

DATA VISUALISATION

ELEMENT 1: INDIVIDUAL REPORTS

YU SHING, LUI – A0201746
TEESSIDE UNIVERSITY | 2020/21

Introduction

The impact of the Coronavirus (COVID-19) since its outbreak in early 2020 has been substantial and has been affecting all walks of life on a global scale. Until 16 December, 2020, almost 73 millions confirmed cases and 1.6 millions people died from Coronavirus (Source: [Financial Times, December 2020](#)) and thus making location analytics more crucial than ever. Nowadays, as everyone can obtain information about Coronavirus through internet, newspapers and all sorts of medias, this makes data visualisation becomes more useful and comprehensible in presentation and illustration. In this report, a wide range of sources about Coronavirus data will be explored and a list of quality criteria will be developed.

Quality Criteria

Before starting to evaluate and suggest in these graphs, this section will elaborate more about the quality criteria and it will be a framework to appraise the following graphs.

- 1) Data-ink ratio – The ratio of data-ink is one of criteria to affect the quality of the visuals. Creators have to pay attention how much data-ink they use in the visuals. Edwards Tufte mentioned that a large share of ink should present data-information on a graphic. If the ink is changed, the data is also changed. Data-ink is the core of a graphic which cannot erase, and the arrangement of non-redundant ink is responding to variation in the numbers represent. That means the graphic will lose the content once data-ink remove from the image (Tufte, 2001). Yuk and Diamond stated that non-data-ink does not represent information in the graphic, but it uses for scales, labels, and edges. Visual content makers should try to include data-ink only in the graphic and they have to erase the things which is not related to data-ink. It can avoid the target audience to draw attention on irrelevant elements (Yuk and Diamond, 2014).
- 2) Reducing clutter – The problem of cluttering will create immoderate and extra load in visuals. Cluttering makes target audience misunderstanding and it should consider to avoid it. The main purpose of reducing clutter is making visuals less complicated than necessary. Nussbaumer mentioned that the existence of clutter will affect or make it worse user experience for the target audience in our visual communication. It will create a visual more complicated than it is. It may make audience lost a passion to understand what a visual present and it will lose the meaning to create visual. Thence creators should distinguish at the visual elements that allow them to use it for communications. Normally, they necessary to find out non-informative value or fulfil informative value to identify and eliminate clutters. It may be one of the most important things on data visualisation (Nussbaumer, 2015).

3) Preattentive attributes in graphs – Preattentive attributes is a powerful tool. If visual content makers use it wisely, they can enable the target audience to see what they want them to see before they even know they are seeing it. Ware indicated that colour, size and placement are important in the preattentive attributes. which they can highlight the most important parts of visual then draw target audience's attention quickly. The target audience will admire that someone provides visuals to make them interact with these data visualisation and will be more generous to provide their time to it than a visual that feels like work to consume (Ware, 2012). Jay also agreed that colour is effective, but it must be used sparingly. It cannot stand out if it has too much variety and it should be making sufficient contrast to draw audience attention. Furthermore, consistent of colour is necessary in some cases. Normally audience needs to take time to get used to the colour's pattern and they will assume the same details for the rest of the presentation. If makers decide to use this colour to represent for this graph, they should avoid use the same colours for other purposes if possible. Otherwise, the audience will be confused because they use the same colour (Jay, 2012).

4) Gestalt principles – The principles are accurate and useful descriptions of visual perception. It can offer useful insights. Nussbaumer stated that these principles have six elements, which are proximity, closure, similarity, continuity, enclosure, and connection (Nussbaumer, 2015). Gordon also mentioned that the principles can define seeing visual information from human brain. It is the most effective way for designer who make visuals to communicate information (Gordon, 2004).

5) Do not overcomplicate – Once the graph looks too complicated, an audience perceives needs more time to understand and they may not interest to spend time to understand it. William pointed out that too much variation can overcomplicate the graph. Legible is necessary and use a consistent, east-to-read font and make data visualization approachable by leveraging visual affordances. Then, using straightforward language to approach the audience because they may not be familiar in it. Do not try to sound smart then make it complicated, as it will annoy audience and it will take risk of making them feel uncomfortable. It is bad user experience for them (William, 1994).

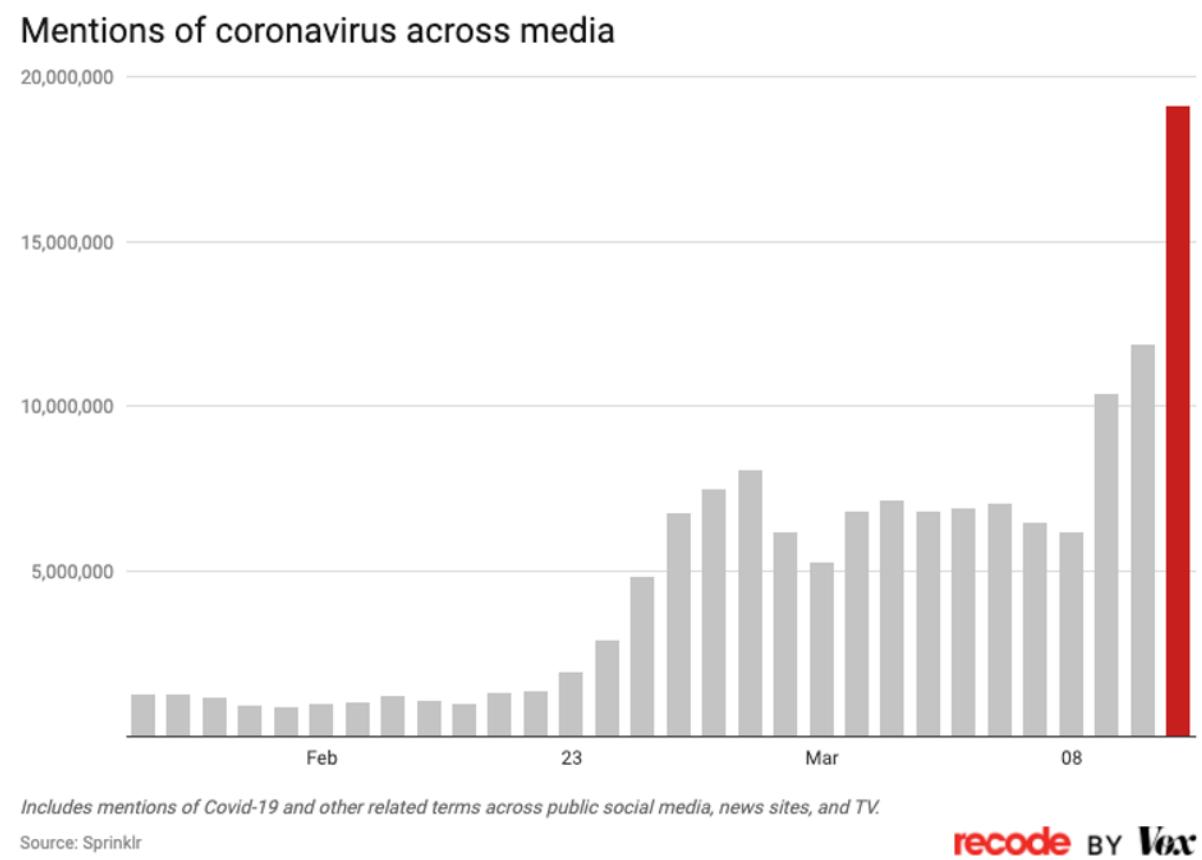
Portfolio

In this section, the table will provide a general idea about following ten graphs. All of graphs come from an official website of organization, newspaper, and commercial etc. It will mention about the title of each graph, type of graph, and the reason to choose this graph to evaluate.

Title of graph	Type of graph	Reason to choose this graph
Mentions of Coronavirus across media	Bar Chart	Exploring how to execute data-ink ratio in bar chart
Italy and Spain's daily death tolls are falling; in the UK and US daily deaths may be plateauing	Line Chart	Discussing how to reduce clutters in this graph
Coronavirus (COVID-19) and Tobacco Use	Infographic	Understanding the importance of Infographic's element
How has Covid-19 changed our spending habits?	Bubble Chart	Getting to know how preattentive attributes to draw attendance's attention
Top 50 countries of confirmed cases of COVID-19.	Radial Bar Chart	Improving a bar ratio of the chart and how to present bar chart in different way
Index Drops Amid the COVID-19 Outbreak	Line Chart	Learning how to use line chart and point out the story line at the same time
The UK government has awarded about 400 Covid-19 contracts	Treemap	Looking for the elements of treemap and express in a proper way.
World's Economic Programs Against the Coronavirus	Choropleth	Discussing how to improve of using colour tone in the chart
The Anatomy of the \$2 Trillion COVID-19 Stimulus Bill	Sankey Chart	Understanding how to use the chart to represent the data in different sections
The Two Fast Tracks of COVID-19	Infographic	Finding out how to make it better in this graph

Graph 1: Mentions of Coronavirus across media

This is a bar chart that is coming from American news website, Vox.com. It accumulated a record to mention Coronavirus-related term from social media. The data collected from mid-February to mid-March 2020 and target audience was public general. From the graph, we can find that people were not much mention about Coronavirus in mid-February, but it was increased dramatically on mid-March.



(Source: <https://www.vox.com/recode/2020/3/12/21175570/coronavirus-covid-19-social-media-twitter-facebook-google>)

Evaluation

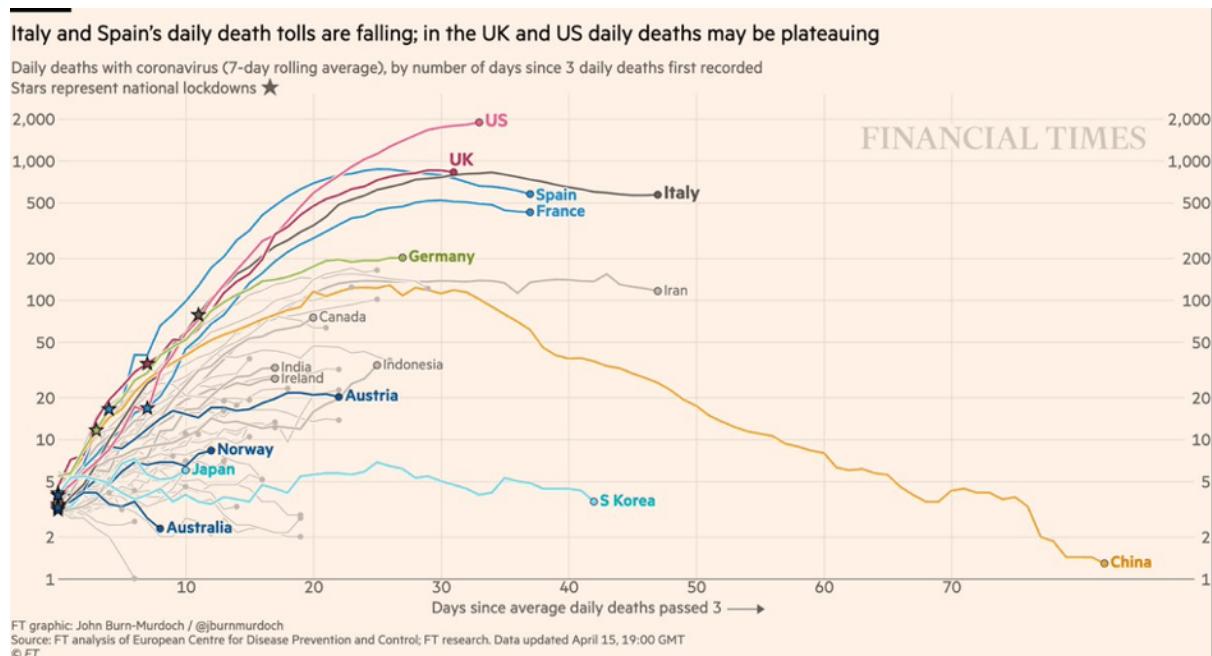
The graph can understand that it is trying to make high data-ink ratio. It uses the data-ink only and cannot find any non-data-ink. The graph does not have any clutter to make it complicated to read. The Gestalt principle of closure is also applied in the graph, which eliminate the clusters. On the other hand, the graph tries to point out the last value in red colour, which want to give target audience to know about the value is changed rapidly. It can make target audience to draw attention easily and let them know where the main point in the graph is.

Recommendations

Target audience may want to know more the number of changes from February to March, but it cannot find any numbers to compare during this period. It may display the number at the beginning, middle and the latest to give audience to know how difference in the short period.

Graph 2: Italy and Spain's daily death tolls are falling; in the UK and US daily deaths may be plateauing

This is a line chart, and it comes from international daily newspaper's website, Financial Times. The graph showed daily death with Coronavirus by number of days, which start from three daily death record. The data started to collect from early 2020 and the last update was 15 April 2020. It wanted to tell public about daily death with Coronavirus by different countries. It illustrated which countries had more successfully mitigated and which they had not.



(Source: <https://www.ft.com/content/c764b98d-ae03-41ac-a2ca-8309f32d5e1c>)

Evaluation

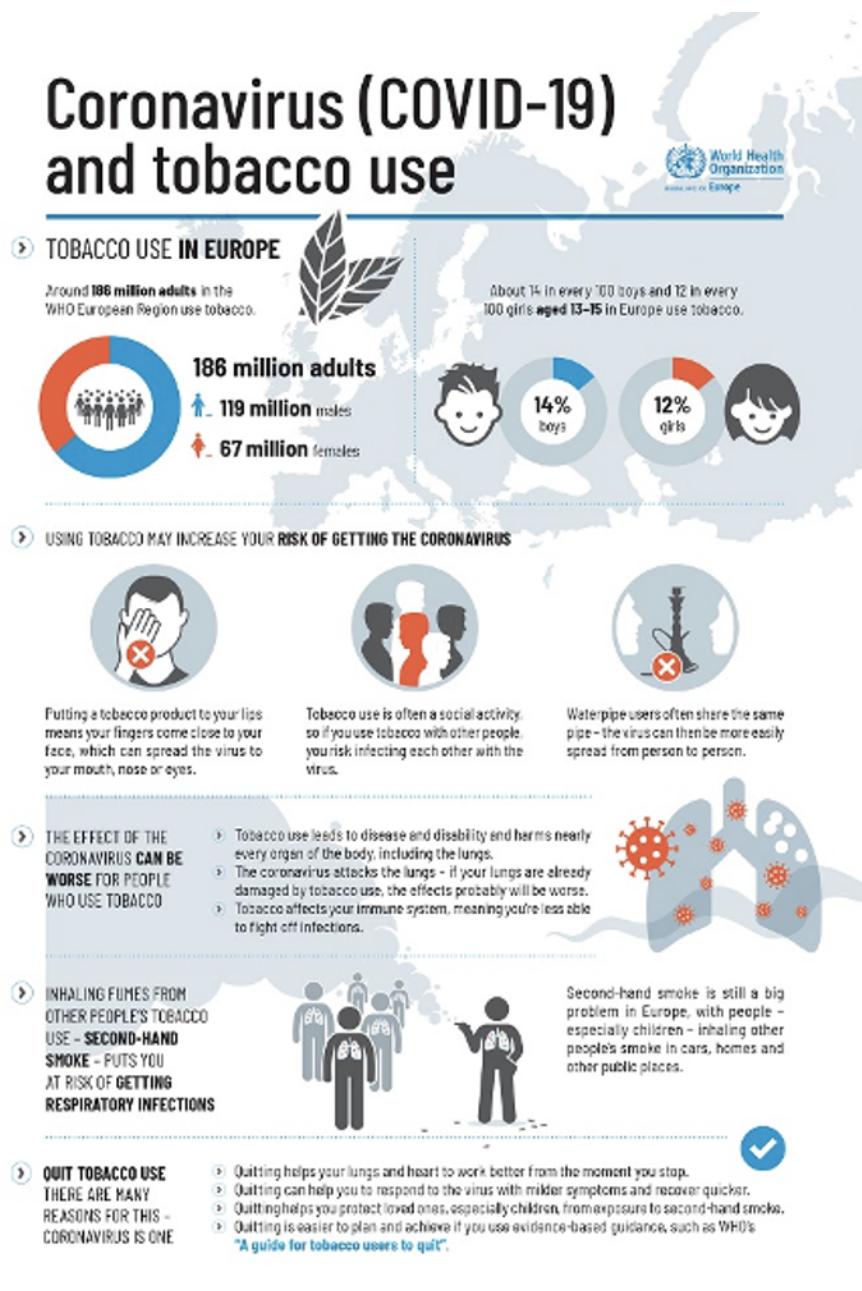
The graph can do it better for the data-ink ratio as it has two sides of y-axis to show how many people death from coronavirus. However, it is not much help the audience to understand the graph. Also, it may try to reduce clutters because extra load in visual will not increase understanding. On the other hand, the designer applies the principle of similarity improperly in colour pattern. It tries to use similar colours to draw audience's attention, but it makes audience more confuse because it cannot find any relationships with them.

Recommendations

Sample and clean is always the key elements for data visualisation. Suggest crossing out one side's label of y-axis, which it can increase data-ink ratio. We also may consider removing star symbols that we cannot find any important meaning for them. The grey lines may also delete as it can make visual less complicate and reduce clutters. Furthermore, colour consistently may need to concern in this visual. It may have too many colour patterns to make audience confuse. It may use one colour to show the highest value and the other one to represent for the rest.

Graph 3: Coronavirus (COVID-19) and Tobacco Use

This is an infographic, and it is coming from the official website of Word Health Organization. It mentions about the relationship between Coronavirus and tobacco use. It delivers the message to general public, especially for some teenagers in Europe. The reason is large amount of people smoke tobacco and over 10 percent of age 13-15 teenagers use it in the European region. The message is mainly concern how high risk of getting the Coronavirus if using tobacco. It also mentions why people should quit using tobacco as soon as possible.



(Source: <https://www.euro.who.int/en/health-topics/health-emergencies/coronavirus-covid-19/publications-and-technical-guidance/noncommunicable-diseases/resources-for-tobacco-use-control-as-part-of-covid-19-response/infographic-coronavirus-covid-19-and-tobacco-use>)

Evaluation

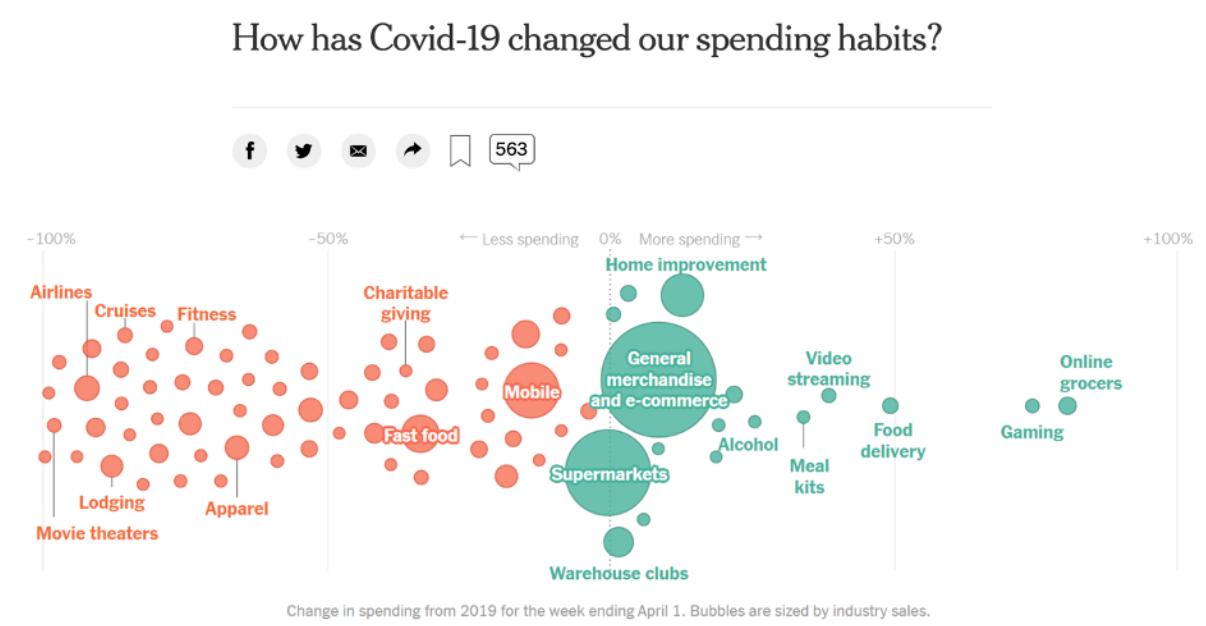
The graph uses high data-ratio as all visuals and texts are related to the main purpose. On the text section, we can find that it uses bold texts to highlight main ideas to draw audience's attention quickly, which is using a part of prettentive attributes. Moreover, the graph tries to use colour consistently and helps audience to focus on a part of population automatically. Once people read though the graph, they are not difficult to understand what message in the graph is.

Recommendations

The graph desire to illustrate how bad of smoking tobacco and hope smokers quit it as soon as possible. Suggest mentioning a main idea to draw audience attention first, that means it may indicate a main idea at the top. In addition, an outstanding title is a good way to capture audiences why people need to quit using tobacco. If a title cannot capture their attention, they may not interest to read though a story afterwards.

Graph 4: How has Covid-19 changed our spending habits?

The bubble chart is coming from the website of The New York Times, which is an American daily newspaper company. The chart illustrated about Coronavirus changes customer spending habits on variety of industries in United States. It delivered the message to commercial companies and they may need to change their sales and marketing plan because of customer spending habits has changed. Bubbles size was a sales volume of the industry. People spent more on online grocers and gaming industry, which it had over 50% increased. However, airlines and movie theatres industry were the most affected in pandemic situation.



(Source: <https://www.nytimes.com/2020/04/16/learning/whats-going-on-in-this-graph-pandemic-consumer-spending.html>)

Evaluation

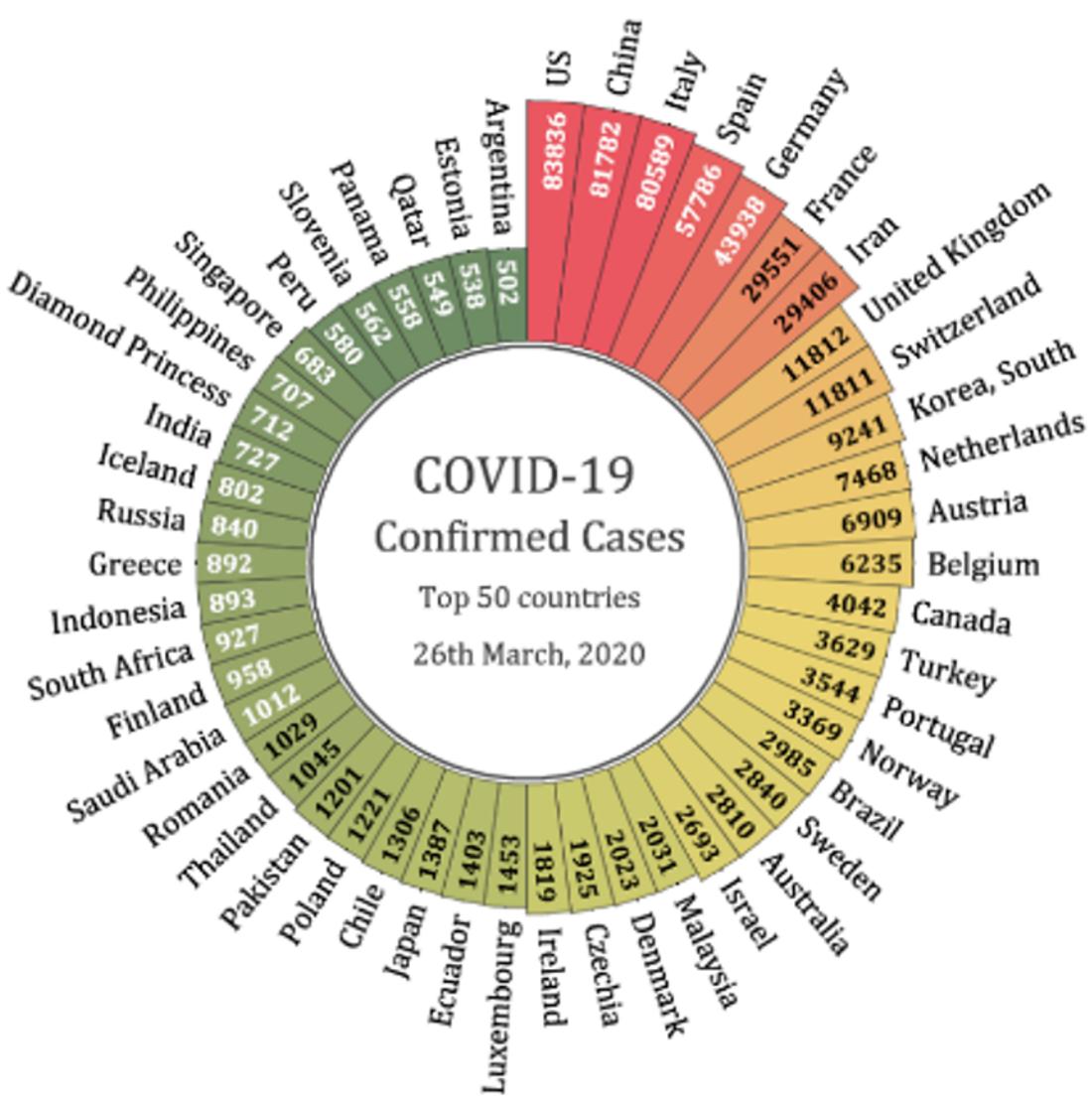
A creator does well on reducing clutter and it cannot find any unnecessary things to interrupt audience to read the data. The bubbles size makes people easy to understand about these industries sales, which is using a prettentive attributes properly. On the other hand, The Gestalt Principle of Similarity is also applied in the chart as it can find two different colours with text to distinguish the differences. The green bubbles represent people spend more and the orange one is people spending less for those industries. The audience will be easy to understand that there are two groups depends on colours.

Recommendations

The section of data-ink ratio may do it better from the chart. It has some bubbles without any labels and it seems does not have any meaning to deliver to target audience. Maybe a designer desire to let audience know that there are many industries are getting less spending from customers. However, it may use the other extra chart to describe more rather than put many bubbles in the charts to make audience confused. Suggest removing the bubbles without labels and attempt to the graph as simple as possible.

Graph 5: Top 50 countries of COVID-19 confirmed cases.

The radial bar chart is created from the website of OriginLab that is a software company in United States. In the scientific and engineering community, they have variety of graphs for people reference. It showed top 50 countries of confirmed cases of Coronavirus till 26 March, 2020. It can find out the highest record was US, China, and Italy. Those countries were the most serious countries at the mentioned period. General public should be the target audience as it desired to deliver a message to let them know the situation of top 50 countries at this moment.



(Source: <https://www.originlab.com/index.aspx?go=products/origin/graphing>)

Evaluation

The chart presents a high data-ink ratio, and it cannot find any unnecessary elements to make target audience confused. It also tries to apply preattentive attributes in the graph, which it uses different text colour to highlight the most and the least serious countries to draw audience attention. Furthermore, the gestalt principle of similarity indicates different groups of similar colours. It attempts to use gradual colour like red, yellow, and green to separate for three groups mainly. Besides, it is not complicated to let the audience read the data in a short time.

Recommendations

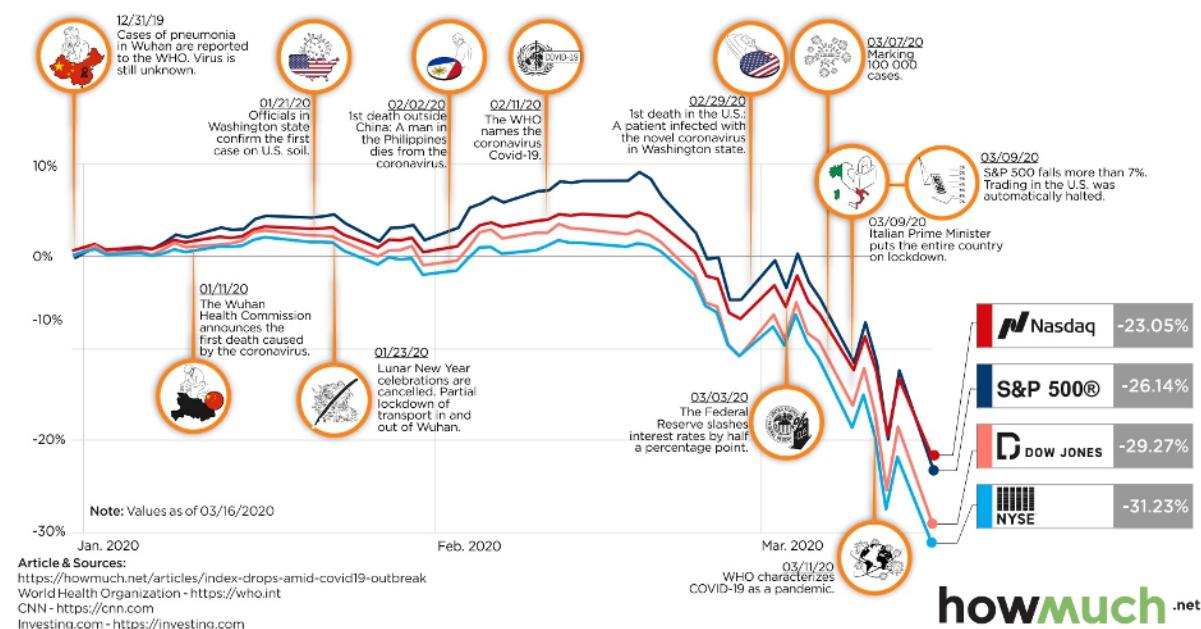
The chart may not deliver how serious in top five countries. It uses red colour to groups the most serious countries, but the bar ratio is not too obvious compare with the least serious centuries. It may use a correct ratio of confirmed cases and it can express how serious for the top five countries at that moment. In addition, the colour of the numbers may not use properly for the least serious countries. Except the top five countries, it may consider using the same colour for the rest, which avoid the target audience draw attention of this part.

Graph 6: Index Drops Amid the COVID-19 Outbreak

The information of the line chart is provided from the website of howmuch.net. The service company provides financial and cost data visualisations to help people make financial decisions. The chart illustrated the impact of Coronavirus (COVID-19) on United States stock indexes. It showed the US stock trend and started from late 2019 until Mar 2020 and it also described the Coronavirus events at the same graph. The graph may deliver the message to general public to know more about how impact of US stock by the coronavirus.

Index Drops Amid the COVID-19 Outbreak

U.S. Stock Indexes Drops since 12/31/2019 (as %)



(Source: <https://howmuch.net/articles/index-drops-amid-covid19-outbreak>)

Evaluation

The ratio of data-ink is high, which it cannot find much noise from the graph and the target audience is easy to read the data. It attempts to highlight four different colours with U.S. stock indexes to indicates about the trend during this period and it applies preattentive attributes in graph quite well. Also, The Gestalt principles of closure is used in the graph that it can avoid clutters, but it can still reflect the message to the audience. On the other hand, it can also find the information about Coronavirus though the timeline. People can have a full picture about the relationship between Coronavirus and U.S. stock indexes at the same time.

Recommendations

For the information of Coronavirus, it is also included the picture of each information. Maybe it can draw audience attention, but it is not necessary. It may distract their eyes to the picture rather the trends of U.S. stock indexes. In addition, it has a pin to point out each information of Coronavirus on the timeline which is good way to let people know what is about. However, it may try to change different colour as it may not distract people attention easily.

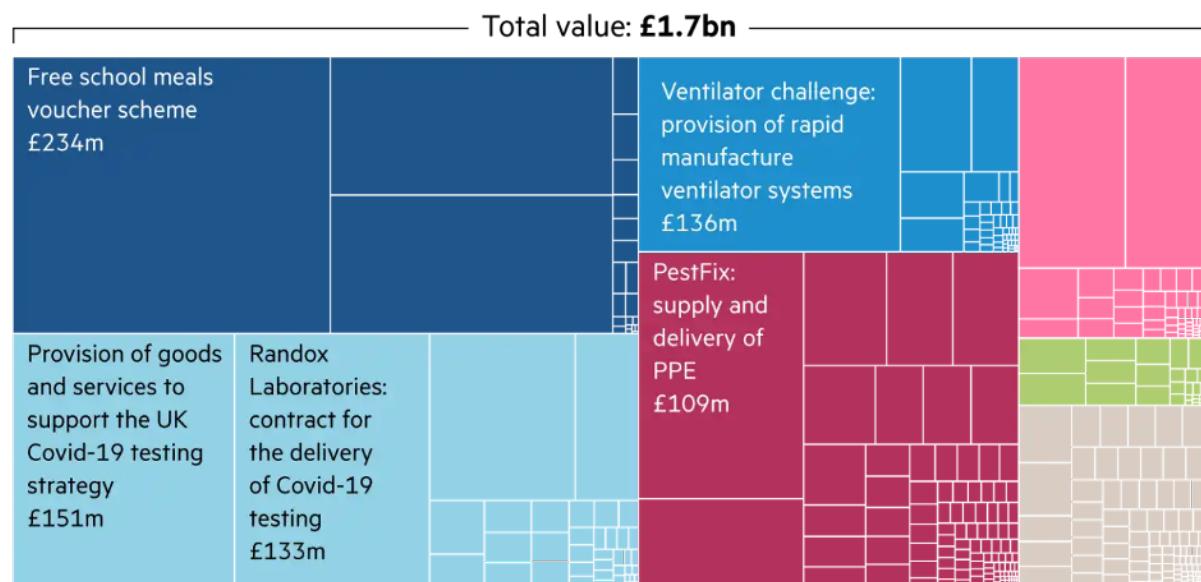
Graph 7: The UK government has awarded about 400 Covid-19 contracts

The treemap chart is sourcing form the international daily newspaper's website, Financial Times. It showed about 400 COVID-19 contracts has agreed by the United Kingdom government. It covered different sections, which were accommodation and food, testing, ventilators, PPE, IT and telecoms and consulting, etc. The graph provided the message to the general public, which UK government wanted people to know what the government spends the money for private companies because of coronavirus.

The UK government has awarded about 400 Covid-19 contracts

Contracts to private companies (£m)

█ Accommodation and food █ Testing █ Ventilators
█ PPE █ IT and telecoms █ Consulting █ Other



Source: Tussell
© FT

(Source: <https://www.ft.com/content/7fe7c2d5-24df-431b-9149-50417fa0236a>)

Evaluation

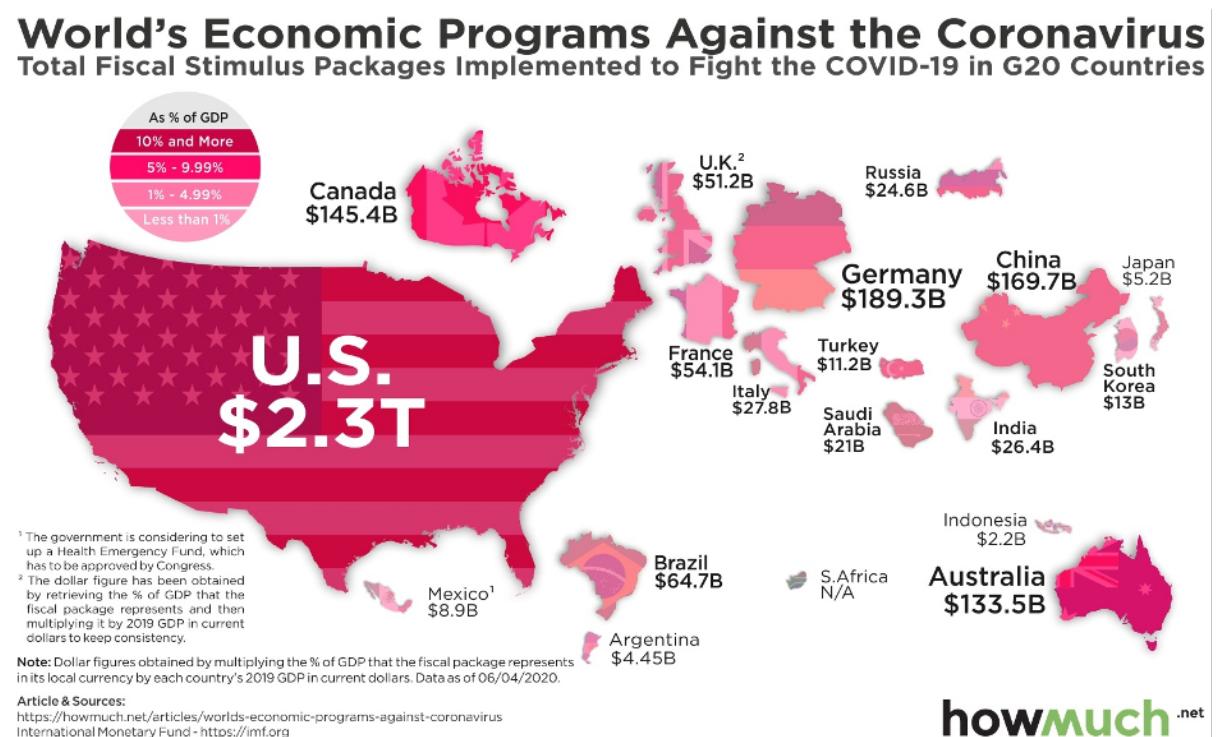
The chart may not have high data-ink ratio. It uses different colours to separate each section, but it also draws some lines to divides the own sections. However, it does not have any meaning of those. Also, the colour of different sections is similar, the target audience may not have found any difference and it cannot draw their attention by colours. The Gestalt principles of enclosure used in the chart and it is good to help an audience to know the size of each sections. Furthermore, the chart does not overcomplicate to read the data including text and people can understand the main idea in short time.

Recommendations

The creator may improve preattentive attributes better. Suggest using contrast colours for each section and it may help the target audience to clarify different sections easier. On the other hand, it may reduce clutter for each section. It should avoid drawing some lines to divides the own sections because it cannot deliver any message to audiences and just make it more complicated.

Graph 8: World's Economic Programs Against the Coronavirus

The information of the choropleth map is provided from the website of howmuch.net. The service company provides financial and cost data visualisations to help people make financial decisions. The map showed the main countries of using budget for helping economic programs against the coronavirus. It can see how much they implemented, and the colour represented the percentage for each country. It hopes to show to the general public that every government spends huge amount of money because of the Coronavirus pandemic and specially Unities States spending over two trillion dollars for the stimulus package.



(Source: <https://howmuch.net/articles/worlds-economic-programs-against-coronavirus>)

Evaluation

High data-ratio is applied for the map that it cannot find any redundant ink to represent information. It does not have any clutter to make it complicated to read and the target audiences read the information easily. In addition, preattentive attributes of size doing well, which it uses the size of image and text to perform how much the government they spend. The audience will be easy to know United States spending the largest amount of package compare with other countries. Overall the map is not overcomplicated, and it can deliver the message to audience directly.

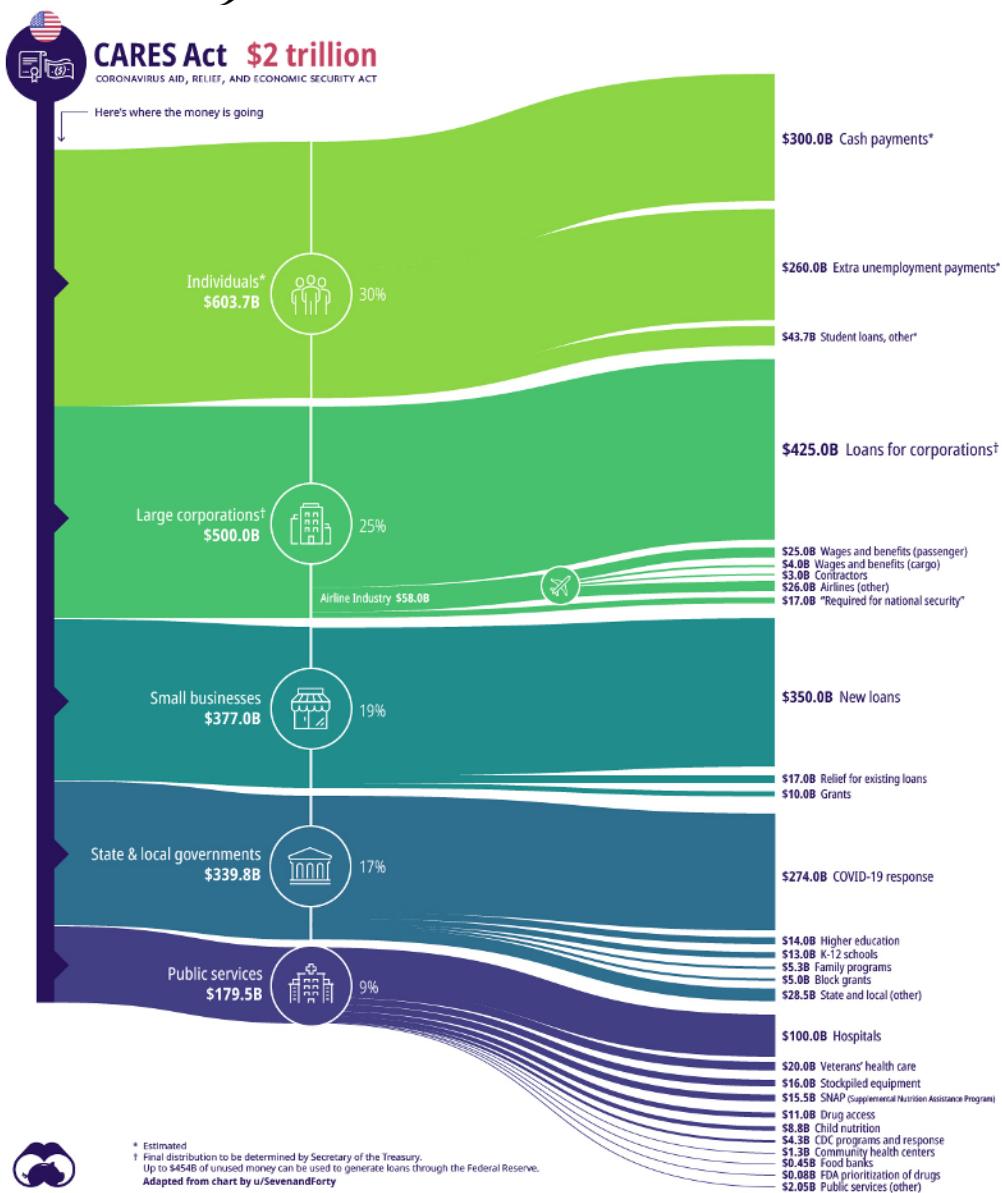
Recommendations

It may do better for using colour. On the top left-hand corner, it shows the percentage of GDP by gradual colour change. Honestly, it cannot see any differences except United Estates, which may not be easy to deliver the message to audience. If the editor desire to mention about United States only that may use only one colour tone for United Estates, the other colour tone for the rest. Or it may change less options of a criteria of percentage of GDP which it can deliver a message to audiences by colour tone clearly.

Graph 9: The Anatomy of the \$2 Trillion COVID-19 Stimulus Bill

The information of a sankey chart is coming from the website of Visual Capitalist. This is an online publisher and it focuses on the topics of markets, technology, energy, and global economy. The chart illustrates the breakdown of the United States government how to use two trillion dollars for a stimulus package because of the Coronavirus pandemic. It can deliver a message to general public about the breakdown of each section, which was like individuals, large corporations, small business, state & local governments, and public services.

The Anatomy of the \$2 Trillion COVID-19 Stimulus Bill



(Source: <https://www.visualcapitalist.com/the-anatomy-of-the-2-trillion-covid-19-stimulus-bill/>)

Evaluation

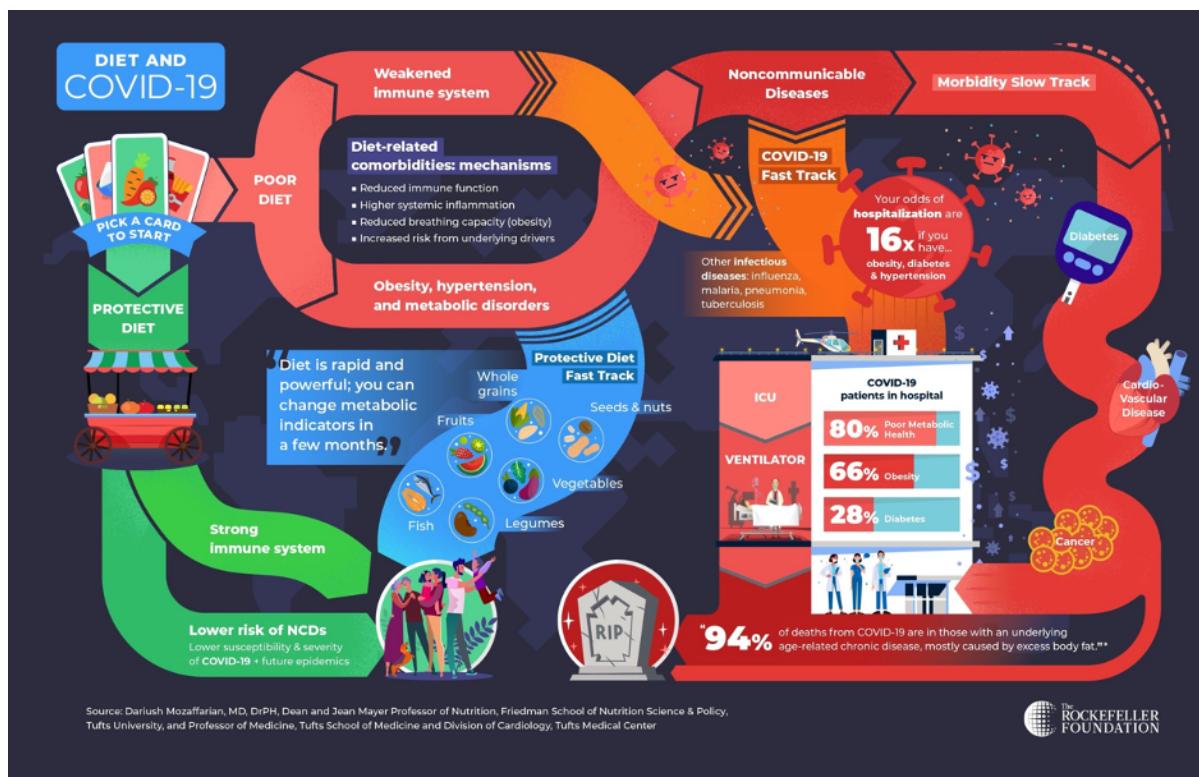
The ratio of data-ink is quite high that it is difficult of find any non-data-ink to distract the target audience. It attempts to use preattentive attributes of colour to highlight the numbers to spend for each section and even breakdown of each numbers for details. It can be easy to draw audience's attention about how much United States government spending on. The Gestalt principles of proximity is used in the chart, which is using the different colours and it contains a suitable distance to separate for each section. The audience is not difficult to follow the flows to read the data without any distraction.

Recommendations

For the sankey chart, it shows the different images of each sections to draw target audience. However, it may try to divert them to focus the text information on the left side. Also, the percentage of each sections show next to the image. It may not be compulsory on the chart. The season is the size of thick is also illustrated, which it can explain about the proportion of each sections. Suggest to cross-out these two elements and attempt to make a chart simpler and reduce clutters.

Graph 10: The Two Fast Tracks of COVID-19

The infographic is sourcing from the website of Rockefeller Foundation. The organisation is an American private foundation based in New York City. It shows two tracks of results if people have protective diet or not. It also delivers the message to public and the importance of the relationship of diet. Then it shows the path how to stay away the Coronavirus.



(Source: <https://www.rockefellerfoundation.org/blog/the-two-fast-tracks-of-covid-19/>)

Evaluation

The infographic illustrates two tracks of results if people have protective diet or not. It is easy to understand if people have protective diet. In addition, it attempts to apply preattentive attributes of size, which uses bold and enlarge some key words and numbers to emphasize the message to draw audience's attention. Gestalt principles of similarity is also applied, that using green and red colour to separate two groups and instruct the target audience to read the information. However, people have to take long time to read the information of result if they have poor diet. It may use redundant-ink for this track because of too many images and the flow of poor diet is bit complicated on the graph.

Recommendations

For the infographic, it has two paths if people have a poor diet. After the option of noncommunicable diseases, there are two paths again. One is COVID-19 Fast Track and the other is Morbidity Slow Track. However, the option of Morbidity Slow Track has not strong relationship with Coronavirus. It may distract the target audience to read the message from the graph. Moreover, the editor desire to provide too much information in a graph. It will make audience does not have enough patients to read through all the data in the graph. Highly suggest reducing some images and part of information, which reduce clutters to make the graph is not too complicated.

Summary

From this report, we create a list of quality criteria to evaluate and suggest in ten graphs of related to Coronavirus. In those graphs, some graphs can be fulfilled the quality criteria and some graphs need to be improvement. After finished the report, we can learn how to clarify the good quality of graph which based on some principles and rules. The key points are: 1) how to draw audience's attention; 2) Do not provide too many information and many kinds of colour in the graph to distract audience; 3) Keep it simple and easy to read.

Basically, the quality criteria are covered wide range of principles. Most of editors will follow the similar principles to justify and criticise a graph when they apply for data visualization. Believe that it is good enough and use the same quality criteria to apply on element two.

Will use this element's quality criteria to find out the problems and try to make it better for the graphs. Even we can cross check our team's graphs and improve everyone quality and discuss how can we make it perfect before submitting ICA.

Reference

Knafllic, C. (2015). Storytelling with Data: A data visualization guide for business professionals. New Jersey: John Wiley & Sons, Inc., pp.19-99.

TUFTE, E. R. (2001). The visual display of quantitative information. Cheshire, Conn, Graphics Press.

Ware, C 2012, Information Visualization : Perception for Design, Elsevier Science & Technology, San Francisco.

Nussbaumer, KC 2015, Storytelling with Data : A Data Visualization Guide for Business Professionals, John Wiley & Sons, Incorporated, Hoboken.

William, C 1994, The elements of graphing data. Murray Hill, N.J., AT & T Bell Laboratories.

Yuk, M, & Diamond, S, 2014, Data Visualization for Dummies, John Wiley & Sons, Incorporated, Somerset.

Friedenberg, J 2012, Visual Attention and Consciousness, Taylor & Francis Group, London.

Gordon, IE 2004, Theories of Visual Perception, Taylor & Francis Group, Hove.