FIT 5147 NARRATIVE VISUALIZATION PROJECT

COVID-19 IN WORLDWIDE AND CHINA

STUDENT NAME: ZHIYIN WANG

STUDENT ID: 31436285

LAB: 24

TUTOR: PRATIK BHUMKAR

CONTENTS

1. INTRODUCTION

1.1 AIM OF THE NARRATIVE VISUALIZATION

2019 Novel Coronavirus (2019-nCoV) is a coronavirus identified as the cause of an outbreak of respiratory illness first officially been detected in Wuhan, China. While coronavirus continues to spread over the world, China the first country that encountered the virus have successfully contained the outbreak. Thus, the report will focus on comparing China's performance to other countries involved in this pandemic.

The narrative visualization aim to convey to readers the following message:

- Overview of the pandemic in world counties
- Comparison between China and other countries
- Spread of the virus over time
- How virus is contained over time in China

1.2 INTENDED AUDIENCE

The visualization is designed for audiences that are interested in the spreading of covid-19 in past month and especially the pandemic in China. There is no target demography of audience, but are expecting more audience from China that are more likely to be interested in finding out the situation in China. The charts in the visualization are designed for non-technical audience with basic analytical skills that can identify clear trends and patterns.

2. DESIGN

2.1 FIVE DESIGN SHEETS

The five design sheets method is used to generate design ideas and compare between alternatives. Prior to create any designs, ideas are generated in the brainstorm phase (Figure 1). In this design sheet, every possible ideas are written and sketched, such as: messages to convey to audience, narrative structure, data source, possible implement of graphs and UI design etc. After brainstorming, the ideas are managed and filtered to be ready to apply in alternative designs in the next phases of FDS.

2.2 INITIAL DESIGNS

There are three alternatives designs created, suitability of each design is accessed before converging on to the final realization design.

The first design (figure 2) includes a line chart and a map of China. Both graphs have options to select and the map also has a data range slide bar available. This is a story-telling design that is intended to give the user an clear overview of the pandemic and how virus spread over time in different countries. User is provided with comparison of three options between countries on the line chart and comparison between provinces in China on the map. The map is designed to be filled with colour according to case numbers in each provinces, thus provides a clear view of how the data spread regionally. However, the design is lack of in depth comparison and there are rooms for more interactive features.

The second design (figure 3) used world map to present covid-19 pandemic in counties and a dot plot showing confirmed cases per day in China. Different to the previous design, a timeline element is added to the design through popups on the dot plot. The popups includes information of government actions and events occur in China during the pandemic on the given date. This design provides more detail regarding China's performance compare to last design. More interactions and information are provided to user, which also makes this design harder to implement.

The third design (figure 4) focus more on providing user with in depth comparison between China and other countries. Bar chart is used instead of line/dot chart with more options. Ratios are provided rather than pure cases counts, countries are compared at a more analytical aspect. The timeline of Covid-19 in China is designed separately from graph, to inform the user how China is able to contain the outbreak successfully. As this is a more analytical design, it is harder for non-technical users to use, users need to understand the calculation and meaning of each ratio beforehand.

2.3 FINAL DESIGN

In consideration of the target audience of the visualization, the final design includes a world map and a dot plot. User interactions includes options to select, slide bar and checkboxes. A map perform well when applying a drill-down-story structure. The world map provide users with an overview of the pandemic in world counties and comparison between the countries. The cases are represented with circles, the size and colour of the circle are determined by the case numbers. The circles are selected to be red colour group, as red can be easily seen and has a high degree of “visual popout”. Human’s visual working memory has a limited capacity. In this case, multiple attributes are encoded into one visual object, more information can be held in user’s working memory. By adjusting the size and colour of circle, the salience, discriminability and accuracy of the visualization increases. The user is able to easily identify location through colour and discriminate/compare values through size. The dot plot has checkbox to tick on/off to avoid the graph become messy and filled up with data. The checkbox also allows comparison between more than two countries. Audience are not expected to have high analytical skills and knowledge, thus only death ratio and recovery ratio is included in the option in order to pass an effective communication. In considered of relevancy, top countries in confirmed cases will be in the selection of checkbox. The timeline of Covid-19 China is highly relevant to the topic. The important regulations and events happened in China is considered as an essential message the design wish to convey and gives more insight about China’s performance in the pandemic.

3. IMPLEMENTATION

3.1 LIBRARIES USED

The visualization is created through Shiny R and following libraries have been used:

dplyr: This library is mainly used to filter the data to be ready for visualization and use of %>% between arguments. It is also used for data transformations and manipulation.

tidyverse: This library is applied to group data by appropriate variable and merge of two dataset.

Shiny: This is the library used to set up web framework and create visualizations on shiny app. Interactive visualizations are achieved through Shiny, such as built in input widgets, CSS themes and JavaScript actions. ("Shiny", 2020)

leaflet: This library is integrate with shiny very well to create interactive maps.

ggplot: This library is used to create charts and plots.

plotly: This is a interactive graphing library which turns plots and graphs into interactive visualizations with features such as: zoom in/out, label etc. ("Plotly R Graphing Library", 2020)

Scales: This is used to show percentage scale on graph.

3.2 DETAIL IMPLEMENTATION

The visualization page contains two section, “Covid-19 World” and “Covid-19 in China”. In order to separate the two section and make them easy for user to access, tab panels are designed into UI through shiny.

In the “Covid-19 World” section, a world map is presented with options to select on a side-bar panel. The world map is set to cover nearly 100% of the web page to allow user to have a clear view of how coronavirus spread around the world. Select options and date range is associated

with the map for user to control the data displayed. The two input options act as filters, filters the data appears on the map and is achieved through the dplyr library. Every time user select on a option or change the date range, a reactive function will filter the data according to the input options and output the data out on to map as user wish. By providing the options to filter the data, user is allowed to freely explore the spread of Covid-19 in the world over time. Values shown on the map are displayed as circles varies in size and in red colour, the more cases in a country the larger and darker the circle become. Labels are created for maps through leaflet. When user hover on the circle in a country, label contains country name, number of cases pops up to help the user better understand the "story" of Covid-19 in the world. Stamen. Toner" map is used in this case, it is a black and white map which makes the red colour circles more stand out in the whole visualization. The option panel is designed to be displayed over the map, therefore the panel is set with an initial low opacity and draggable to allow user to see through the panel. When user hover on the panel to filtered the map with options, the panel will display as normal and become transparent again when hover off. This design is intended to help user to appropriately focus attention on the map. This is achieved through modifying the panel setting through HTML CSS codes. A bar chart showing comparison of top countries is included in the panel, created through ggplot, which can also be filtered by the options in the panel. Top countries are selected by filtering the dataset and reorder the data in descending cases numbers. Legends are added in coding of both map and bar chart to explain the use of colours in each graph.

In the "Covid-19 in China" section, options to select are also available for users to filter the data. Different to the first section, checkboxes and date range selection are implemented through shiny. Checkboxes are designed to avoid too much information to be displayed on the plot simultaneously and allow for free comparison between countries as user wish. A interactive plot demonstrates death/recovery ratio comparison is created through ggplot and plotly library, the use of plotly adds labels and zoom features to the plot. A pie-chart showing cases in China provinces is created with ggplot, with percent/total labelled on corresponding provinces displays on top. The percent value is calculated through scale library and pie-chart is created by transform a bar chart to display in circle shape. The colours in bar chart and dot plot is determined by groups with legends showing each group. Timeline is applied as label on the plot through annotate in ggplot. The information regarding timeline is extracted from WHO timeline website. ("WHO Timeline – COVID-19", 2020)

4. USER GUIDE

4.1 OVERVIEW

The visualization contains two tab sections, user can click on “Covid-19 world map” and “Covid-19 in China” in top panel to switch from the two sections.

4.2 COVID-19 WORLD MAP

- Click on the buttons in “Measurements” including: “Confirmed Cases”, “Death”, “Recovered” to see corresponding case numbers display on the map and bar chart. Only one option can be selected.

- Drag the date slide button and slide from left to right to show cases in the selected date range on map and bar chart.

- Use the selection “Select numbers of top countries for comparison” to change the numbers of countries show on the bar chart. Top 3 = the top three countries with the most cases. The cases show depends on the selection of measurements. Only one option can be selected.

- Zoom in and out on map using scroll button on your mouse or the “+” and “-“ button on the top left side of the map.

- Hover on the circles in the map to see the country name and number of cases at the location.

4.3 COVID-19 IN CHINA

- Select measurement from side bar to compare the data on “China vs other countries” plot, only one measure can be selected.

- Click on checkboxes to select the countries wish to compare in the “China vs other countries” plot, multiple countries can be selected.

- Click on dates in the date range selection to choose the begin and end date of the data shows in the “China vs other countries” plot and pie-chart

- Hover on the dot/line in the “China vs other countries” plot to see details

- Click on interactive feature on top right side of the plot to zoom in/out, move or save the graph etc.

- Hover on the timeline label to see the exact date of the action being taken during the pandemic

5. CONCLUSION

Although it is perceived that China has overact to the outbreak of coronavirus compare to other countries, the coronavirus still spread widely and quickly across the country from January to February. Luckily, China’s overaction and regulations implement have managed to contained the outbreak by the end of April. On the other hand, the pandemic seems out of control in the rest of the world, coronavirus began to spread in a rapid pace in other countries. The outbreak in other countries started in March with increasing confirmed numbers and death ratio surpasses China by the end of March.

The visualization is created through Shiny R, interactive maps, charts and input widgets is achieved by implementing varieties of R packages. The integration between R packages such as leaflet and shiny, ggplot and plotly was widely used to create interactive visualizations. In this project, R studio is mainly used on providing a “storying telling” visualization to the audience. However, the visualizations can also be done in an different aspect using other tools such as D3. Despites, creating visualization with D3 requires more development time compare to R studio, D3 offers more flexible integration with analysis, visualization can be designed to be more analytical on D3 ("D3 and R Shiny: The When & What of Data Viz, Part II", 2020). In future projects, use of animations may be applied in a narrative visualization, to guide audience through the background information and add more context to the story.

Nevertheless, interactive poster is another tool that maybe applied in the visualization to demonstrate different aspects of the story to the audience.

6. BIBLIOGRAPHY

Data source:

Novel Corona Virus 2019 Dataset. Kaggle.com. (2020). Retrieved 20 June 2020, from https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=covid_19_data.csv.

World capitals gps. Kaggle.com. (2020). Retrieved 20 June 2020, from https://www.kaggle.com/nikitagrec/world-capitals-gps.

WHO Timeline - COVID-19. Who.int. (2020). Retrieved 20 June 2020, from https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19.

Library:

Shiny. Shiny.rstudio.com. (2020). Retrieved 20 June 2020, from https://shiny.rstudio.com/.

Plotly R Graphing Library. Plotly.com. (2020). Retrieved 20 June 2020, from https://plotly.com/r/

Other:

D3 and R Shiny: The When & What of Data Viz, Part II. Medium. (2020). Retrieved 20 June 2020, from https://medium.com/opex-analytics/when-what-data-viz-part-ii-b4f76f5d0a29.

WHO Timeline - COVID-19. Who.int. (2020). Retrieved 20 June 2020, from https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19.

7. APPENDIX

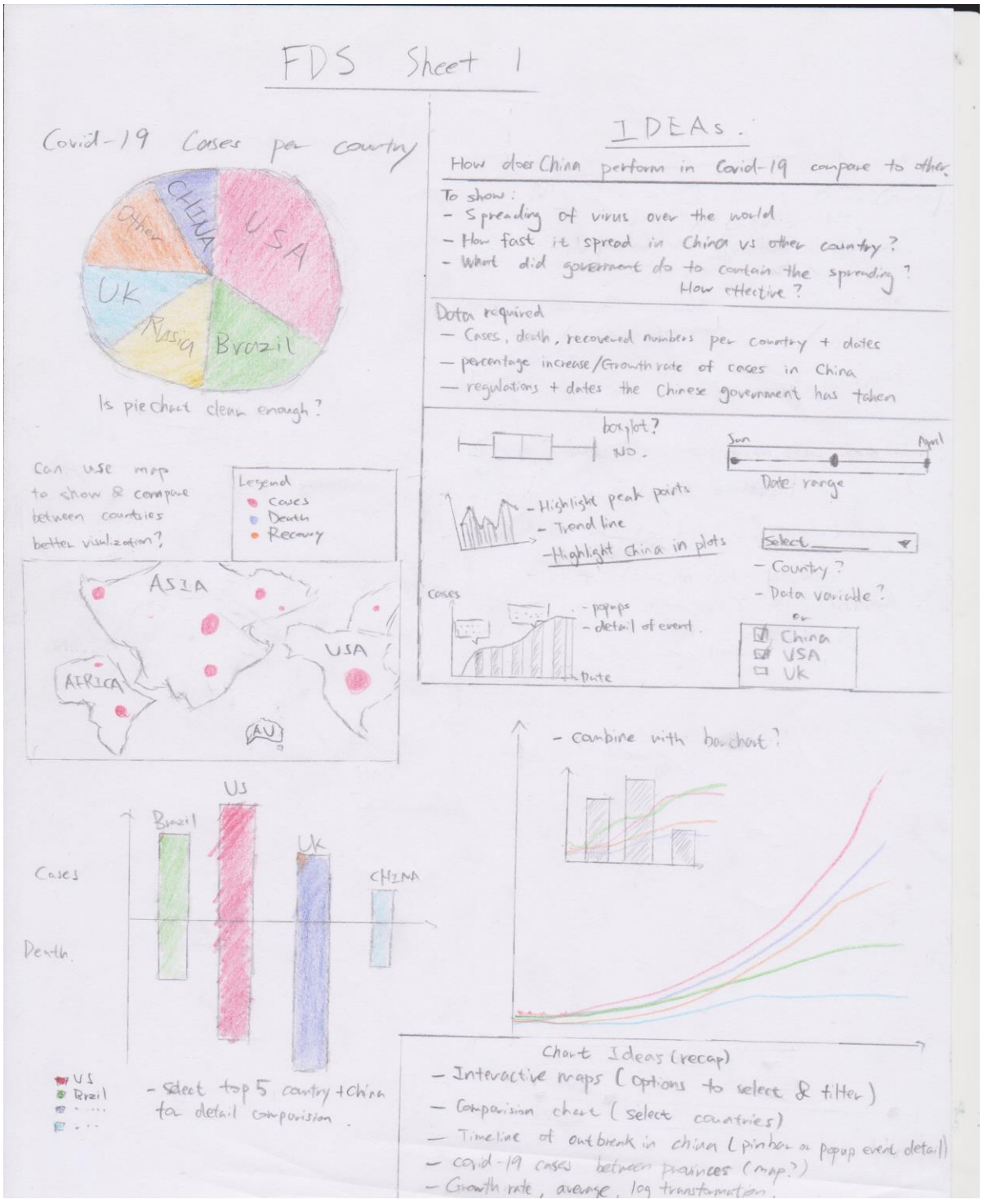

Figure 1 Brainstorm phase

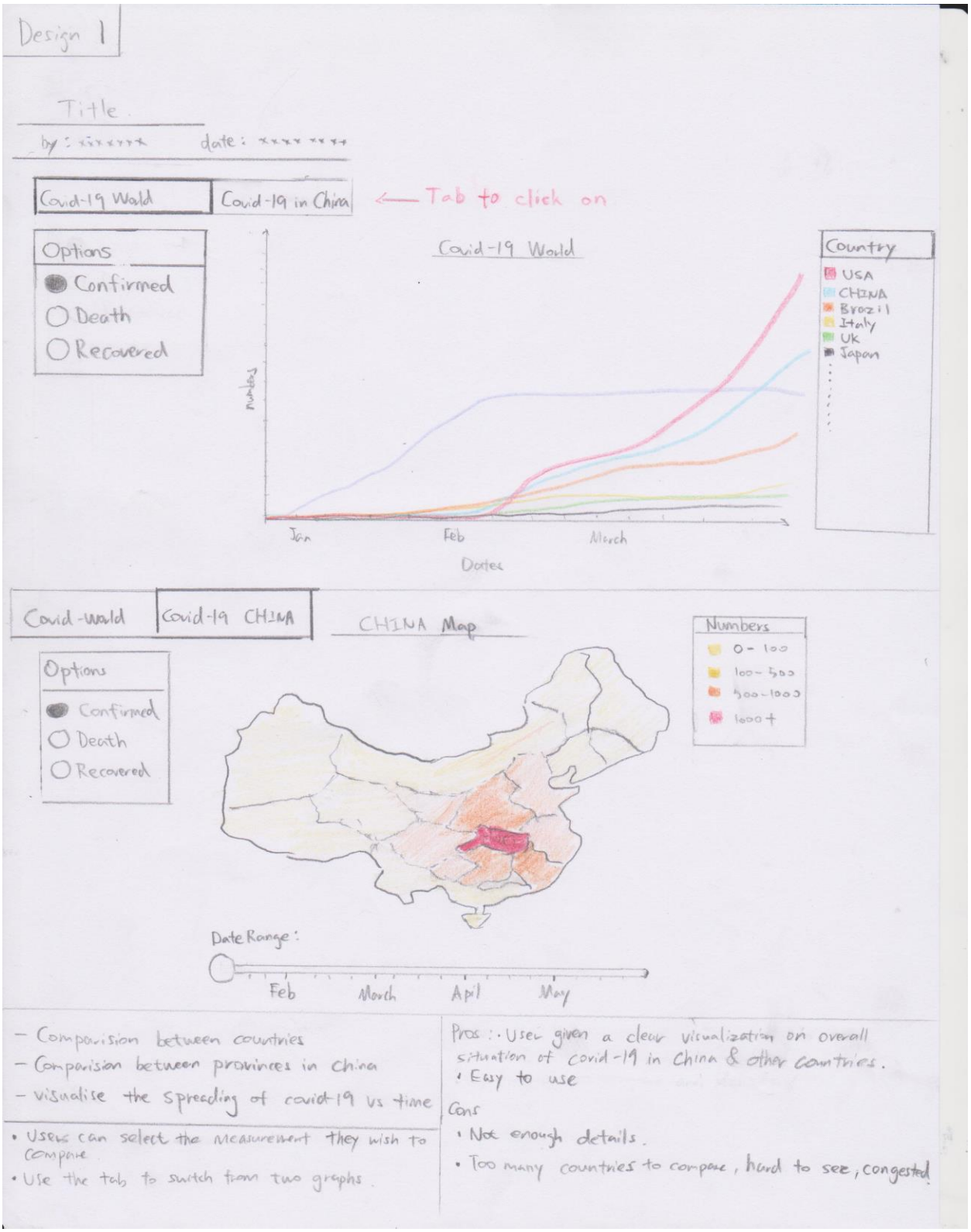

Figure 2 Alternative design 1

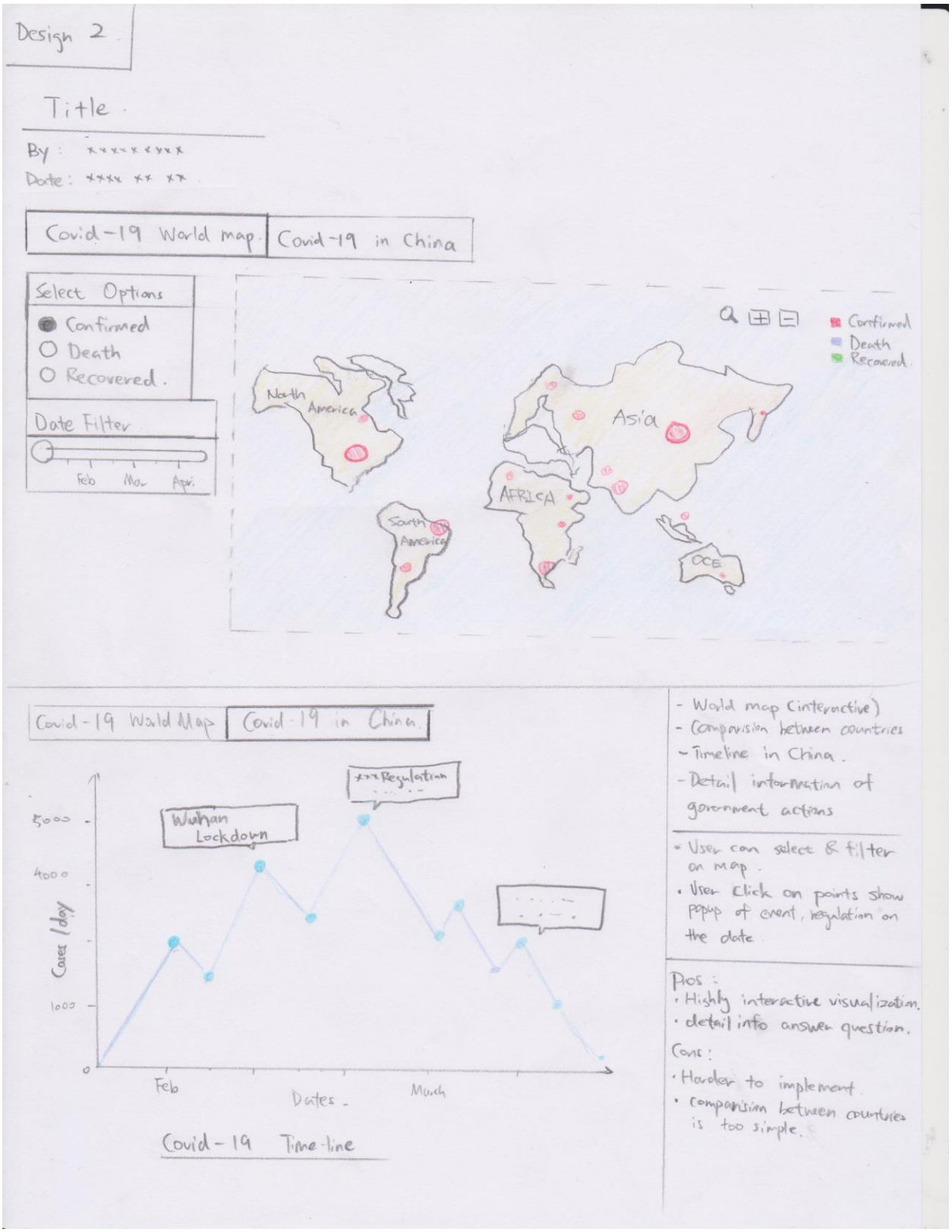

Figure 3 Alternative design 2

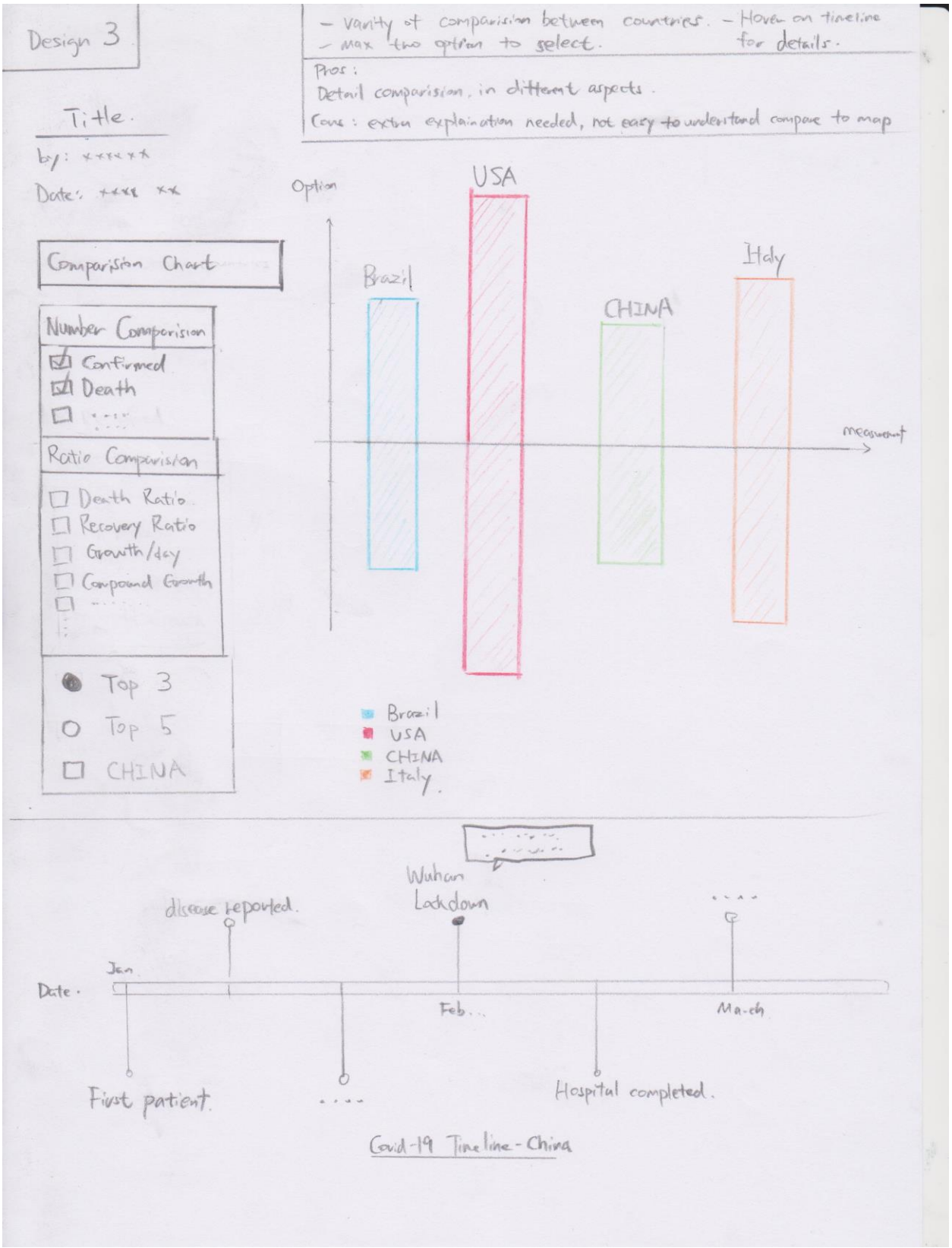

Figure 4 Alternative design 3

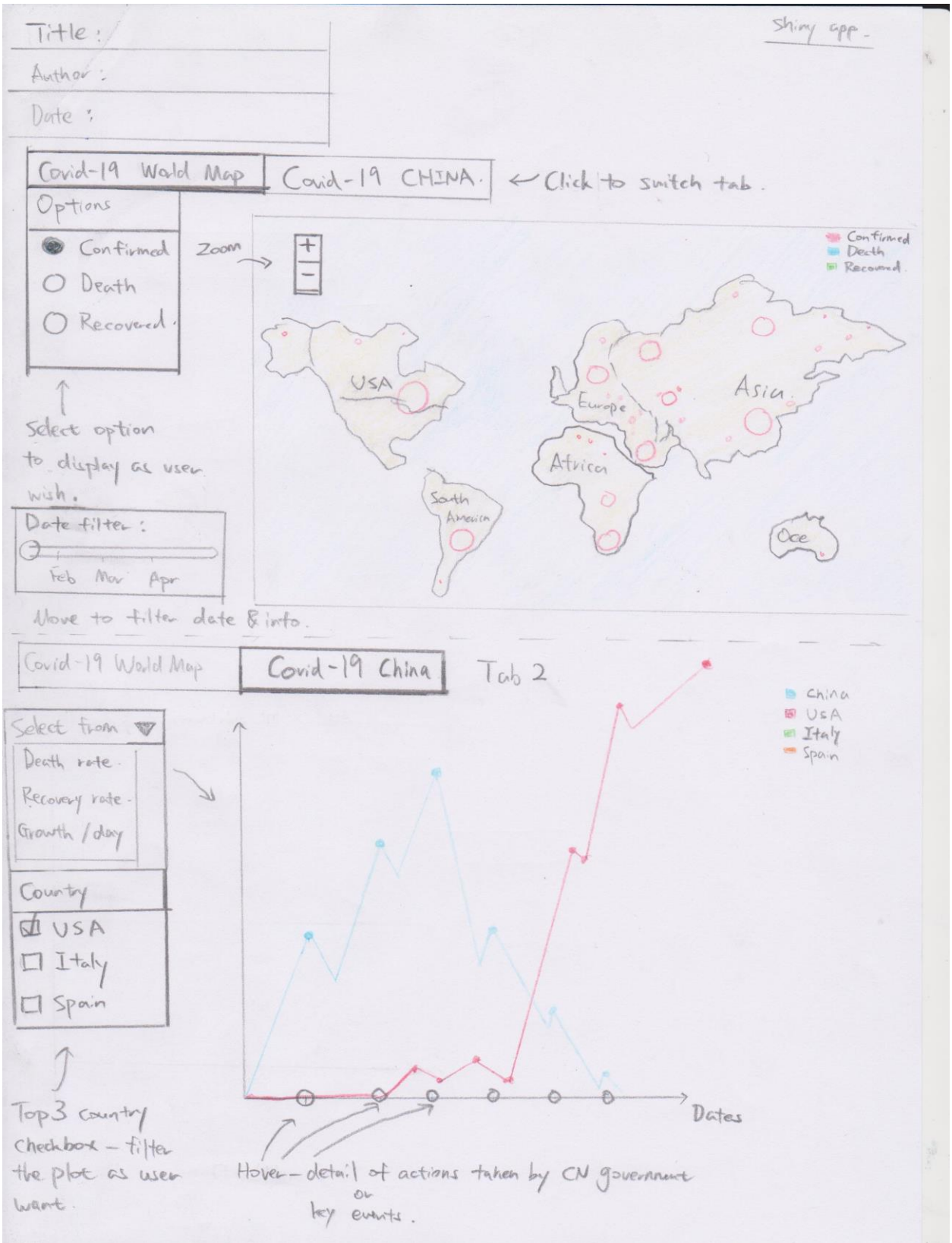

Figure 5 Final design