

WESTERN SYDNEY UNIVERSITY



Visualisation

Rabina Padhy – 19884657

Assignment -2

Multidimensional Data Visualisation

Declaration

I hold a copy of this assignment that I can produce if the original is lost or damaged.

I hereby certify that no part of this assignment/product has been copied from any other student's work or from any other source except where due acknowledgement is made in the assignment.

No part of this assignment/product has been written/produced for me by another person except where such collaboration has been authorised by the subject lecturer/tutor concerned.

Table Contents

Introduction	4
Tableau	4
Advantages and Disadvantages.....	4
R Statistics	5
Advantages and Disadvantages.....	5
Dataset using Multiple Techniques	5
Insight about Covid-19 Data Set.....	5
Tableau Visualisation Technique.....	5
Map Chart	5
Technical Details.....	5
Bar Chart	7
Technical Details.....	7
Line Chart.....	11
Technical Details.....	11
Pie Chart.....	15
Technical Details.....	15
Tree map	16
Technical Details.....	16
Story and Dashboard	17
Technical Details.....	17
R Statistic Visualisation Technique.....	18
Map Chart	18
Technical Details.....	18
Bar Chart	19
Technical Details.....	19
Line Chart.....	20
Technical Details.....	20
Comparison Analysis on Different Tools.....	21
Conclusion.....	21
Reference.....	22

Introduction

High quality of documents received from multiple sources and we need intuitive visualisation tools to explode data efficiently. In 1990's, basic data visualisation used to understand the data. But the time has changed, now people use multiple visualisation methods and tools in data visualisation to get the desired outcomes. The traditional techniques were shaped to some extent however, more enhanced technologies are available. All the latest technologies are very interactive and high configuration so that it easily extracts the hidden pattern of the dataset (Martin and Marinescu, 1998).

Using visualisation method, we can achieve multiple goals and benefits (Tejada, Minghim and Nonato, 2003).

- ❖ User's creative thinking are enhanced by use of visualisation in presenting the datasets in a meaningful manner.
- ❖ This programming technique helps to readily enhance perception and recognise the resources.
- ❖ Visualisation will give an idea about law of attraction, surrounding circumstances and boundary forces to do a graph.
- ❖ To achieve organisation goals, visualisation helps to build your internal motivation.

We need visualisation to discover hidden patterns and analyse multi-dimensional datasets. Enterprise users generate data from different source systems in various formats such as .txt, .xlsx, .csv, .xml. Presenting the data with custom analytics is a major challenge. The approach of visualisation helps to signify the data in numerous manner such as bar chart, line chart, pie chart, images and other natural display ways.

This paper will discuss the technical details, their advantages and disadvantages of multiple visualisation technique. This could be achieved using Tableau and R tool to draw different charts such as Bar-charts, Pie-charts, World Map and Line charts. Tableau and R Statistics tools are used to analyse the multi-dimensional dataset for better visibility.

Tableau

Tableau is the swiftest-growing data visualization tool which helps one convert textual and numerical information to interactive and beautiful dashboards. By using tableau, we can effortlessly transform raw data into simple understandable data. It is a powerful business intelligence tool and widely used in all small and big organisation to get better insight by doing analytics on multi-dimensional datasets. The main advantages of tableau is, it doesn't require any kind of technical or programming skills.

Advantages and Disadvantages

Tableau provide benefits as high performance with large multi-dimensional dataset. Tableau is very user friendly and platform independent. Tableau also gives extensive quality customer service and easy to upgrade. User can write SQL query and multiple condition in a dataset for better analytics. By using Tableau, we can analyse the trends and forecast the dataset with time (Visualization, Projects, Kumar and Kumar, 2020). However, there are heavy cost for licenced product at an enterprise level. Tableau has no automatic refresh of reports and not available for 3D representation of dataset.

R Statistics

R is most popular programming language for statistical modelling and analysis. By using R statistics, we can predict future and also, forecast data.

Advantages and Disadvantages

R statistics is open source and data wrangling. By using R, we can draw quality 3D plotting and graphics design. R helps to achieve various machine learning operations such as classification and regression. It has many in-built packages which are used to represent data in a graphical manner (R Code- java point, 2020). However, R also have disadvantages such as its security model is quite preliminary and non-enterprise compatible. It's highly complicated to understand the code behind the model. If we use large dimensional datasets, then it provides lower speed in R as compare to others tool. This speed is reduced since all the data is stored in memory of R Statistical tool. Although accuracy is high in R Statistical tool but fetching that information is complicated process.

Dataset using Multiple Techniques

Insight about Covid-19 Data Set

The most discussed about topic in the news, is the Novel Corona virus effect on today's world and aftereffects. Corona virus effect started in late 2019 at Wuhan and till now it continues affecting many countries. This covid-19 situation also reflects political tension and economic crisis in various countries. The following abstract is to understand more about Covid-19 affected counties and how it spread out globally within short span of time. Let's do analysis using Tableau and R Statistic tool to predict the future by using 13,286 records in a dataset (Coronavirus Source Data, 2020). We took large multi-dimensional covid dataset which include attributes as location, date, total_cases, new_cases, new_death_cases, total_death_cases, total_case_per_million, new_cases_per_million, total_test, death_cases_per_million, new_test, total_test_per_thousand and new_tests_per_thousand.

Tableau Visualisation Technique

Map Chart

We can build multiple types of maps for our geographical analysis using Tableau. To build maps, we do have some precondition viz. dataset should have at least single attribute as location, country or region. That will help to generate longitude and latitude automatically.

Technical Details

- ❖ We have location in our dataset, so navigate to worksheet.
- ❖ Find out location is named as dimension, so right click and select geographic role then select country/region option.

- ❖ Then it will automatically generate longitude and latitude. Drag and drop longitude in column shelf and latitude in row shelf.
- ❖ From the Show Me menu, select appropriate maps.
- ❖ Drag and drop the total_cases_per_million data into color option in the Mark card and similarly drag location into details option.
- ❖ To edit the color, select the color option then edit, select the appropriate one. Try to use the advance option as there one can select start, end and centre points on colouring.
- ❖ To show the labels with count of cases, drag the attribute and place it in Label option.

Covid Case Count Per Million

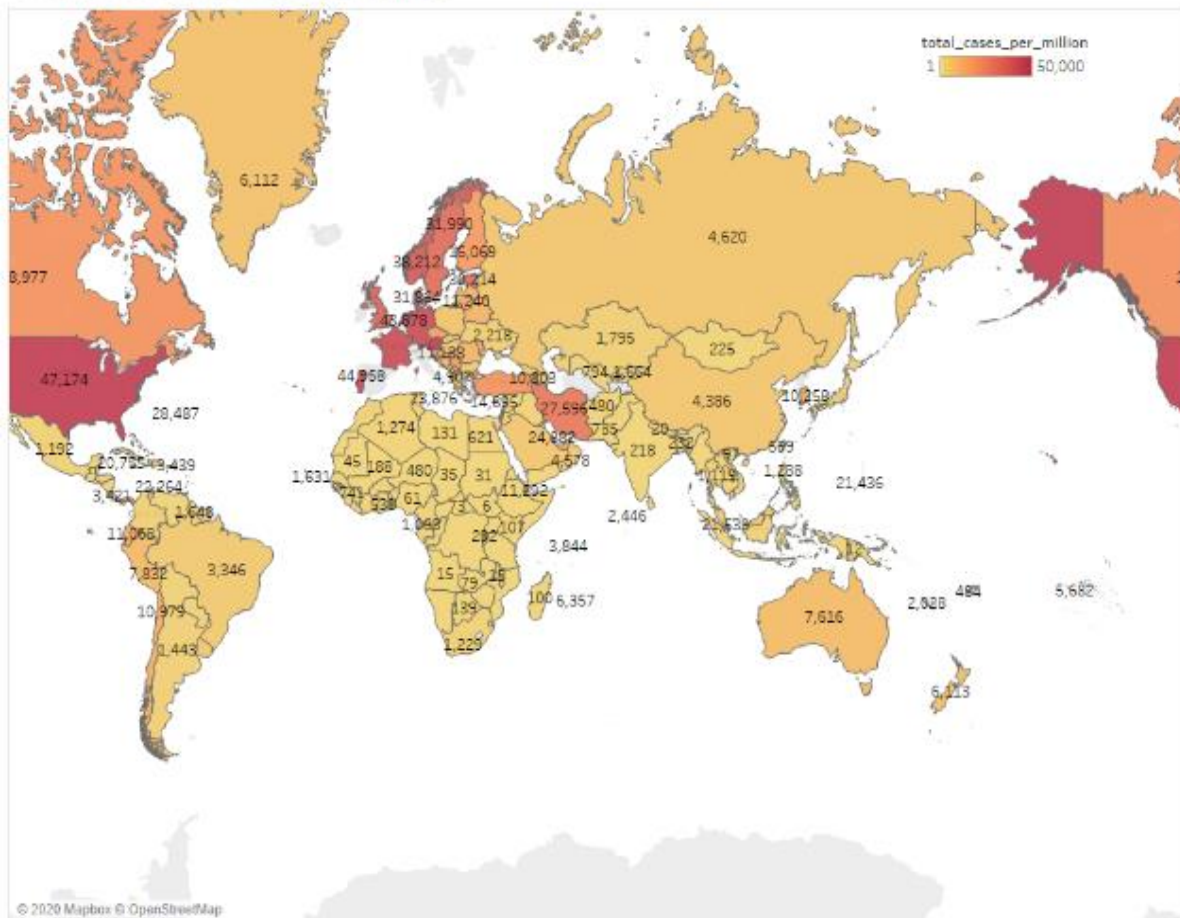


Figure: 1 World Map labelling Covid Cases Per Millions

From the figure1, we can easily conclude that United States has 47,174 covid cases per millions register until April. United states have highest amount of cases, followed by Germany with 42861 total cases per millions. By April-20 some countries had feeble cases registered. The color code Red and Orange shows highly impacted areas whereas Yellow and Mustered Yellow color code shows less impacted areas in the world map.

Bar Chart

To compare data across categories, it is better to use bar charts. We can visualise bar chart by placing dimensions on rows shelves and measures into the columns shelves by selecting bar chart option from Show me menu. We can do many modifications in a bar chart like dual axis with different data, combined bar chart on date wise and multiple view bar charts.

Technical Details

- ❖ From date attribute create two new calculative fields as Year and Week No in the worksheet using Week and Year function in Tableau.
- ❖ Combine this two new field and get single field in the worksheet and add as per condition like $\text{Year}([\text{Date}]) - \text{Week}([\text{Date}])$ in the tableau tool.
- ❖ Select the week and year combined field in columns and new_cases in rows shelves.
- ❖ From Marks card, select week as color and new cases as label. We can change the alignment Label by selecting edit of Labels.
- ❖ From the Analytics section, select the reference line, add an Average line of total_new_cases registered entire world.
- ❖ We are excluding International and World data from the location. Remove Week 18 from the week numbers. Refer figure 2.
- ❖ Now considering sum (new_deaths) and sum(new_cases) in row shelves.
- ❖ We can see in the Marks card, have 3 different tabs. 1st represent ALL, 2nd for sum (new_deaths) and 3rd for sum (new_cases).
- ❖ Change new_cases to bar chart, new_deaths to line chart. Add the labels as per respective fields and change color as per Year & Week No field. Refer to figure 3.
- ❖ Right click on the axis and select Dual Axis option. So that I will merge the both bar chart and line chart.
- ❖ Left-hand side scale represents total of new cases registered and right-hand side represents as total of deaths case due to covid in the whole world. Refer figure 4 for better visualisation. This visualisation helps to compare the row shelves.
- ❖ We also visualise data in stack bar chart where top 5 with originated counties are listed. We are plotting graph against with sum(new_cases) register per week. Refer figure 5.

New Cases Registered Per Week

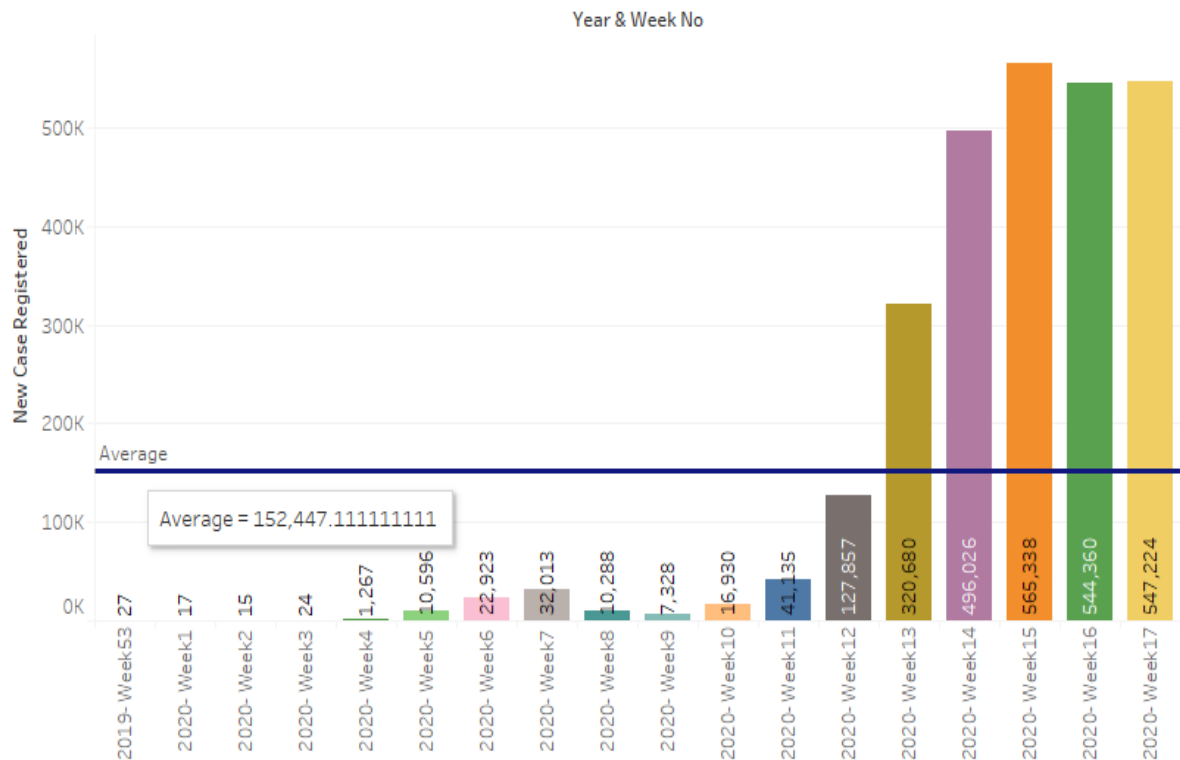


Figure 2 Week wise New Cases Registers

From figure 2, We observe a reference line also known as Average amount of new cases register i.e. (count = 1,52,447.11) in the world from 2019 to till April 2020. This average value is showing less value as compare to today's record because it counts from Dec 2019 where very few cases registered till April 2020 where more cases registered. This bar chart also represents in every week, how the new cases counts are increasing. However, from Week 15, counts are slightly decreasing. This decrease in new cases count also say that there is an awareness in public regarding the Nobel Corona virus. Every country is declaring their lockdown periods. People started maintaining social distancing. This helps to decrease the counts further worldwide. Although it decreases but not in a margin rate.

New and Death Cases Count

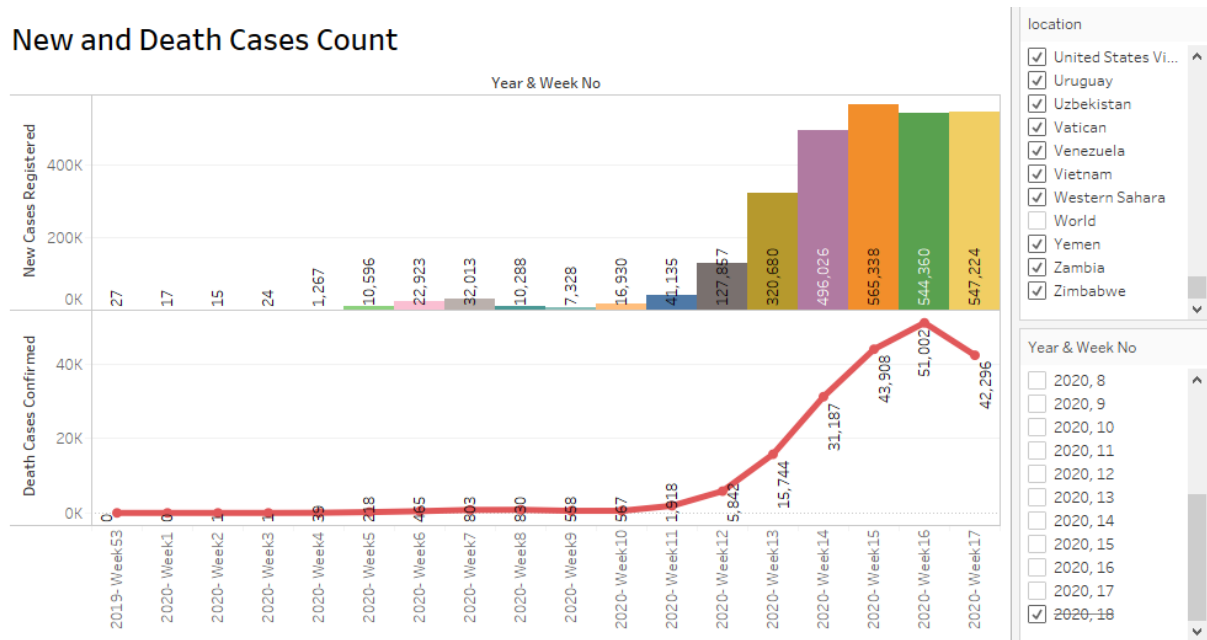


Figure 3 Multiple View of New cases and Death Cases Counts

New and Death Cases Count

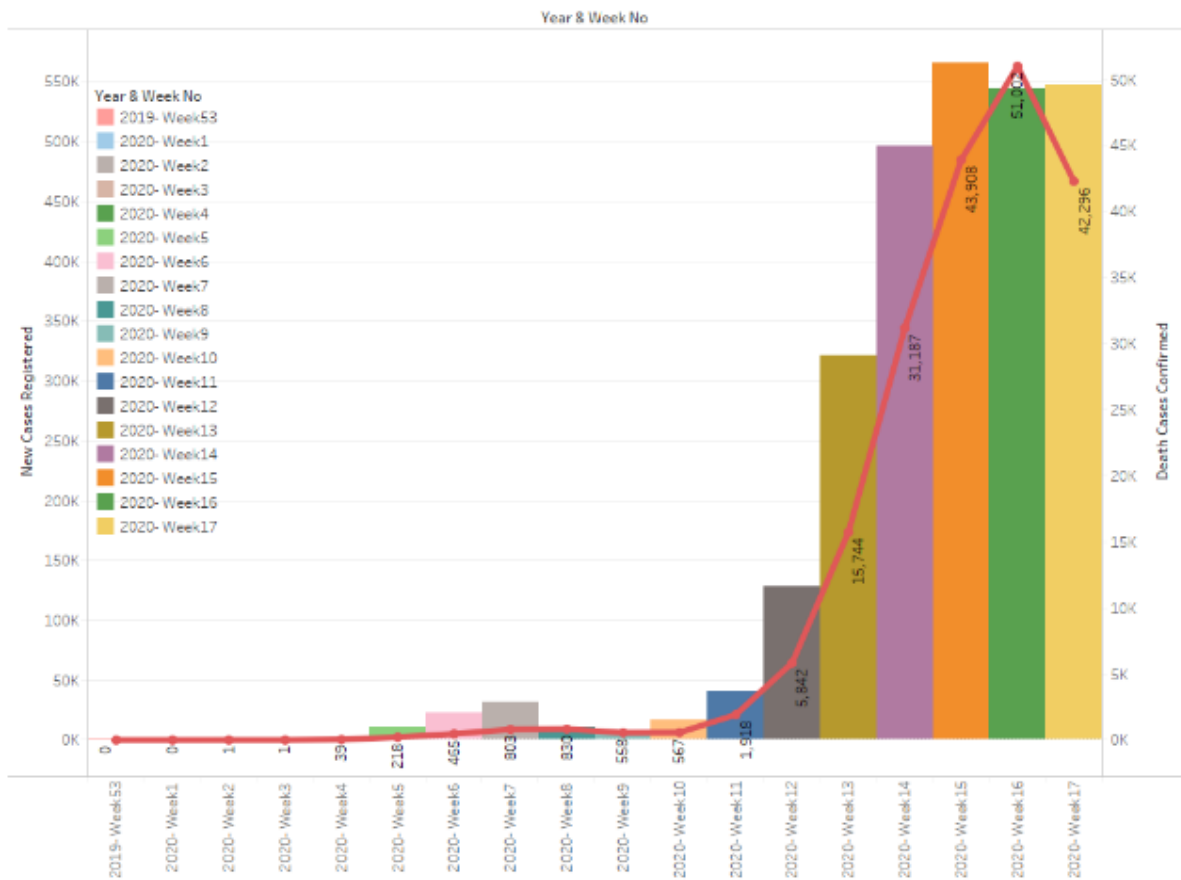


Figure 4 Total New and Death Cases in Dual axis Graph

In the figure 3, we can see multiple views with two different types of charts. The first panel representing new cases register in every week due to covid-19. We selected bar chart to visualise the data because it has categorical data i.e. Year & Week combined Field and observe that there is slight decrease in new cases from Week 15 onwards. However, second panel represents the line chart between Weeks numbers verses total death cases confirmed in the world. Till Week 16 the death counts are increasing rapidly due to Nobel Corona virus whereas in the Week 17 there is slight decrease in the death cases. This death count can be controlled if we spread early awareness about Corona virus in the public platforms. As explained earlier lockdown in every country brought down the new cases count. We are showing one the Tableau feature by selecting dual axis on death counts resulted figure 4. If we observe clearly in figure 4, we can conclude that the proportional ratio of total of new cases to total of death cases is approximately 7.5% per week wise.

Top 6 Countries Covid Cases Registered

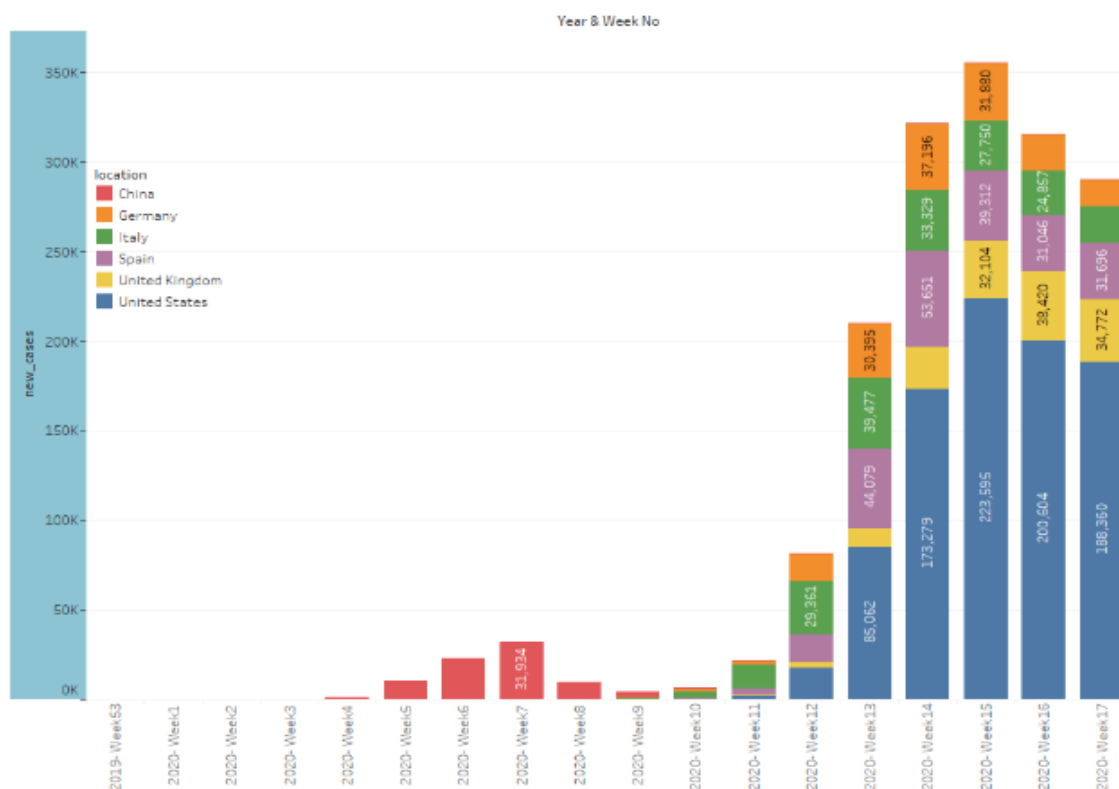


Figure 5 Comparing Top 6 countries COVID cases with each other

We are listing top six countries such as (China, Germany, Italy, Spain, United Kingdom and United States) officially registered Corona virus cases during each week. In the figure 5 showing data as a stacked bar, we can evidently see that, Nobel Corona virus originated from China. China faced highest Corona virus cases during Week 7 whereas there are no cases registered in any top countries in those weeks. From Week 12 onwards, there is a hasty growth of cases in all other countries except China. That means due to lack of information sharing between countries, people started traveling from one place to another. Due to which United states confronted highest cases during Week 15 i.e. 2,23,595. After that it is decreasing but not in a marginal rate. In Germany, Italy and Spain there are increase in new cases count until Week 14 after that it started decreasing the new case counts. From Week 10 to 17, there are very minimum cases registered in China. So, from the graph we can conclude that, if China handled this virus

seriously in the initial weeks and locked down the country for specific period then this virus would not be spread in other countries. From Week 8 onwards there was huge drop of new cases count in China. This drop counts obviously indicating that China figure out a solution to control the virus but not sharing same information with other countries.

Line Chart

From decades ago, people present continuous data in a line chart format to show their valuable data. This is a most basic type of charts used throughout the industries and educational institutions for better understanding and simple representations of data. This line charts are very useful for comparison of multiple datasets or multiple columns in a single dataset. Line charts always provides quick analysis of data (Add Trend Lines to a Visualization - Tableau, 2020). It also tells about any gaps or clusters in the data set and how it changes as per times.

Technical Details

- ❖ By using line chart, we can do new cases and death cases forecasting.
- ❖ Select Sum(new cases) and Sum(death cases) in rows shelves and added Week(date) in columns shelves.
- ❖ We can rename and format the axis.
- ❖ Can draw dual line charts with right click on axis and select the Dual Axis option.
- ❖ Go to analytics section and select the forecast option.
- ❖ In the forecast option, we can specify the upcoming weeks forecasting. In the figure 6, we forecast till September 2020.
- ❖ We also can check Describe Forecast option where we can see the modelling their prediction techniques. Refer figure 7.
- ❖ We can also add trends analysis by selecting Trends Lines by right click on the graph. Refer to figure 6.
- ❖ Select Trends describe model after applying Trends Lines which will pop of a window with their trend analysis details. Refer to figure 8.

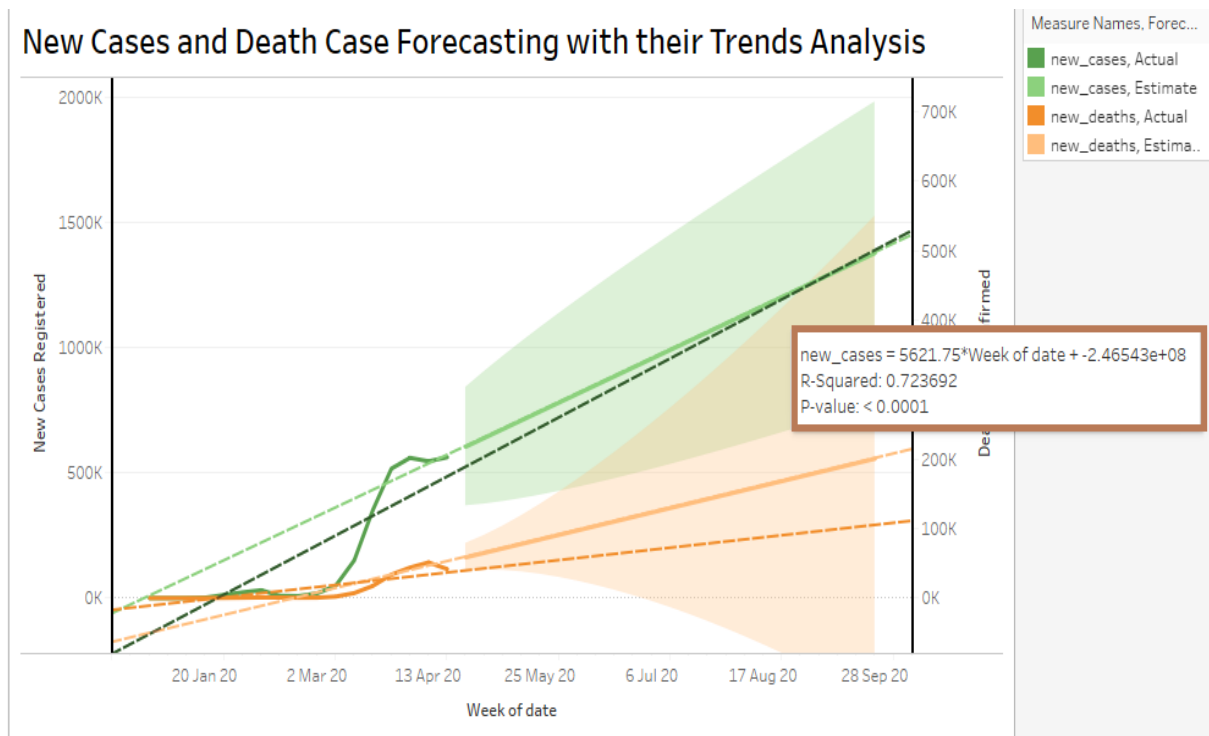


Figure 6 Forecasting and Trend Analysis of New and Death cases in the world

In the figure 6, we found that there is speedy increase in new cases from 17 Feb 2020 to 30 Mar 2020. Exactly after two weeks that means from 2 April 2020 onwards there is enormous growth of death count also. So, from the graph we can conclude that this deadly virus needs two weeks of time to make that person functionally dead. As we used forecasting option and predicted till 28 Sep 2020, we found that the new cases count will increase drastically in upcoming weeks. Similarly, death cases count also increase but it will be low as compare to new cases. In the figure 6, we can see both the forecasting and trend analysis at one graph. Looking at the forecasting intervals, we can say that new cases registered in worldwide will be varied from 750k to 2000k whereas death cases varied from minimum to 500k. This interval is obtained as 95% confidence intervals from the model. However, if we observe the trends then it clearly shows that till 28 Sep 2020 new and death cases will be increase drastically.

Options Used to Create Forecasts

Time series: Week of date

Measures: Sum of new_cases, Sum of new_deaths

Forecast forward: 23 weeks (27 April 2020 – 28 September 2020)

Forecast based on: 30 December 2019 – 20 April 2020

Ignore last: No periods ignored

Seasonal pattern: None (Not enough data to search for a seasonal pattern recurring every 13 Weeks)

Sum of new_cases

Initial 27 April 2020	Change From Initial 27 April 2020 – 28 September 2020	Seasonal Effect		Contribution		Quality
		High	Low	Trend	Season	
604,862 ± 236,743	774,154	None		100.0%	0.0%	Poor

Sum of new_deaths

Initial 27 April 2020	Change From Initial 27 April 2020 – 28 September 2020	Seasonal Effect		Contribution		Quality
		High	Low	Trend	Season	
58,591 ± 19,921	142,089	None		100.0%	0.0%	Poor

Figure 7 Forecast Model Description

Forecasting means predicting the future of a measures. We can use many mathematical models for forecasting. However, in Tableau there is default forecasting model known as exponential smoothing. In this, exponential smoothing is considered recent observations and are relatively more weight than the older observations. There are two important concepts for forecasting in tableau such as Trends and seasonality. The trend is increase or decrease in data over time whereas seasonality is repeating variation in values over determined period of time. This seasonal and repeating variations are predictable and provide accurate forecast. In the figure 7, we can observe that seasonal accuracy is 0% which makes the quality accuracy also poor. Tableau tool forecasting accuracy is low due to black box implementation. We also observe that from 27 April 2020 to 28 September 2020 change in new cases registered values is high around 775. However, in death cases confirmation change is comparable low such as 142.08. Hence proved that in future death cases confirmation will be less as compare to new cases registered.

Describe Trend Model

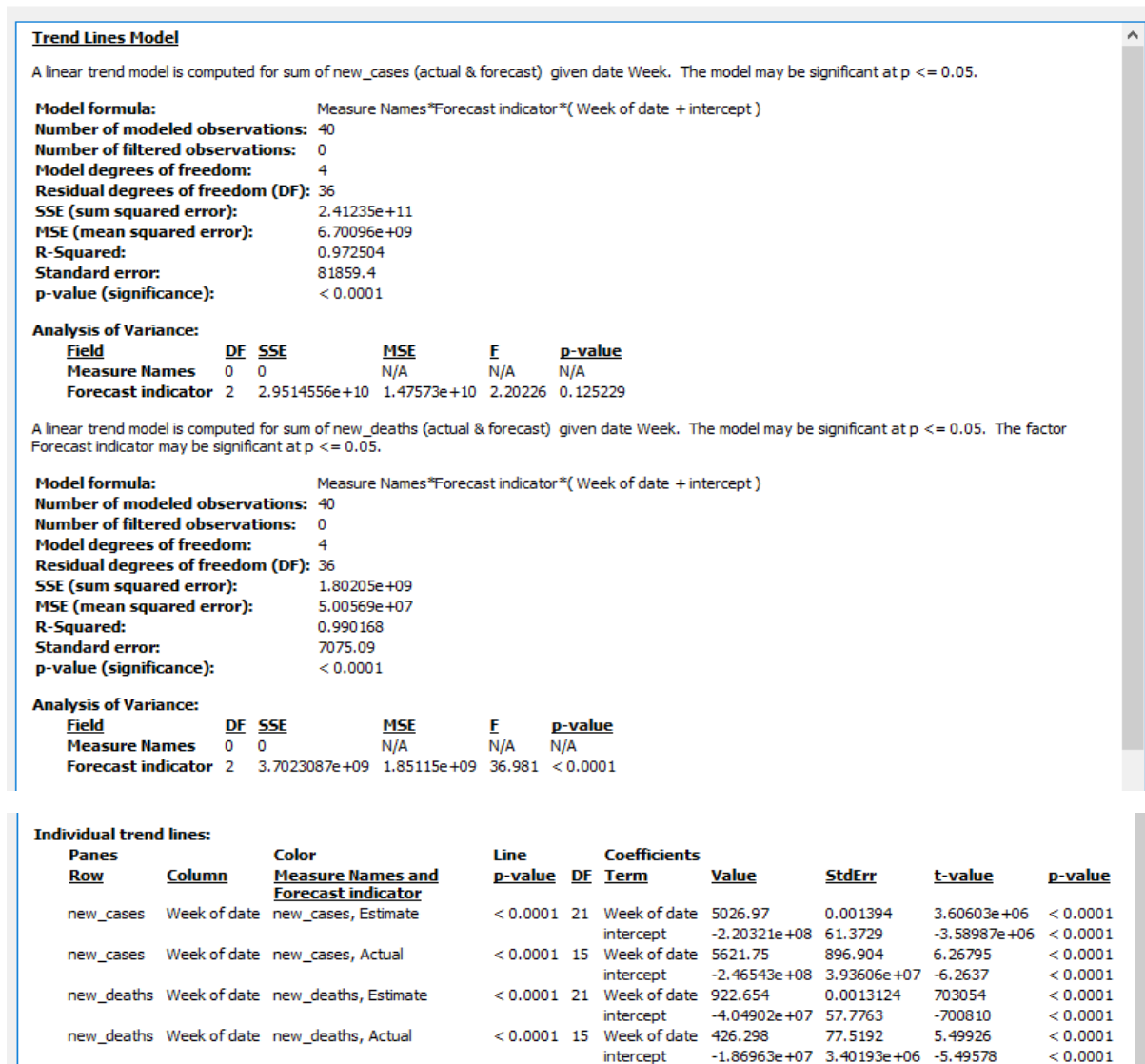


Figure 8 Described Trend Model

Trends lines are used to predict the trends of a variable in a dataset. It also helps to figure out the correlation between two variables by observing the trends. In Tableau, we have many mathematical models such as Linear, Logarithmic, Exponential and Polynomial. In figure 6, we used Logarithmic modelling to fetch the trends. In logarithmic modelling, formula is:

$$Y = b_0 + b_1 * \ln(X)$$

Because a logarithm is not defined for number less than zero, any marks for which the explanatory variable is negative are filtered before estimation of the model. One should avoid using a model that discards some data unless one knows that the data being filtered out is invalid. In the trend description report it was mentioned that R-squared value, Mean Squared Error, Degree of Freedom and P-value. In the figure 8, if we observed then the trend follows R^2 value as 0.97 which is really high as p value is < 0.0001. Here we are considering significant of p-value as 0.05.

Pie Chart

In a Pie chart, to calculate the percentage each slice is worth, measure the angle of each slice and divide this by 360 then multiply it by 100. To find the number of pieces of data each slice represents, multiply the percentage that each slice is worth by the total number of the data sets. Pie chart is good for presenting well in 6 or less categorical data.

Technical Details

- ❖ Create an Excel file in which only location and total cases count of South America.
- ❖ Do an inner join with existing Data set with location. Refer figure 9
- ❖ Change the location data type to geographical role.
- ❖ Select longitude and latitude for South America.
- ❖ Create calculative field called Week No
- ❖ In the Mark card, select Pie option and drag Sum(new_cases) to Angle and Size.
- ❖ Select Week No field to colour option in Mark card.
- ❖ Under label select location Name and Total_count of cases. Refer figure 10

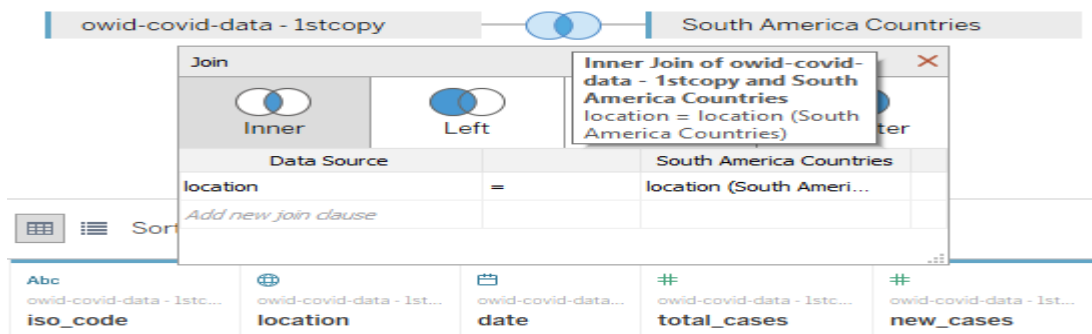


Figure 9 Inner joins in Tableau with new excel file

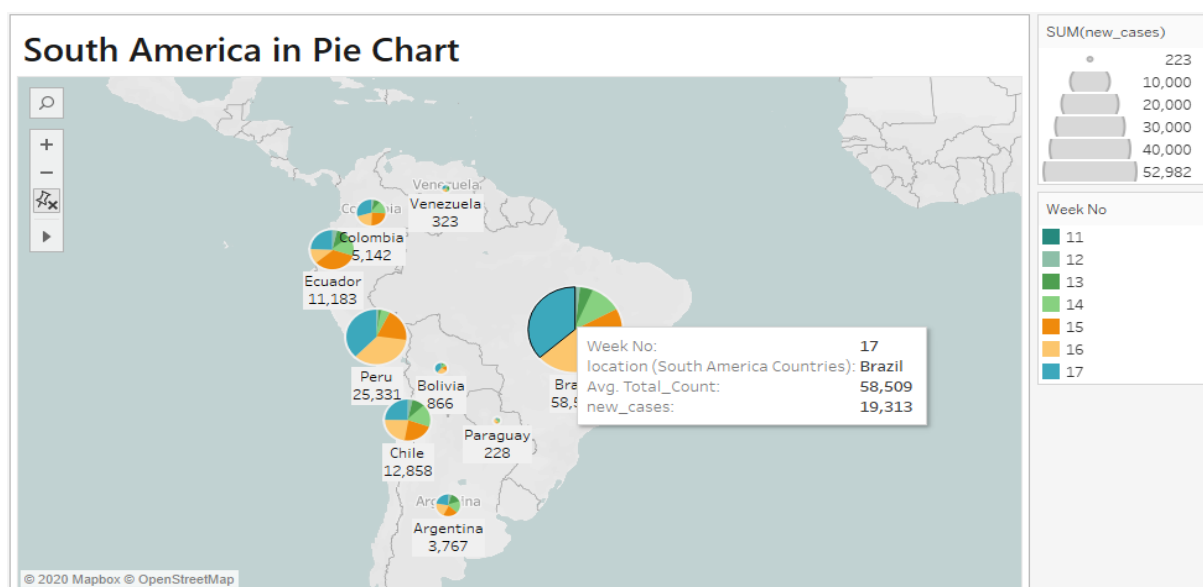


Figure 10 South America Covid Impacted Locations and their Counts

From the figure 10, we can conclude that in South America most impacted country is Brazil by see their size. By labels we can say that total cases registered by April in Brazil is 58,509. The second position is occupied by Peru with count of 25,331 cases in total. If we observe closely, more cases are registered during Week 15 to 17 in South America. In Paraguay, less cases were registered by April as compare to other countries in South America and same things reflecting by size in the figure 10. We can add transparency to view data in a better manner as currently we can't see small countries geographical locations in South America. This could be one of the limitations of Tableau tools.

Tree map

In Tableau tool, Tree map populates the data in a rectangular nested loop. The dimension defines the structure of the Tree map and size and color defines the individual rectangles (Shneiderman, Dunne, Sharma and Wang, 2011). We can create Tree map with one or more dimensions and one or more measures using Tableau tool.

Technical Details

- ❖ Select the combination field of Year and Week in the column shelf and sum(new_tests) in the row shelf.
- ❖ Under the Mark Card, select the location as colours, size for sum(new_tests) and select location, sum(new_tests) and Combination field of Year and Week as Labels.
- ❖ As we have limited data, so can be filtered out those values are null.
- ❖ Rename the Titles.

Tree Map Visualisation for Testing Dataset

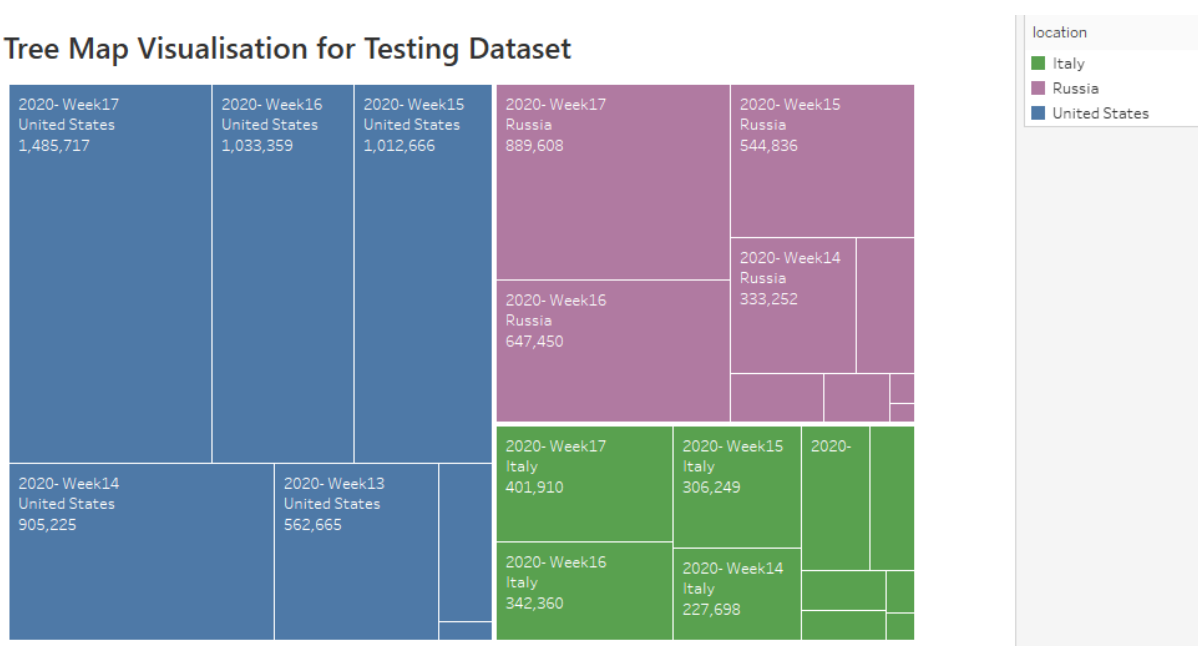


Figure 11 Total Test cases of Countries during Week 13 to 17

From the Tree map, we can see that more cases are tested during week 17 that means there are high probability of chances that in week 18 or 19 more cases will be registered. As we have limited data in our dataset like few countries such as United States, Russia and Italy. So, it's hard to predict by visualising this Tree Map.

Story and Dashboard

In Tableau we have advanced features to add a story or dashboard for reporting purpose. This feature helps the user to add multiple views at one place. We can add captions and text data into this.

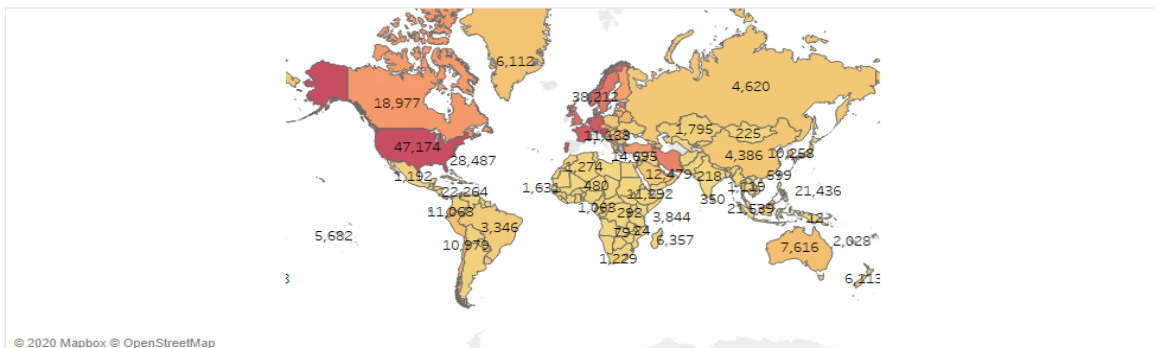
Technical Details

- ❖ Create a Dashboard, select the appropriate worksheets to populate.
- ❖ Select horizontal and vertical divisions in the dashboard.
- ❖ Go to create a new story and add the dashboard into it.
- ❖ Add the captions like current situation and forecasting.

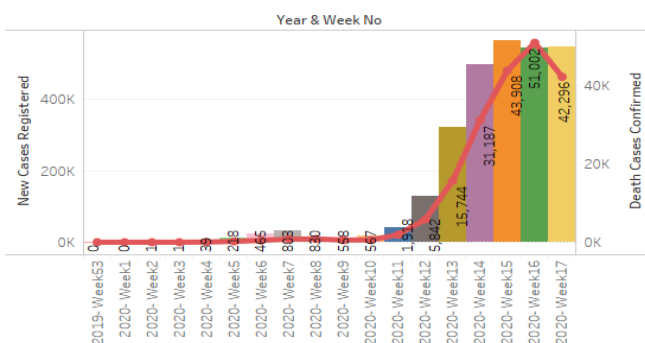
Coronavirus Story

Current Situation and Prediction of Coronavirus Worldwide

Covid Case Count Per Million

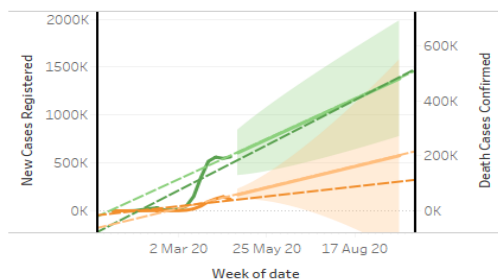


New and Death Cases Count



Current Situation

New Cases and Death Case Forecasting with their Trends Analysis



Forecasting Till September

Figure 12 Located and Comparing the Current situation and Forecasting the futures

Looking at the figure 12, we can conclude that in United States corona virus patients are high as compare to other countries. Corona Virus cases started in the Week 2 and continue to increase till Week 17. We also confirmed that Death cases count is multiple of approximately 7.5% to new cases count as per current situations. Same point has been proved during forecasting as well. As per forecasting till September 2020 it will increase the new cases count as well as death cases count. We have 95% intervals covered while doing modelling. If we closely observe Trends in figure 12, it is more linear line appear between Week no - New cases registered and Week No – Death cases confirmed. If we see that initial weeks then we can see, small variation between actual data predicted data.

R Statistic Visualisation Technique

Map Chart

By using R programming code, we can draw world map by using library Rworldmap. Maps will show as per geographical location.

Technical Details

- ❖ From original data set, need to convert into csv format with Country and Total_Cases_Count columns.
- ❖ We can read the csv file by using read.csv function.
- ❖ By using rworldmap library, we can show total_cases_count as per country.
- ❖ As for better representation, we add color bar to show from lowest to highest value.

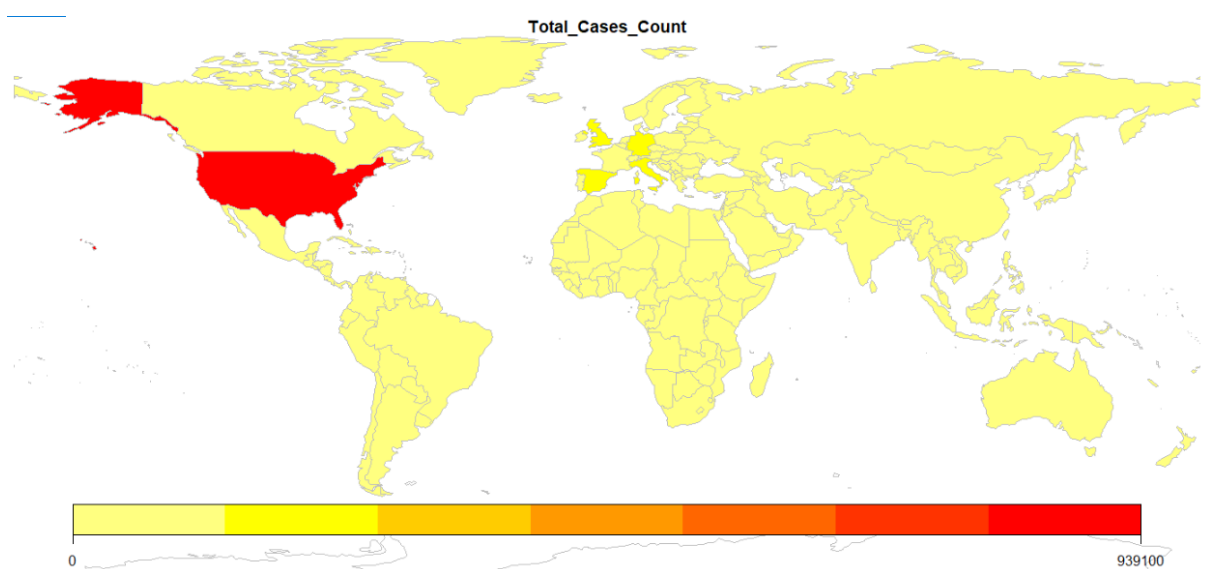


Figure 13 Covid impacted Countries as per their total Count

From the figure 13, we can easily conclude that United States has high covid cases registered throughout the world as the color code is red. Similarly, Germany, United Kingdom and France has second highest count of covid cases as per yellow color representation. Remaining countries has registered cases but comparably lower. There is huge difference between United States and Germany because Germany is proper yellow color coded whereas United States is dark red color coded.

Bar Chart

When we have categorical data, the best way is to represent in the bar chart. We can visualise bar plot in R statistics in two ways such as barplot function and ggplot function. To represent top 10 countries and their week wise increase in total cases count in a bar chart.

Technical Details

- ❖ Install the ggplot2, dplyr packages in your R.
- ❖ From original dataset, create new csv file with top 10 countries week wise data.
- ❖ Create y matrix by using ddply function.
- ❖ Use this y in the ggplot and reorder as per Country, week respectively.
- ❖ Add the legend as per Week wise such as 1 to 17. Even we can apply colors for the same.

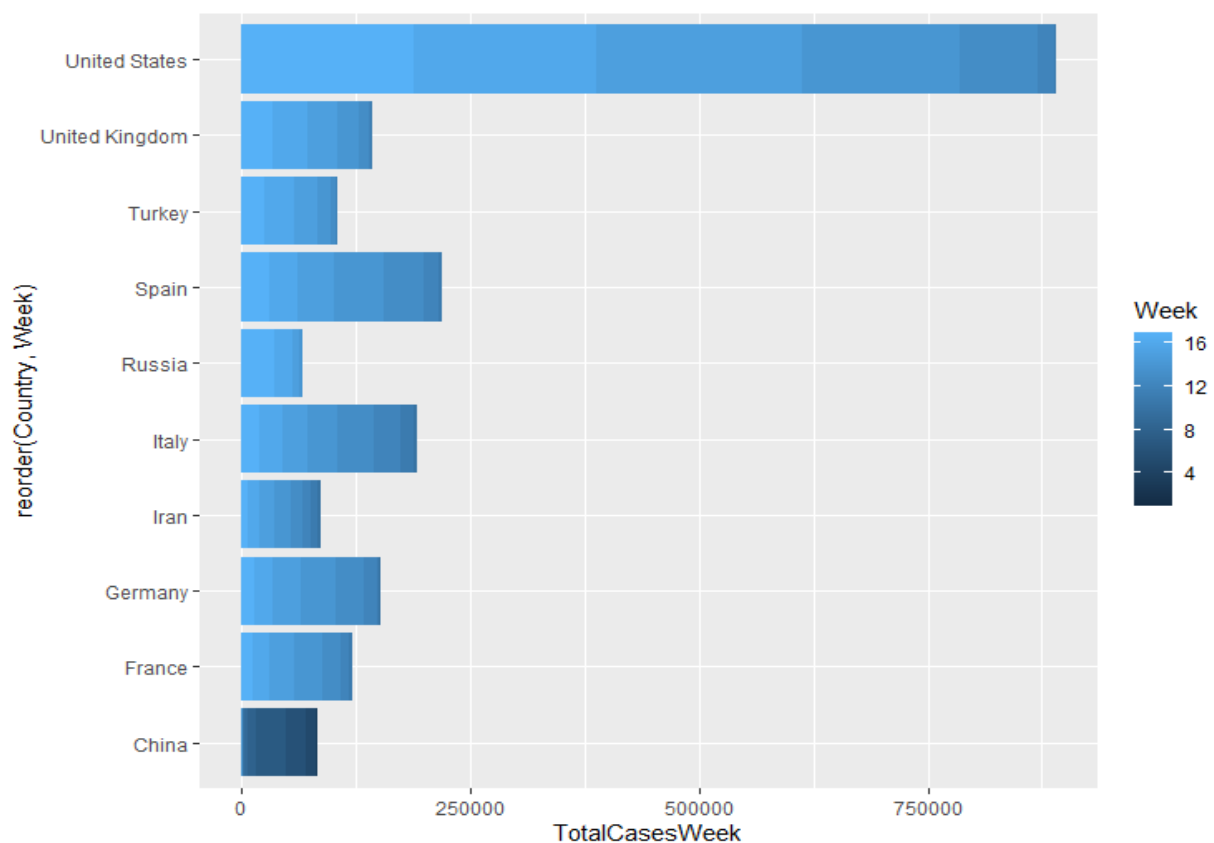


Figure 14 Top 10 Countries Covid cases registered per Week

We have represented top 10 country's total covid cases data in a bar chart. From the figure 14, we can easily conclude that United States is the highest cases registered and maximum cases registered during Week 13 to 17. The last position occupied by Russia. If we observe China, then more cases registered in the initial weeks such as week 4 to 7 after that its starts reducing. Hence proved that Novel Corona virus was originated in China. The second position occupied by Spain followed by Italy.

Line Chart

Using R statistic, we can also draw traditional Line chart as well. We can do multiple line charts by using R codes.

Technical Details

- ❖ From original data set, need to transform into csv format with Country, Week and Total_Cases_Week_Wise columns.
- ❖ Create X matrix by using ddpby function.
- ❖ Use this X in the ggplot and aes as per week respectively.
- ❖ Add the legend as per Country wise such as top 5 countries. Even we can apply colors for the same.

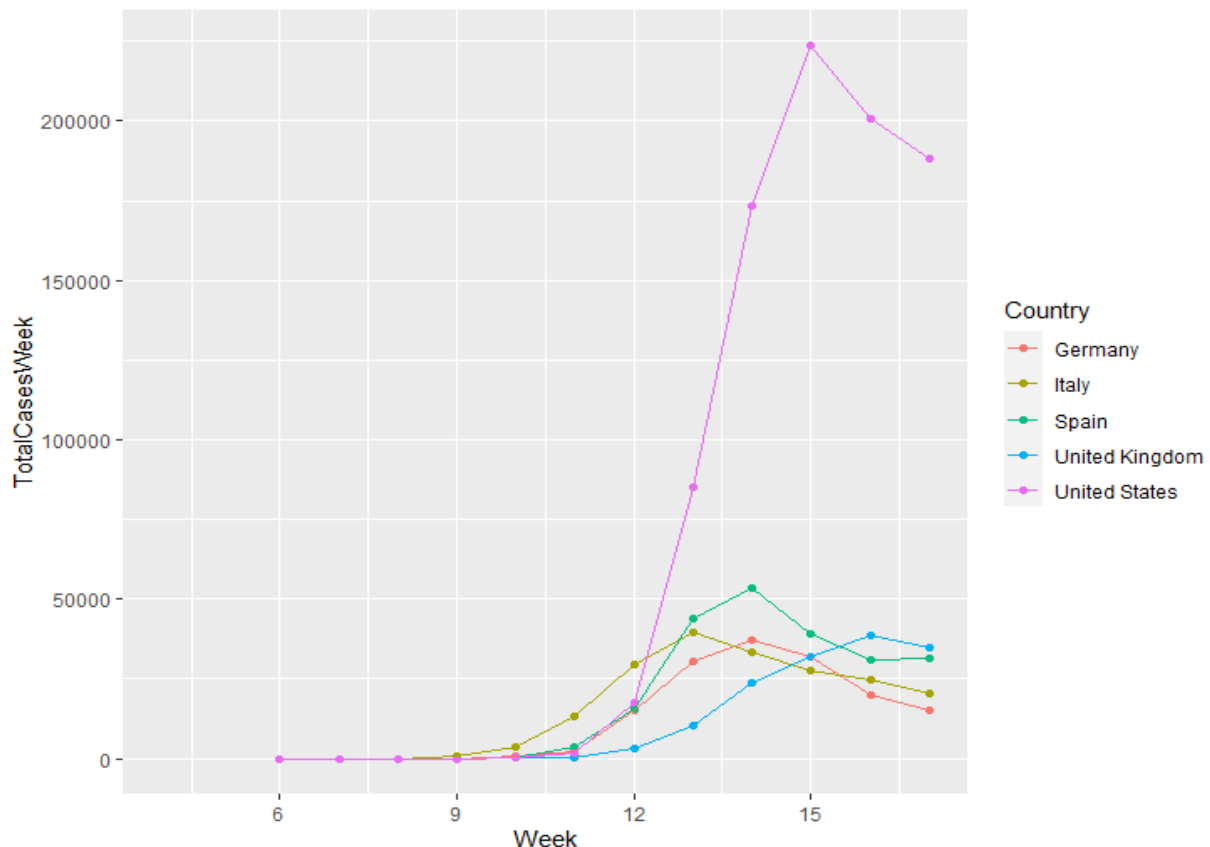


Figure 15 Top 5 countries Covid cases Count

From the figure 15, we can see that United States has highest count as per compare to others. There is huge increase in cases from Week 13 to 15. After that there is decrease in cases but not in a margin rate. This lower down could be people starts aware of Novel Corona Virus. During Week 13 to 16 there are a greater number of cases registered in all other countries as well.

Comparison Analysis on Different Tools

By analysing two different tools with multiple techniques such as Tableau and R statistics, we can conclude that Tableau tool is much faster and easier as compare to R statistics. Because Tableau tool is very much user friendly so any non-technical person in data visualisation can easily understand and draw the graph. Whereas it's hard to understand R code for a non-technical people. We used multiple techniques in Tableau to represent different data like Map chart, Bar Chart, Pie Chart, Line Chart and Tree Map. However, we have very limited option such as Map chart, bar chart and line chart in R statistics. If we compare the Map chart from Tableau and R statistic, we can easily say that Tableau Map chart more interactive than R code generated one. Similarly, we found that bar chart and line chart of R code also have less interactive than Tableau Bar and Line chart. We also observe that both graphs created by Tableau and R Statistics provides similar results and conclusion to user. This helps user to figure out the hidden pattern of the information.

Conclusion

We used Covid data set in the Tableau tool to analyse their hidden pattern. We found that during Week 12 to Week 17 we have highest number of Covid cases registered in the entire world except one country viz. China. In China we have more cases during week 4 to 8 but later that was lower down and flat. Either China found the solution for Novel corona virus otherwise as it was initiated over there so immediate lockdown helps them. Throughout the world cases increase, this could indicate that during week 8 or 11 more people travel from China to different countries as this virus can be transfer from human to human only. Because of long term lockdown and Corona virus impact analysis says that there is huge economic crisis coming up in near future as the cases increases in whole world. In the Line chart of figure 6, if we compare the new cases and death cases count then it takes exactly 2 weeks to show their deadly symptoms. We can analyse more if we have each person's details like their age, their located countries and their travel history who has tested positive in covid cases. In our current covid dataset, we have highest countries who gets impacted most due to Covid is United States, Spain, Germany, Italy, United Kingdom and France. If we look into their forecast or trends analysis, then we can conclude that in future also there are high chance of increase Corona virus cases.

Reference

1. Martin, I. and Marinescu, D., 1998. Data visualization: Concurrent computation and data visualization for spherical-virus structure determination. *IEEE Computational Science and Engineering*, 5(4), pp.40-52.
2. Visualization, B., Projects, P., Kumar, A. and Kumar, A., 2020. Top 8 Amazing Benefits of Using Tableau for Your Projects. [online] Knowledge Hut Blog. Available at: <<https://www.knowledgehut.com/blog/business-intelligence-and-visualization/tableau-advantages-disadvantages>> .
3. www.javatpoint.com. 2020. R Advantages and Disadvantages - Java point. [online] Available at: <<https://www.javatpoint.com/r-advantages-and-disadvantages>> [Accessed 3 June 2020].
4. Herman, I., Melancon, G. and Marshall, M., 2000. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1), pp.24-43.
5. Our World in Data. 2020. Coronavirus Source Data. [online] Available at: <<https://ourworldindata.org/coronavirus-source-data>> [Accessed 3 June 2020].
6. Help.tableau.com. 2020. Add Trend Lines to A Visualization - Tableau. [online] Available at: <https://help.tableau.com/current/pro/desktop/en-us/trendlines_add.htm#:~:text=Add%20trend%20lines%20to%20a%20view,Polynomial%2C%20or%20Power%20model%20types.> .
7. Yu, C., Yudofsky, D. and Xu, T., 2011. Visual Data Mining: An Exploratory Approach to Analysing Temporal Patterns of Eye Movements. *Infancy*, 17(1), pp.33-60.
8. Tejada, E., Minghim, R. and Nonato, L., 2003. On Improved Projection Techniques to Support Visual Exploration of Multi-Dimensional Data Sets. *Information Visualization*, 2(4), pp.218-231.
9. Shneiderman, B., Dunne, C., Sharma, P. and Wang, P., 2011. Innovation trajectories for information visualizations: Comparing tree maps, cone trees, and hyperbolic trees. *Information Visualization*, 11(2), pp.87-105.