male student in his 20s with high technical proficiency, a female in her 30s with medium technical proficiency, and a female in her 50s with limited technical proficiency.

During the interviews, each participant was asked to navigate and explore the site, and provide their thoughts and feedback on its design, functionality, and content. The results of the interviews provided valuable insights into the user experience, and identified areas for improvement.

For example, the male student in his 20s appreciated the visual design of the site and found the scrollytelling format to be engaging and effective for presenting data journalism. However, he also pointed out some technical issues with the site, such as slow loading times and difficulties with the navigation.

The female in her 30s with medium technical proficiency also enjoyed the scrollytelling format but noted that some of the content was difficult to understand due to a lack of clear explanations and context. She suggested adding more explanatory text and visual aids to help readers better understand the data and its implications.

The female in her 50s with limited technical proficiency struggled with the scrollytelling format and found it difficult to follow the story. She suggested simplifying the navigation and making the content more accessible for readers with less technical knowledge.

Additionally, there were a few things that were mentioned by all participants. One of the main issues mentioned by all participants was regarding the color choice for the RWI map visualization. Specifically, the dark dots representing the areas with a relatively high RWI value were too dark, and could be mistaken for the background, thus confusing the reader. There are also areas with no data - these areas are simply represented with no dot. This could lead to confusion in whether the specific area had no data (e.g. because of a body of water, which in turn was also almost impossible to distinguish from land) or was an area with a high RWI index.

Other things that were mentioned by the interviewees, and then changed to reflect this feedback, include:

- Add figure text to avoid the reader having to gaze over a whole paragraph for a simple explanation

Figure 6: Before user testing

Figure 7: After user testing

Figure 8: RWI visualizations' background color changed to reflect user testing

- Remember users' scroll position when changing between cities (see code snippet 5.3.2)
- Alter visualization 'trigger point'. Visualization initially changed as soon as the user scrolled from one text box to another

Overall, the user testing and interviews provided valuable insights into the strengths and weaknesses of our scrollytelling site, and will help guide further development and improvements in an iterative process.

6 FINAL PRODUCT

In this section, we present the final version of our scrollytelling site, which explores the impact of COVID-19 lockdowns on different socioeconomic groups in four different cities in low-medium-income countries. Through a series of interactive stories and data visualizations, our site aims to provide a nuanced and engaging look at the results presented by Keerd (2021). Included below are several screenshots of the site, along with brief descriptions of its key features and functions. In developing the site, we received valuable feedback from user testing, as mentioned in section 5.4, which we have incorporated into the final version to enhance its usability and impact.

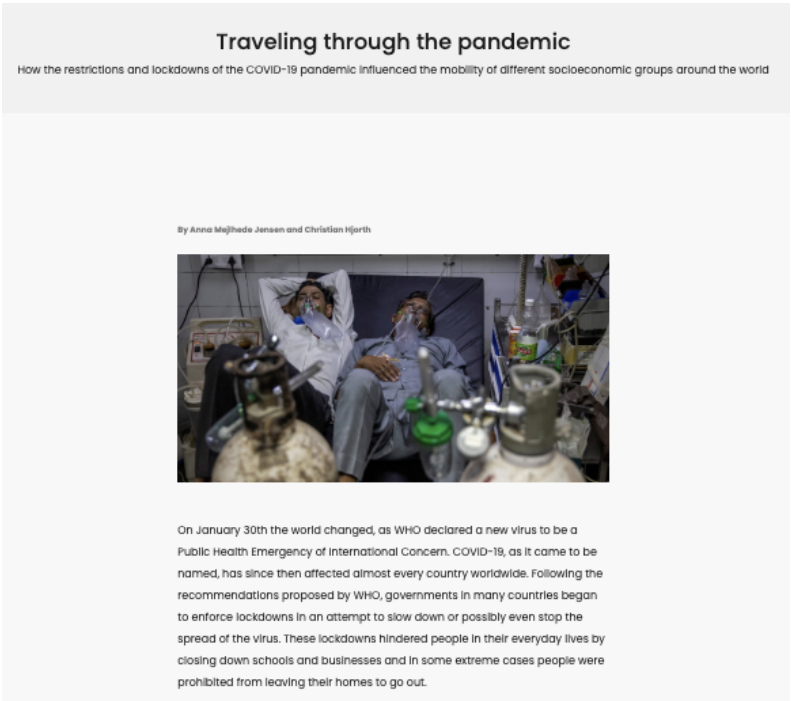

Figure 9: Text box and visualization side-by-side

As shown above, the reader is first met with the headline and a subheading. The headline is attention-grabbing and highlights the main topic of the article, as well as being a wordplay, as "traveling" both refers to transportation specifically, and, more broadly, how society moved through the pandemic. The subheading further introduces the topic and provides more detail about the focus of the article. These two elements, as well as the GitHub icon linking to our repository (not shown in

figure 9), are all wrapped in a header, a vertical bar colored differently than the background to make it stand out more. Below the header, the article starts with a picture. The picture grabs readers' attention and provides a visual representation of what the article is about. Then, the article starts with a general introduction, shared across the four cities' sites.

When scrolling through and past the introduction, the user is met with the navigation bar, where they can navigate to the article corresponding to the city they wish to follow.

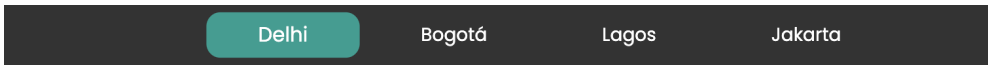

Figure 10: Navigation bar

The navigation bar is relatively simple - a click on the city redirects the user to the chosen city. When clicking the city that has already been chosen, the button shrinks a little in a quick animation, to signal that the city has already been selected.

Scrolling further down, the user is met with the actual scrollytelling, consisting of text box to the left with a 'normal' scrolling behaviour, and visualizations on the right, seemingly unaffected by scrolling. However, the shift to the next visualization is triggered when the text box reaches a certain y-position.

Figure 11: Header and introduction to the article

The text box contains the text associated with the visualization, with some words in bold, to signify extra importance and draw the reader's eyes to those words or phrases. The visualizations also include the option for the user to interact with them. In this case, the user can hover the mouse to see a specific value. In the bottom right, a small arrow can be seen - this is a back-to-top button. As there is a fair amount of content to scroll through, the user can click this button, and instantly navigate to the top of the site. As mentioned briefly, the visualizations offer the user

the option to interact with them (to the extent that Flourish allows).

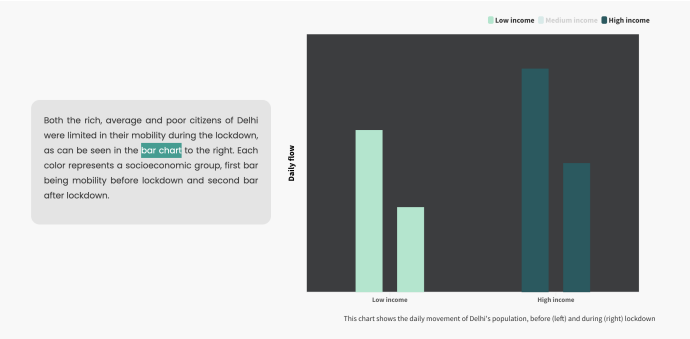

Figure 12: User interaction + text selection color

As seen above, the user can select and de-select value by clicking the legend, to simplify the visualization, e.g., for comparing two specific groups. Furthermore, as seen in the text box, the color of the highlighted text has been changed to reflect the general color palette.

7 DISCUSSION AND LIMITATIONS

The aim of this project was to communicate the results of a relatively technical study by Keerd (2021), which investigated the socio-economic effects of COVID-19 lockdowns. The results of the original study were published in a traditional scientific format, but we wanted to make them more accessible and understandable to the general public. To achieve this goal, we used data visualization techniques to create a series of interactive graphs and charts that allowed the user to explore the data in a more intuitive and engaging way. The visualizations were presented in an article format, with accompanying text and explanations that provided more context and detail. We also used various techniques to engage the user and catch their attention, reflected in the choice of color palette, styling of the website, image choices, etc.

Overall, we believe that we succeeded in our goal of making the study results more understandable to a broader audience. The visualizations were more accessible and engaging than the traditional academic paper, and they allowed the user to explore the data in a more interactive and intuitive way. The article format also provided a convenient and engaging way to explain the results and provide more context. Specifically, the visualizations were also less cluttered, making them easier to understand in less time, which is essential for a good visualization. Additionally, we also made specific color palette choices to enhance the beauty of the visualizations, again making it more interesting for the reader, as well as making it more accessible to those affected by colorblindness.

We do realize that although we believe we succeeded in the aim of this project, the audience is still limited to a certain level of technical proficiency. Through user testing, we learned that some of the visualizations were still relatively complex, and thus, hard to understand. While still more understandable than the original paper, some visualizations still required additional explanations. This led us to add figure texts to most visualizations, as an extra, explanatory layer. In general, this can be represented by the term 'data literacy' - the reader's ability to read and understand data as information. A concrete example from the visualizations is the dual y-axis line graph containing lockdown severity and daily flow. Here, it was decided to calculate a 7-day rolling average of daily flow, in order to 'remove' the spikes that occur on weekends. By not providing a detailed description of what a rolling aver-

age is, we assume that the reader has a certain level of data literacy.

While working on the visualizations, we also learned that there were some limitations and shortcomings to our approach. One of the main challenges faced was the limited customization offered by Flourish, in comparison to coding them from the bottom. While choosing a no-code visualization tool like Flourish has its advantages, such as simplicity, the ability to easily embed it in a website, etc., it certainly also has its limitations. In contrast to coding the visualizations from the bottom, one can only customize exactly what Flourish wants the user to customize. This has led to some situations where there was a desire to change something relatively simple, but as it was simply not a feature in Flourish, this was impossible. For example, we, naturally, intended to keep exactly the same style and color of map in each visualization which included a map. However, after finding the style that fit our needs for the RWI visualization, we realized that this exact style and color palette couldn't be used on the other map visualizations. There was nothing to do about this - had we coded this from the bottom, it would have been as simple as changing the color reference in the code.

In addition to this, being reliant on a third party to "supply" the visualizations is not ideal. Several times, after not touching the visualizations for a few days, many things had changed (e.g., some of the maps had suddenly rotated 30° overnight - without rotating the overlaid content) without a reason or any interaction. Furthermore, one could imagine that Flourish, like other companies, could be the victim of a cyber attack or just a server or power outage, thus potentially rendering our data journalism piece essentially useless. Due to the time constraint associated with a project of this nature, a simplified data visualization tool was necessary for the completion of the project, but it's important to take note of the limitations associated with this.

When writing the article and researching related information, we also encountered another hindrance. In general, it has been a challenge to find information and data regarding the society in the countries and cities featured in this project. This could be due to several factors, such as under-reporting from the government, many people living 'under the radar', and, in general, fewew resources to conduct studies, report on the population, etc. This could, potentially, be an issue in general for studies investigating societal topics, like COVID-19 lockdowns and their impacts, in low-middle-income countries. This was an issue that we had not considered before starting this project, but it did make the process of gathering information quite complex.

8 CONCLUSION

As mentioned, our aim in this project was to visualize the complex results of a study in a way that was more accessible to a general audience. We believed that by using data visualization techniques and principles, we could create visualizations that were more engaging and intuitive than traditional academic papers.

By using a variety of tools and techniques, we were able to create interactive graphs and charts that allowed users to explore the data in a more intuitive and engaging way. The accompanying text and explanations provided more context and detail in an approachable format, and the article format made it easy to navigate and understand the results of the study.

However, we also encountered some issues and shortcomings along the way. One of the main challenges we faced was the limited customization of the visualization tools we used, which limited our ability to create more complex and sophisticated visualizations. Additionally, the study itself was still quite technical and complex, which made some of the results difficult to understand even with the visualizations and explanations we provided.

Despite these challenges, we still believe that our project was successful in making the results of the study more accessible and engaging to a broader audience. The visualizations were more accessible and intuitive than the original study, and the article format provided a convenient and engaging way to explain the results. We believe there is a certain importance related to the field explored in this project, as many studies, with quite interesting results, simply are not presented well enough to catch the attention of the general public.

Had we had more time, we would have liked to refine our approach and address the challenges we faced in this project. By improving the visualizations and addressing the limitations of our tools and techniques, we believe that it would be possible to create even more engaging and accessible ways to present complex research results to the public. In the next section, we will discuss some of the potential future works that we are considering.

8.1 FUTURE WORK

In the future, if we were to have more time to work on this project, we would focus on the construction and development of our visualizations. As previously mentioned, our limited time and experience led us to utilize a third-party visualization tool, Flourish, rather than building our visualizations from the ground up. While this approach allowed us to easily create visualizations, it also limited our ability to customize them. With more time, we would be able to learn and experiment, ultimately allowing us to create a more comprehensive and engaging narrative.

If we were to construct our visualizations from the ground up, we would be able to incorporate features that are conceptually simple but cannot be implemented using Flourish due to its limitations. Some of these features that we plan to implement in the future include:

- Present RWI data as a heatmap instead of individual dots
- Interactive elements, such as hover-over text and clickable buttons, to provide a more engaging experience for our audience
- Customized color schemes and designs to enhance the overall aesthetic appeal of our visualizations
- Show original Daily Flow data 'behind' 7-day rolling average in a faded color
- Customize the x and y axes more than what's possible in Flourish
- Remove borders around some of the visualizations

In addition to the aforementioned improvements, we also plan to explore the "Economic Access indicator (EA)" developed by Keerd (2021). This measure provides insight into the economic value accessible from a given "cell" or micro-region, making it a useful tool for understanding how different socio-economic groups and their economic opportunities were impacted by COVID-19 lockdowns. However, due to the complexity of calculating the EA and the limited scope of this project, we chose to focus on revising the other visualizations in our study. Nevertheless, we believe that if possible, visualizing the EA could greatly enhance the value of our data journalism piece.

Due to the amount of text in our piece, we would like to develop a navigator that could 1) show how far in the story the reader is and 2) allow the reader to click the section they want to navigate to. The conceptual idea is drawn and shown in figure 13 below.

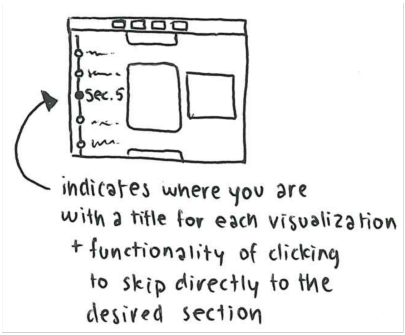

Figure 13: "Bus-stop" navigator

Another possible approach would be to use the Lockdown Severity and Daily Flow line graph as a timeline, dynamically changing as you view the story. This would require a change in the story's structure as well as a lot of code, but would add a new dimension of interaction.

Lastly, as shown in section 4.1.2, there is an unused column from the RWI data, containing an 'error' value. This is not in the scope of the thesis by Keerd (2021), but visualizing this uncertainty within the data would certainly be an interesting future point of research.

BIBLIOGRAPHY

Bernard, S., D. Blood, C. Tilford, M. Harlow, C. Nevitt, Ændrew Rininsland, A. Smith, M. Stabe, and A. Wisniewska (2022). Lockdowns compared: tracking governments' coronavirus responses. https://ig.ft.com/coronavirus-lockdowns/. Last accessed 13/12-2022.

Cairo, A. (2016). *The Truthful Art - Data, charts, and maps for communication*. Asian Development Bank.

Ghebreyesus, D. T. A. (2020, 3). Who director-general's opening remarks at the media briefing on covid-19.

Ghosh, A., S. Nundy, and T. K. Mallick (2020). How india is dealing with covid-19 pandemic. *Sensors International 1*, 100021.

Hale, T., N. Angrist, R. Goldszmidt, B. Kira, A. Petherick, T. Phillips, S. Webster, E. Cameron-Blake, L. Hallas, S. Majumdar, and H. Tatlow (2021, 03). A global panel database of pandemic policies (oxford covid-19 government response tracker).

Harris, R. (2016). Svelte. https://github.com/sveltejs/svelte.

Jackson, J. K. (2021, 11). Technical Report, Congressional Report.

John Branch (2012). Snow fall. https://www.nytimes.com/projects/2012/snow-fall/index.html#/?part=tunnel-creek. Last accessed 13/12-2022.

Keerd, E. (2021). The impact of covid-19 lockdowns on the access to economic opportunities in the metropolitan areas of low and middle-income countries.

Mathieu, E., H. Ritchie, L. Rodés-Guirao, C. Appel, C. Giattino, J. Hasell, B. Macdonald, S. Dattani, D. Beltekian, E. Ortiz-Ospina, and M. Roser (2020). Coronavirus pandemic (covid-19). *Our World in Data*. https://ourworldindata.org/coronavirus.

Moere, A. V. and H. Purchase (2011). On the role of design in information visualization. *Information Visualization 10*(4), 356–371.

Mogaji, E. (2020). Impact of covid-19 on transportation in lagos, nigeria. *Transportation Research Interdisciplinary Perspectives 6*, 100154.

Nations, U. (2021). Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all.

Riley, K., C. Aguilar García, and E. Hollinghurst (2019, September). How sky news revealed flourish's secret scrollytelling feature. https://flourish.studio/blog/scrollytelling-flourish-sky-q-and-a/. Last accessed: 04/12-2022.

rusellsamora (2017). scrollama. https://github.com/russellsamora/scrollama.

Segel, E. and J. Heer (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics 16*(6), 1139–1148.

sitiaish (2021). Chronicle story template. https://gist.github.com/sitiaish/366003de97e9118eecb9a7b93b03d013/revisions. Cloned October 2022.

UserTesting.com. 7 gestalt principles of visual perception.

WHO (2022, 3). Covid-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide.

A APPENDIX

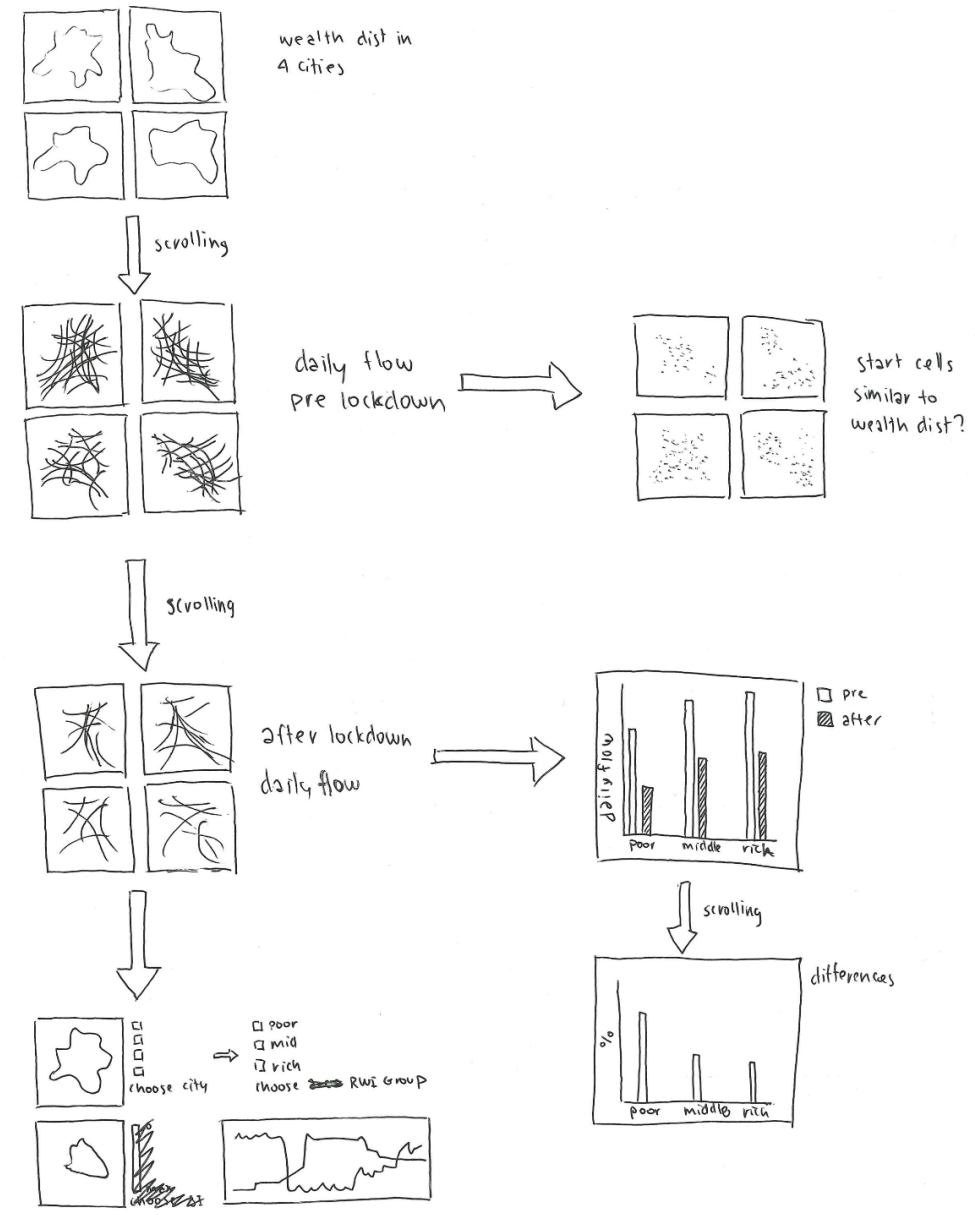

Figure A1: Drawing/storyboard of how we wanted to tell the story

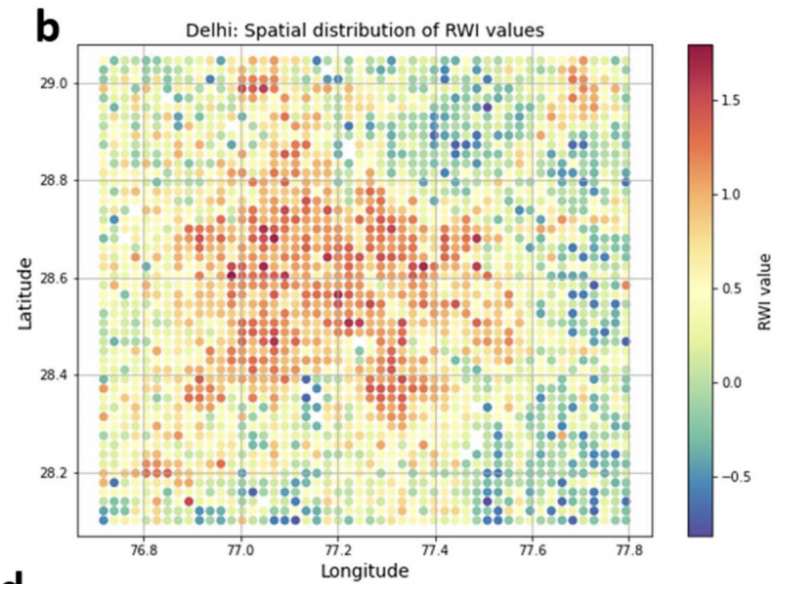

Figure A2: Visualization of the RWI visualization/ heatmap in the original master thesis

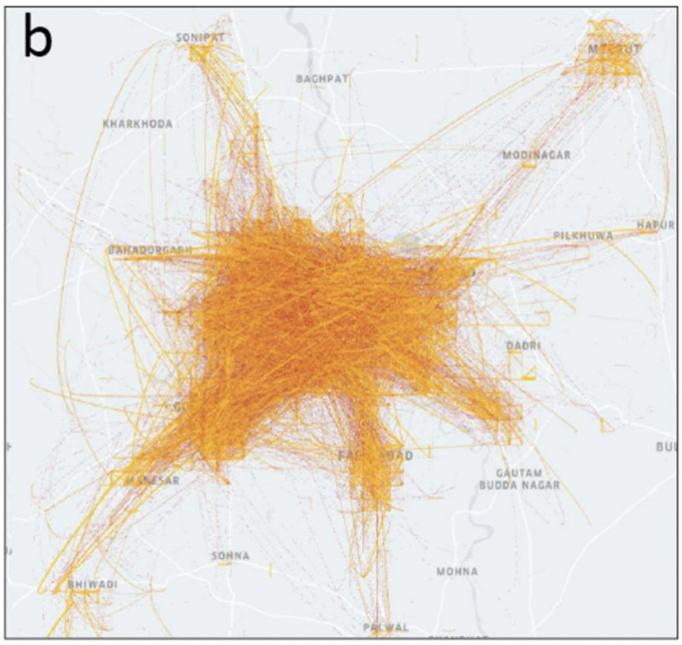

Figure A3: Visualization of the mobility graph during lockdown in the original master thesis

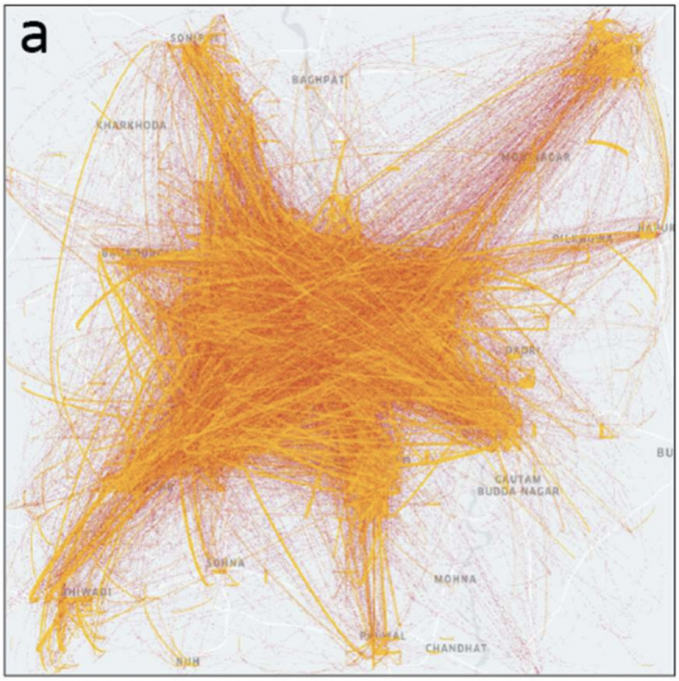

Figure A4: Visualization of the mobility graph before lockdown in the original master thesis

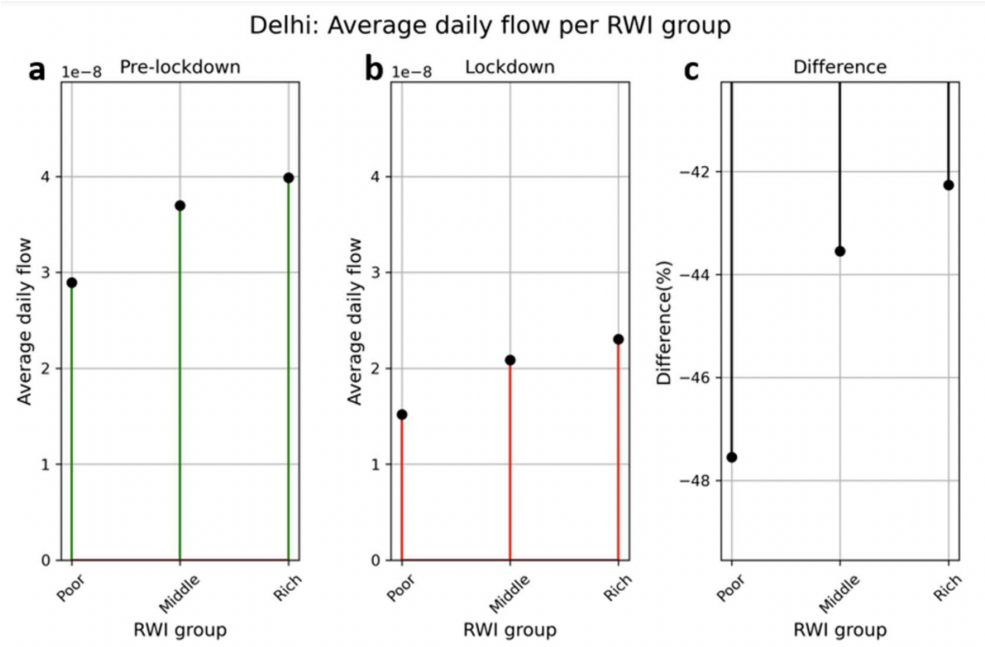

Figure A5: Lollipop chart in the original master thesis visualizing the mobility decrease for each socioeconomic group

Figure A6: Visualization of the map of Delhi in the original master thesis

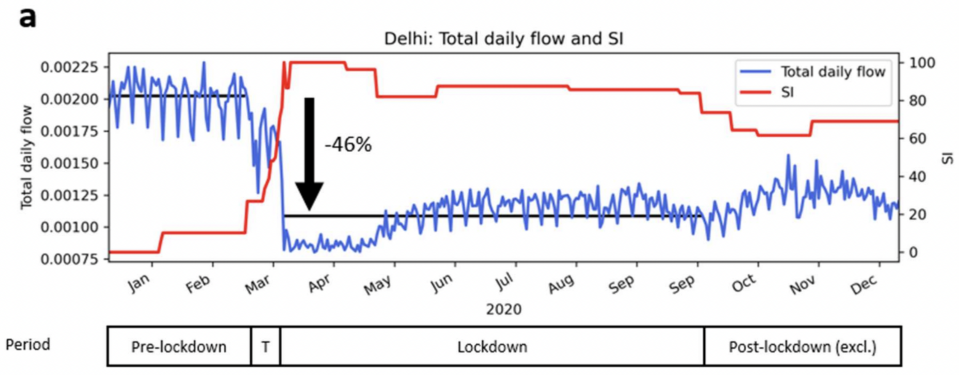

Figure A7: Original visualization made in the master thesis of how mobility decreased in Delhi as the lockdown began.

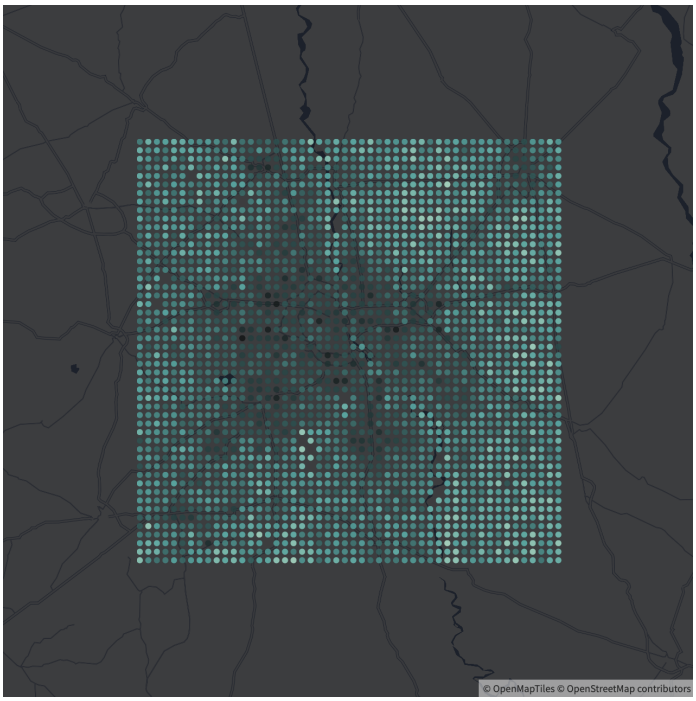

Figure A8: The redesigned heatmap visualization of Delhi and its surrounding metropolitan area using the Relative Wealth Index measure

Figure A9: The redesigned map of Delhi showing the city and its surrounding metropolitan area

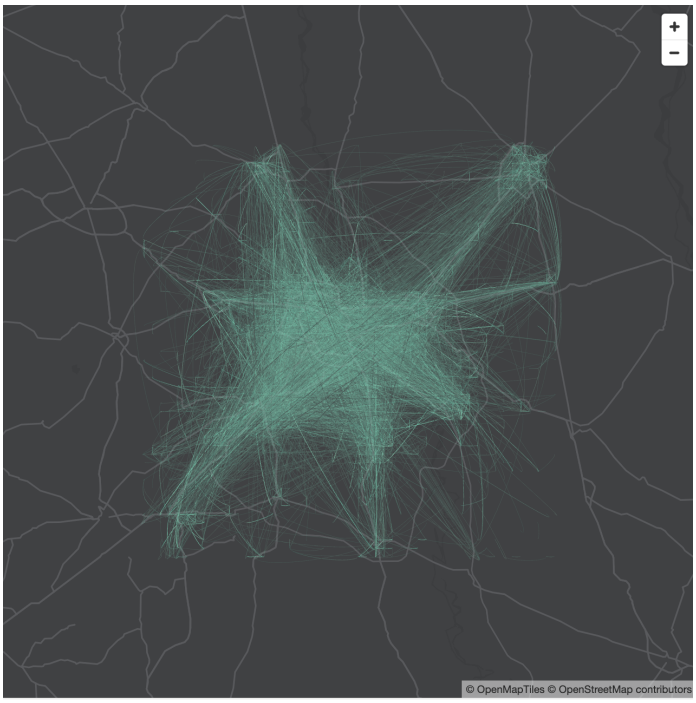

Figure A10: The redesigned visualization of mobility in Delhi and its surrounding metropolitan area prior to the lockdown.

Figure A11: Our final color palette consisting of teal and gray hues

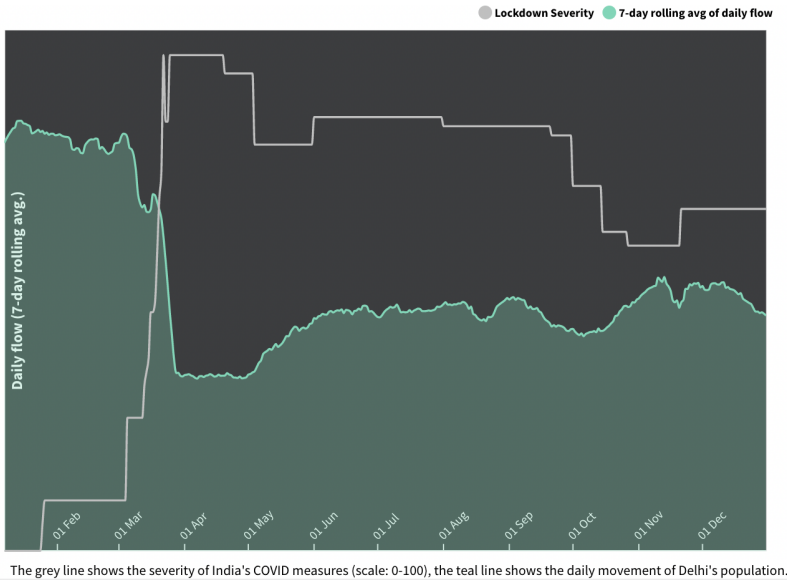

Figure A12: The improved and redesigned visualization of Lockdown Severity and Daily Flow in Delhi, India