

Title: Pandemics: Data Evolution & Visualization

Srigouri Vasuki Bikkina, Bayer, Hyderabad, India

Poojitha Mandava, Bayer, Hyderabad, India

ABSTRACT

With the evolution of the mankind through the various stages of technological growth, it has encountered rampant pandemics starting with Antonine plague to the present day COVID 19. Throughout this journey, there has been a rise of disparate data sources and analyzing tools with advent of various programming languages. The trend of predicting the factors and consequences of a disease to forecast and forewarn has become the norm.

In our research, we are trying to incorporate the copious amount of data amassed into Heat maps to represent the spread of the COVID 19 through the different climate zones and incidence rates of the disease. Additionally, we are trying to present the impact of the factors influencing the disease and its effect on human growth and development visually and picturizing the raise of the data analytics through time.

Key words: COVID 19, data visualization, pandemic.

INTRODUCTION

As human population evolved over ages, they had suffered from many Pandemics in the history starting from Plague to COVID 19 that created catastrophic damage to the population. In other way there exists a direct proportionality between the civilization of humans and pandemics^[1].

The first instance of using the data to predict & analyze the number of cases that can be impacted in a pandemic happened in 1854 during a cholera outbreak by Dr. John snow, who concluded the spread was due to transmitted water and also displayed the mortality data visually on a graph^[2].

PANDEMICS

An illness is a disease or a state of being sick that impacts body or mind. The illness sometimes might be contagious and hence is termed as an **outbreak**, when it effects larger population that last for several days or years, having an unknown cause and which needs time to get it identified or treated. If this outbreak spreads across in a region and impacts more lives than expected it is said to be **epidemic**.

A **Pandemic** is different from epidemic where the former spreads across the countries, continents and disrupts larger population than epidemic.

DATA AND ITS EVOLUTION:

A set of values [qualitative or quantitative variables] which could be a symbol, number, word, character. This can be an information when it is organized, structured and processed in a context. This data is very important & useful as it helps in decision making, improve QOL of people, predictions and simulations. The evolution happened in various stages.

- The patient data was hand-written earlier and even the test reports were handwritten. With the evolution of data, the entire medical history of one patient/person is documented and stored electronically.
- Earlier data was collected in form of tally or tick marks which were collected to track use of inventory or food. Later Abacus was invented to help with those calculations. Even the astrological studies and their data resulted in scientific discoveries. Thus, even in the earliest of times, many tools were discovered to collect and analyze data. One example for this data evolution is the census-a large source of data that started in early 1800.^[3]
- Hollerith Tabulating Machine was thus invented by 'Herman Hollerith' who understood the need for collection and analysis of census that can tabulate data from census & analysis it from years to months.
- Later with the evolution of computer in 1900's there was a point of storage crisis. 'Howard Dresner' an analyst at Gartner with his discovery of magnetic disk storages by 1980's solved it. With internet discovery, the revolution reached its peak.

At the end, it is about how the data is collected and analyzed and this needs faster & cheaper ways to analyze and store it.

DATA ANALYSIS

It is a process of cleaning, transforming and modeling data to extract useful information from data & take the correct decision. The data analysis can be categorized^[4] into:

PhUSE EU Connect 2020

1. Descriptive Analysis: describes the data statistically -what happened or what is it about
2. Diagnostic Analysis: helps us to take a deep insight into the data and in root-cause analysis
3. Predictive Analysis: this gives us the prediction of what will/might happen in the future with the retrospective data present.
4. Prescriptive Analysis: this analysis will take the predicted data to further level by suggesting what action can be taken
5. Cognitive Analysis: Cognitive analytics combines a number of intelligent technologies like artificial intelligence, machine-learning algorithms, deep learning etc.to apply human brain like intelligence to perform certain tasks.

DATA VISUALIZATION

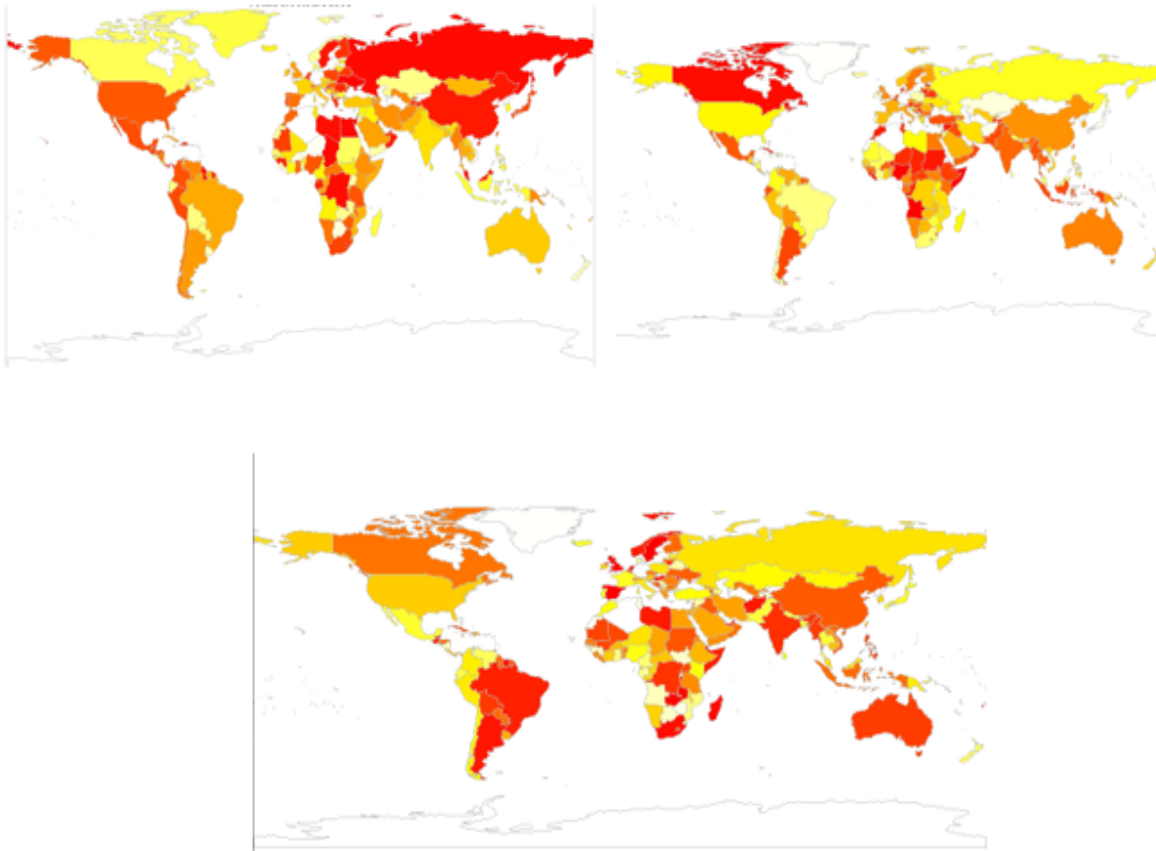
The common saying "Use a picture. It's worth a thousand words," which first appeared in a 1911 newspaper article, is often used without much thought regarding its meaning or origin ^[5]. Data visualization is the graphical or pictorial presentation of data form which helps the reader to understand the information easily. In other terms, with the massive data that we have collected it will be difficult to analyze it or obtain required information and hence we often try to display the results in the form of charts, graphs etc.

One of the first instances of relying on geography and statistical analysis and visualization was in mid-19th century London, during a cholera outbreak. In 1854, Dr. John Snow came to the conclusion that cholera was spreading via tainted water and decided to display neighborhood mortality data directly on a map- dot map. He also illustrated the hypothesis of quality of water and cholera cases using statistics. Data visualization is concerned with the design, development and application of computer-generated graphical representation of data. Irrespective of the domain or industry every field benefit from understanding data like in government, finance, marketing, history, consumer goods, service industries, education, sports, and so on.

HEAT MAPS

Heat maps is a data visualization technique which shows the magnitude of an event or occurrence through the intensity of the color. The COVID 19 data for the heatmaps was taken for the WHO statistics. Using R packages; a) rworldmap, b) ggplot2, the below the heatmaps are created for the data :

1. total number of affected cases (figure at the top left)
2. total number of deaths (figure at the top right)
3. total number of active cases (figure at the bottom middle)



PhUSE EU Connect 2020

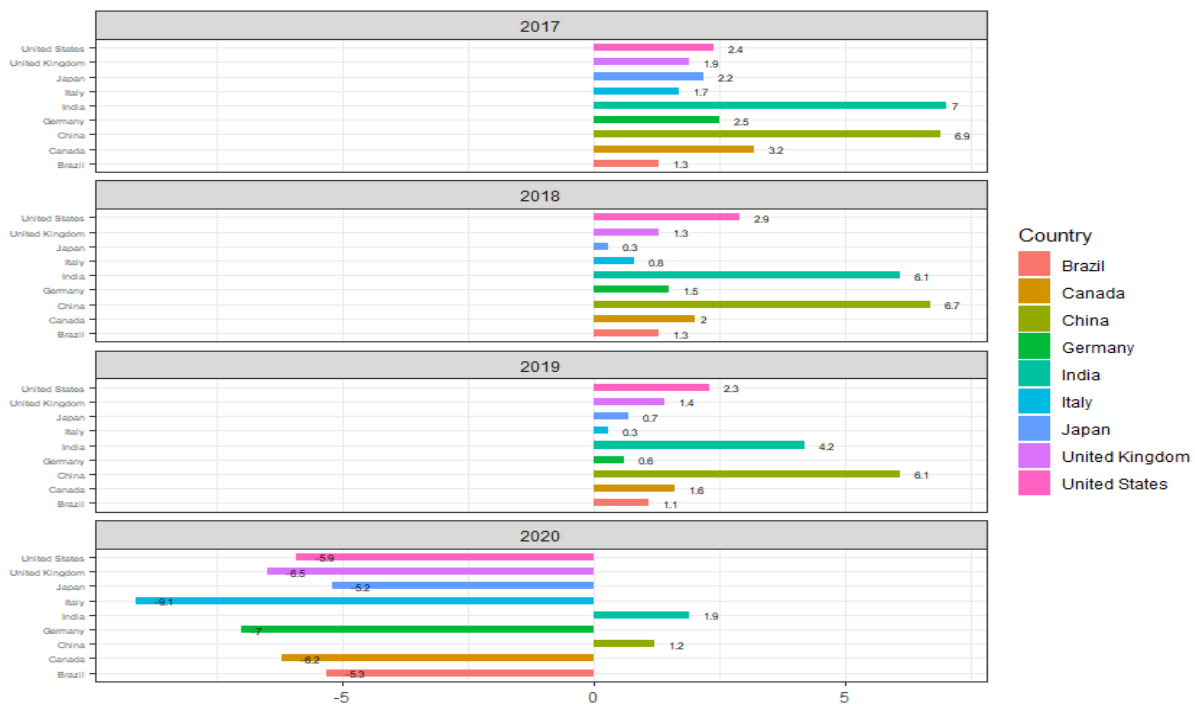
The following observations can be made from the heatmaps:

- The countries like USA, which have been more affected, have a darker shade of color then the countries with less effected cases.
- In USA there is also a major difference in the total affected cases and active cases, indicating recovery of many people which can be correlated to better medical services or better recovery rate
- Few countries with lesser travelers, have lower number of cases

GDP GROWTH RATE (ANNUAL PERCENT CHANGE) PLOT:

This below plot shows the effect of COVID 19 on GDP rate of few of the countries. As you can see when compared to last three years the GDP has massively declined in the year 2020. This is mainly by all the sectors like industrial, entertainment, tourism, IT etc. being affected by the COVID and also the lockdown imposed.

The GDP has taken a turn to the negative scale and it will take an extended time period for the affect to decline and the GDP's to return to the pre-COVID state.

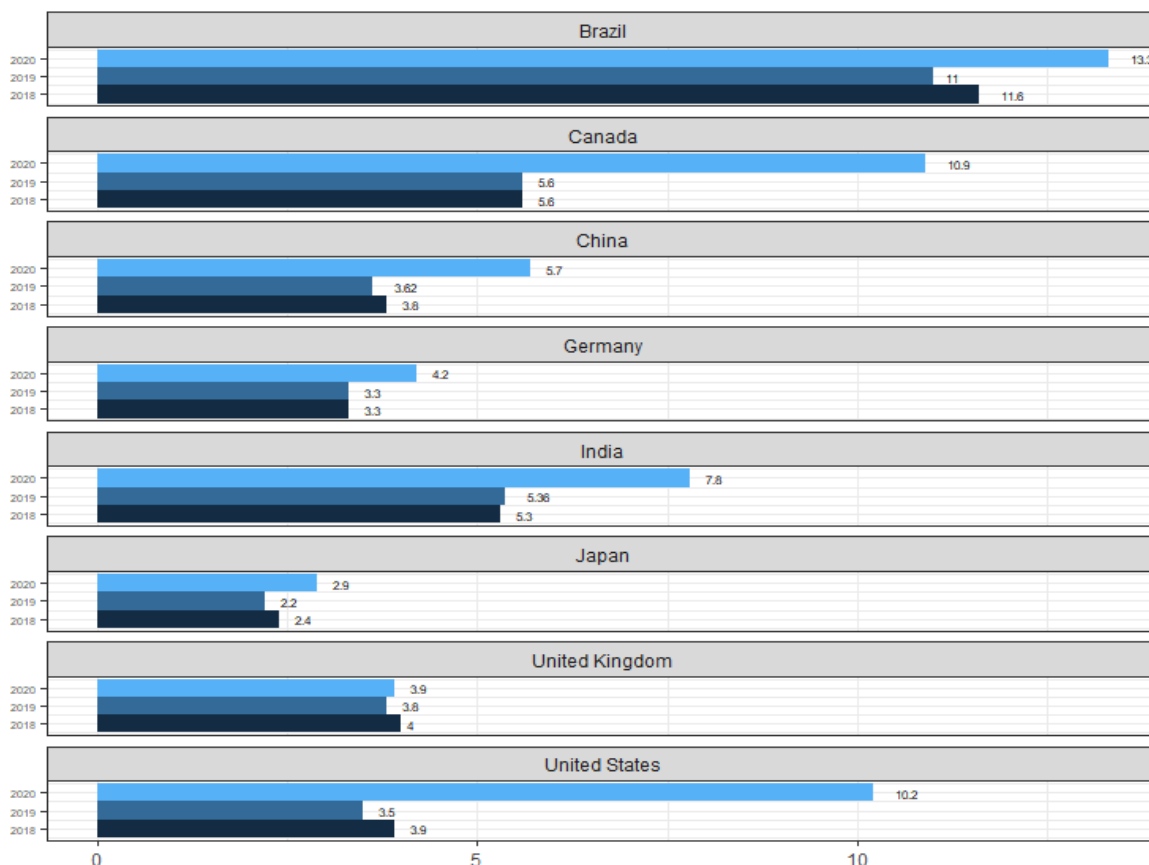


UNEMPLOYMENT RATE STATISTICS VISULIZATION

As all the sectors have taken a critical blow, the organizations are laying off huge number of employees as strategic approach. This increase in the unemployment rate at a steep incline would in turn adversely affect the peoples mental and financial stability and lead to unrest in the country.

In the below graph we can witness that in USA the unemployment rate has jumped from a 3.5 percent to 10.2 percent. The end result is again the decline in the medical services available to the unemployed people leading to higher COVID 19 cases.

PhUSE EU Connect 2020



Pandemic has shook the world and has changed the perspective of the human race to newer norm in coexisting with a deadly virus.

CONCLUSION

With boom of the data analysis, visualization techniques and massive data accumulating every day, the objectives like prediction of the outcomes of events, management of patient info, inferences of the retrospective data for future occurrences are swiftly becoming the focus of the world. This might lead to future predominated of predictions of the natural disasters with upcoming data visualization procedures.

REFERENCES

1. <https://www.weforum.org/agenda/2020/03/a-visual-history-of-pandemics>
2. <https://www.nytimes.com/interactive/2020/06/10/magazine/covid-data.html>
3. <https://www.gutcheckit.com/blog/a-history-of-data/>
4. <https://www.logianalytics.com/predictive-analytics/comparing-descriptive-predictive-prescriptive-and-diagnostic-analytics/>
5. <https://bdataanalytics.biomedcentral.com/articles/10.1186/s41044-020-00047-z>
6. https://www.researchgate.net/publication/311597028_DATA_VISUALIZATION
7. <https://www.imf.org>
8. <https://www.worldometers.info>
9. <https://countryeconomy.com>

ACKNOWLEDGMENTS

I would like to thank my manager N Jayapandian for encouraging us to write the paper and our colleague K. Harsha for helping with analysis of the data.

DISCLAIMER

The content of this paper is the work of the authors and do not necessarily represent the opinions, recommendations, or practices of Bayer Pharmaceuticals Pvt. Ltd

PhUSE EU Connect 2020

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author or co-author at:

Author Name : Bikkina Sri Gouri Vasuki
Co-author Name : Poojitha Mandava
Company : Bayer Pharmaceuticals Pvt Ltd.
Address : Ohris Tech Park, VBIT Park Road, Madhapur
City / Postcode : Hyderabad, Telangana 500081
Work Phone : +91 9052390343, +91 8790000898
Email : vasuki.sri02@gmail.com , poojithamandava9@gmail.com
Web : <http://www.bayer.in>

Brand and product names are trademarks of their respective companies.