Analysis of Covid-19 and Vaccination Data

Syama Ravi Teja Jerrypothula-015906098

Department of Applied Data Science, San Jose State University

Data 230: Data Visualization

Andrew H. Bond

May 22, 2022

GitHub Link: https://github.com/syamrt/Data-Visualizations

Abstract

The year 2020 is the nightmare for the world. The reason behind that unexpected nightmare is covid -19. This disease effected the whole world in many ways. It doesn't effected only the health of humans but also living style, priorities, and many other things. In that pandemic, governments of countries tried to face this virus in many ways. In process of that data played a main key role. From predicting intensity of cases in terms of numbers to develop the vaccine, data is used to cross check the decisions. Depending on the analysis done on the covid-19 related data, governments are making decisions on the constraints, policies and etc. This project is about collecting data about covid-19 and making analysis of that data. We collected data from two reliable sources one is from World Health Organization (WHO) and other is Our Data in World. Using that data, will create meaningful visualizations in power bi and derive useful conclusions.

Project Goals and Motivation

The impact of the pandemic is more than enough as a motivation to select this topic. The goal of this project is to analyze the covid-19 related and to come up with useful conclusion. To achieve that we will create several dashboards by dividing into different parts. We will make a dashboard for providing the current information regarding cases of covid with respective countries and continents. In addition to that will make dashboards which are interactive to see trend of cases reporting, recovering, deaths. Considering vaccination data, dashboards are made similarly as covid-19 dashboards which tells about number of vaccinations happened until now as single/first dose vaccination, second dose vaccinations, and booster shots.

Objectives

- Total count of countries effected with covid.
- A total number of cases reported, recovered, deaths in the world.
- A total number of cases reported, recovered, deaths in each country.
- Number of cases reporting with respective year.
- Number of cases recovered, deaths with respective year.
- Geological Maps projecting intensity of cases reporting.
- Trend analysis of cases reporting, recovering, deaths in each country with date filter.
- Number of cumulative cases, deaths in a country of a particular date.
- Share of each continent in cases reporting, recovering, deaths, testing in the world.
- Total number of countries vaccinations are initiated.
- Total number of vaccinations in the world.
- Total number of vaccinations with respective countries.
- Total number of Second dose, Booster dose vaccinations in the world and country.
- Number of Vaccinations in a state of US
- Number of vaccinations in particular age group in US
- Details of vaccination in a particular country.
- Integrating python with power bi and creating visualizations.
- Number of vaccinations in each state of US.

Dataset Description

There are several datasets are available publicly about the covid-19. John hopkins is the first organization started collecting data regarding the virus like number of cases region wise. But as this issue changed into pandemic, many organizations including World Health Organization (WHO), governments started maintaining data about virus. This data is useful in many ways to governments to make decisions in order to face the virus. We collected multiple datasets to cover all kinds of wide data. In process of that we have collected data from WHO and our world in data. we downloaded data from both the organizations in csv format.

The dataset from WHO consists of 2,01,924 rows with 8 columns, from our world in data consists of 1,81,684 with 50 columns. Second dataset have many columns indicating different kind of measures. But we will consider only necessary columns which are useful for analysis. The columns of datasets which are used in this project is describe as shown below.

The columns from both datasets described as below:

Date: date is the reported/recorded date of information. So, this field is in the date format. The first date in the dataset is January 1 2020 and latest is May 06 2022. All the dates are almost unique.

Age group: categorical column which tells about the age group does it belongs to.

WHO Region: Code of the region labeled by WHO. This column contains seven unique values.

Country: It contains the name of the country from where information recording. There are 217 unique values in the dataset.

Country code: The code of the country. There are 217 unique values.

New_Cases: Number which indicates number of new cases reported.

New_deaths: Number of new deaths reported on particular date.

Continent: The name of the continent from which cases are reported.

Hosp_pateints: number of patients hospitalized on that day.

Icu_pateints: number of patients joined/taking treatment in the icu.

Total vaccinations: Number of vaccination recorded until now.

Partial Vaccine: The number of people vaccinated partially (1st dose) on the reported date.

Booster_shot: The number of people vaccinated with booster shots.

Hand washing facility: This field indicated about the number of facilities arranged for hand washing.

Gdp_per_capita: A decimal value indicates the gdp value of that country.

Life_expectancy : The avg age of the person in that country.

Reproduction_rate: The number tells about the reproduction rate of that country on reported date.

Source for the 1st dataset: Download

Source for the 2nd dataset: Download

Raw Data Sample

Figure 1

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
iso_code	continent	location	date	total_case	new_case	new_case	total_deat	new_deatl	new_deatl	total_case	new_case	new_case	total_deat	new_deatl	new_deatl	reproducti	icu_patien	icu_patien	hosp_patie	hosp_patie	weekly_icu	weekly_icu
AFG	Asia	Afghanista	24-02-20	5	5					0.126	0.126											
AFG	Asia	Afghanista	25-02-20	5	0					0.126	0											
AFG	Asia	Afghanista	26-02-20	5	0					0.126	0											
AFG	Asia	Afghanista	27-02-20	5	0					0.126	0											
AFG	Asia	Afghanista	28-02-20	5	0					0.126	0											
AFG	Asia	Afghanista	29-02-20	5	0	0.714				0.126	0	0.018										
AFG	Asia	Afghanista	01-03-20	5	0	0.714				0.126	0	0.018										
AFG	Asia	Afghanista	02-03-20	5	0	0				0.126	0	0										
AFG	Asia	Afghanista	03-03-20	5	0	0				0.126	0	0										
AFG	Asia	Afghanista	04-03-20	5	0	0				0.126	0	0										
AFG	Asia	Afghanista	05-03-20	5	0	0				0.126	0	0										
AFG	Asia	Afghanista	06-03-20	5	0	0				0.126	0	0										
AFG	Asia	Afghanista	07-03-20	8	3	0.429				0.201	0.075	0.011										
AFG	Asia	Afghanista	08-03-20	8	0	0.429				0.201	0	0.011										
AFG	Asia	Afghanista	09-03-20	8	0	0.429				0.201	0	0.011										
AFG	Asia	Afghanista	10-03-20	8	0	0.429				0.201	0	0.011										
AFG	Asia	Afghanista	11-03-20	11	3	0.857				0.276	0.075	0.022										
AFG	Asia	Afghanista	12-03-20	11	0	0.857				0.276	0	0.022										
AFG	Asia	Afghanista	13-03-20	11	0	0.857				0.276	0	0.022										
AFG	Asia	Afghanista	14-03-20	14	3	0.857				0.351	0.075	0.022										
AFG	Asia	Afghanista	15-03-20	20	6	1.714				0.502	0.151	0.043										
AFG	Asia	Afghanista	16-03-20	25	5	2.429				0.628	0.126	0.061										
AFG	Asia	Afghanista	17-03-20	26	1	2.571				0.653	0.025	0.065										
AFG	Asia	Afghanista	18-03-20	26	0	2.143				0.653	0	0.054										
AFG	Asia	Afghanista	19-03-20	26	0	2.143				0.653	0	0.054										

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
AFG	Asia	Afghanista	22-12-20	50888	210	200.571	2089	15	16.286	1277.456	5.272	5.035	52.441	0.377	0.409	0.93
AFG	Asia	Afghanista	23-12-20	51070	182	195.286	2105	16	14.857	1282.025	4.569	4.902	52.842	0.402	0.373	0.92
AFG	Asia	Afghanista	24-12-20	51357	287	204.286	2126	21	15.571	1289.229	7.205	5.128	53.37	0.527	0.391	0.92
AFG	Asia	Afghanista	25-12-20	51595	238	199	2135	9	14.714	1295.204	5.975	4.996	53.596	0.226	0.369	0.9
AFG	Asia	Afghanista	26-12-20	51764	169	186.857	2148	13	15.857	1299.446	4.242	4.691	53.922	0.326	0.398	0.89
AFG	Asia	Afghanista	27-12-20	51848	84	187.429	2158	10	14.857	1301.555	2.109	4.705	54.173	0.251	0.373	0.88
AFG	Asia	Afghanista	28-12-20	52007	159	189.857	2170	12	13.714	1305.546	3.991	4.766	54.474	0.301	0.344	0.87
AFG	Asia	Afghanista	29-12-20	52147	140	179.857	2182	12	13.286	1309.061	3.514	4.515	54.775	0.301	0.334	0.86
AFG	Asia	Afghanista	30-12-20	52330	183	180	2189	7	12	1313.655	4.594	4.519	54.951	0.176	0.301	0.85
AFG	Asia	Afghanista	31-12-20	52330	0	139	2189	0	9	1313.655	0	3.489	54.951	0	0.226	0.85
AFG	Asia	Afghanista	01-01-21	52513	183	131.143	2201	12	9.429	1318.249	4.594	3.292	55.252	0.301	0.237	0.84
AFG	Asia	Afghanista	02-01-21	52586	73	117.429	2211	10	9	1320.081	1.833	2.948	55.503	0.251	0.226	0.84
AFG	Asia	Afghanista	03-01-21	52709	123	123	2221	10	9	1323.169	3.088	3.088	55.754	0.251	0.226	0.84
AFG	Asia	Afghanista	04-01-21	52909	200	128.857	2230	9	8.571	1328.19	5.021	3.235	55.98	0.226	0.215	0.84
AFG	Asia	Afghanista	05-01-21	53011	102	123.429	2237	7	7.857	1330.75	2.561	3.098	56.156	0.176	0.197	0.83
AFG	Asia	Afghanista	06-01-21	53105	94	110.714	2244	7	7.857	1333.11	2.36	2.779	56.332	0.176	0.197	0.82
AFG	Asia	Afghanista	07-01-21	53207	102	125.286	2253	9	9.143	1335.67	2.561	3.145	56.558	0.226	0.23	0.82
AFG	Asia	Afghanista	08-01-21	53332	125	117	2257	4	8	1338.808	3.138	2.937	56.658	0.1	0.201	0.81
AFG	Asia	Afghanista	09-01-21	53400	68	116.286	2264	7	7.571	1340.515	1.707	2.919	56.834	0.176	0.19	0.81
AFG	Asia	Afghanista	10-01-21	53489	89	111.429	2277	13	8	1342.749	2.234	2.797	57.16	0.326	0.201	0.81
AFG	Asia	Afghanista	11-01-21	53538	49	89.857	2288	11	8.286	1343.98	1.23	2.256	57.436	0.276	0.208	0.81
AFG	Asia	Afghanista	12-01-21	53584	46	81.857	2301	13	9.143	1345.134	1.155	2.055	57.763	0.326	0.23	0.81
AFG	Asia	Afghanista	13-01-21	53690	106	83.571	2308	7	9.143	1347.795	2.661	2.098	57.938	0.176	0.23	0.81
AFG	Asia	Afghanista	14-01-21	53775	85	81.143	2314	6	8.714	1349.929	2.134	2.037	58.089	0.151	0.219	0.82
AFG	Asia	Afghanista	15-01-21	53831	56	71.286	2324	10	9.571	1351.335	1.406	1.79	58.34	0.251	0.24	0.82
AFG	Asia	Afghanista	16-01-21	53938	107	76.857	2336	12	10.286	1354.021	2.686	1.929	58.641	0.301	0.258	0.82

Sample of Prepared Data

Figure 2

Continent	Country,Other	TotalCases	NewCases	TotalDeaths	NewDeaths	TotalRecovered	NewRecovered	ActiveCases	Serious,Critical	Tot Cases/1M pop	Deaths/1M pop	TotalTests
Asia	Georgia	1655221		16811		1637293		1117		416391	4229	16920079
Europe	Belarus	980828		6951		928536		45341		103861	736	13171246
Asia	Nepal	978836		11951		966674		211		32524	397	5642144
Asia	UAE	899033		2302		882591		14140		88934	228	156338294
Asia	Azerbaijan	792579		9709		782786		84		76902	942	6819005
Asia	Sri Lanka	663408		16505		646129		774		30745	765	6486117
Asia	Myanmar	612914		19434		591847		1633		11128	353	7970408
Europe	Bosnia and Herzegovina	377087		15765		192218		169104		116260	4861	1758191
Africa	Kenya	323792		5649		318002		141		5788	101	3602311
Europe	North Macedonia	309769		9280		299865		624		148697	4455	2022649
Europe	Iceland	185579		119		75685		109775		537451	345	1953616
Africa	Namibia	158811		4025		153662		1124		60498	1533	1007690
North America	Jamaica	130299		2967		83689		43643		43654	994	1004132
Africa	Rwanda	129812		1459		45522		82831		9588	108	5279288
Africa	Angola	99194		1900		97149		145		2856	55	1499795
Africa	DRC	87023		1337		50930		34756		920	14	846704
Africa	Senegal	86005		1966		84019		20		4900	112	1070306
Africa	Ivory Coast	81957		799		81112		46		2972	29	1501863
South America	Suriname	79336		1328		49401		28607		133059	2227	235950
Europe	Channel Islands	74383		169		73224		990		420730	956	1252808
North America	Barbados	70793		398		66662		3733		245793	1382	651301
Africa	Mauritania	58694		982		57695		17		12044	202	804490
Asia	Syria	55825		3150		52201		474		3054	172	146269
Africa	Gabon	47602		304		47285		13		20502	131	1593622
Africa	Seychelles	42474		166		41727		581		427042	1669	
Africa	Burundi	39998		38		773		39187		3190	3	345742
Africa	Mauritius	37691		990		35962		739		29547	776	358675

Sheet1 (211 rows)

Challenges Faced

Data is available publicly very well but the challenging part to get the correct data. For that I cross verified the figures with many other sources like world meter whether data is reliable or not. After that, dataset consists of so many attributes which are not necessary for our analysis. So, filtering them is also kind of challenging part. Power BI is a new tool for me to make dashboards. So, It took some time to understand the environment of Power BI and to work on that. Apart from this, integrating python with the Power BI is a challenging part which deals with server and code part.

Deliverables

In this project, we will create dashboards of different aspects from the dataset which will help to understand and analyze the data. This is an important deliverable of the project. Presentation will be presented in class. A report in APA format about the project will be submitted. Apart from that, all the source files of dashboards are uploaded to GitHub and link will be provided.

Scope of Work

As we are having huge data related to covid and vaccination in online. We will have a wide range of options to draw conclusions. So, that we can use all the options from Power BI for dashboards.

Visualizations

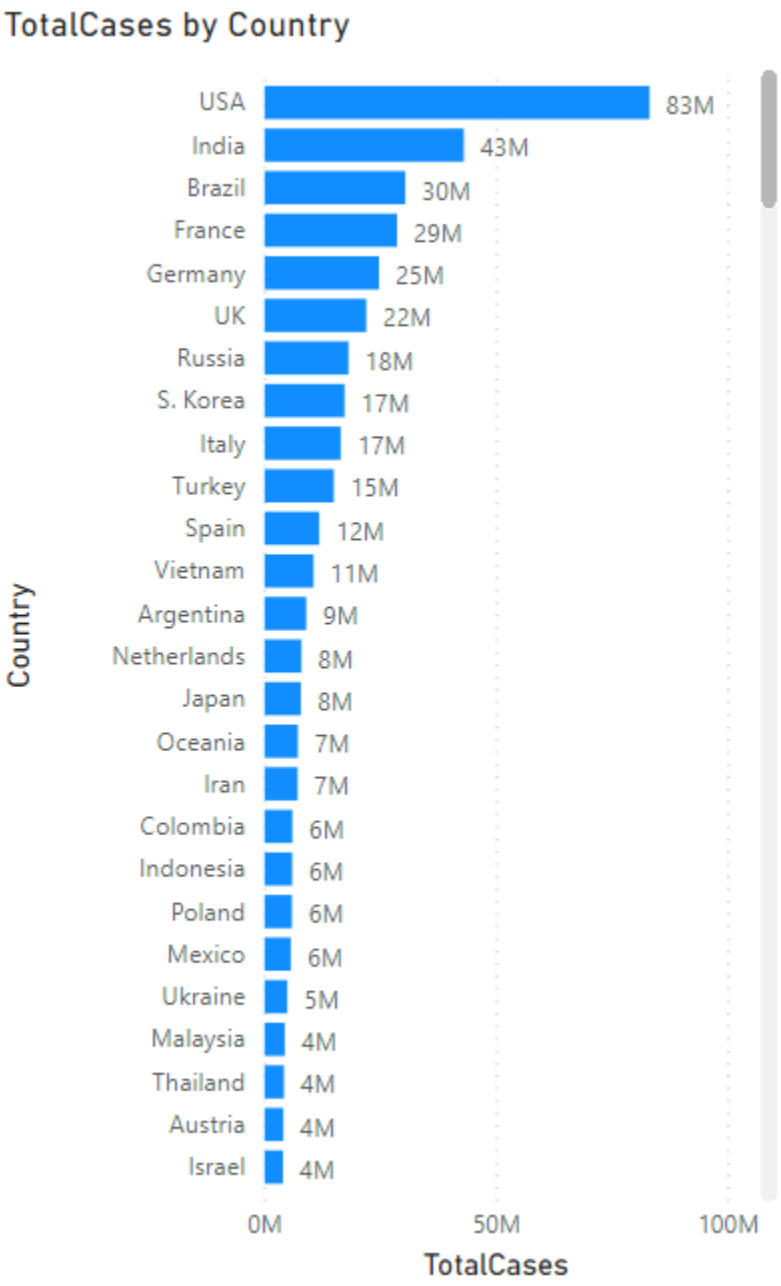
Score Card of Covid Global Data

The below figure 3 shows visualizations about the basic statistics related to covid-19. The type of the above visualization is score card. Score cards are the small sized display cardboards showing the figures in numbers or any names that you want to display. Before going to see about the numbers we need to note that all the statistics displays are according to the dataset. For example if you want shown number of unique countries in the dataset you can just drag countries column on the score card and select distinct count. Then you can see number of unique countries in the dataset. In the above figure there are five score cards. First one display number of countries effected with the covid-19 and procedure to get that discussed earlier. In the dataset there are 210 countries from where cases are reported. Second score card display the total number of cases reported from 2020 to until now which is 514 million. This figures includes all the cases recovered, deaths, active cases. The third score card displays number of cases recovered from covid-19 virus until now which is 547M. The fourth score card displays number of people died due to virus. The data tell that until now 6M people died with the virus. The fifth score card displays the number of active cases now which is 50M.

Figure 3

Bar Chart of Covid Cases Country Wise

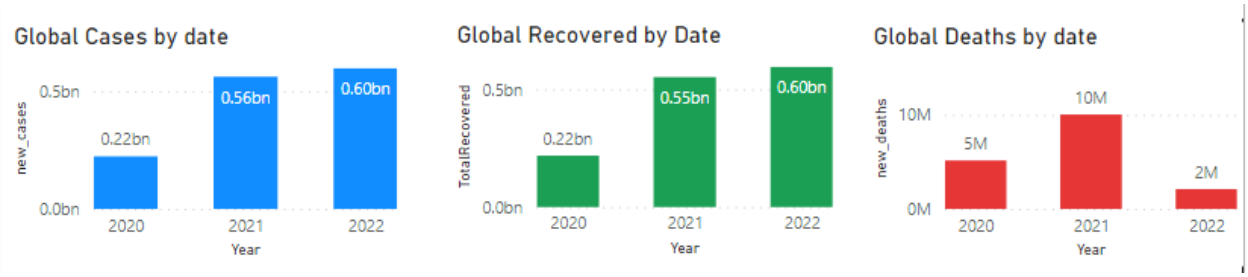
Figure 4



The above visualization is the stacked bar chart. Stacked bar chart is the bar chart which can be plotted with labels on each bar in the graph. Therefore it is called as stacked bar chart. The above bar chart shows the information regarding total number of cases reported until now

country wise. According to the graph the top most country with highest recorded cases until now in the world is USA with 83M cases. The next highest country is India with 43M cases. The least or lowest country with the covid cases is Saint Helena with 2 cases until now. Which is not surprising due to limited population of around 4 thousand and within area of 121.8 sq kilometers.

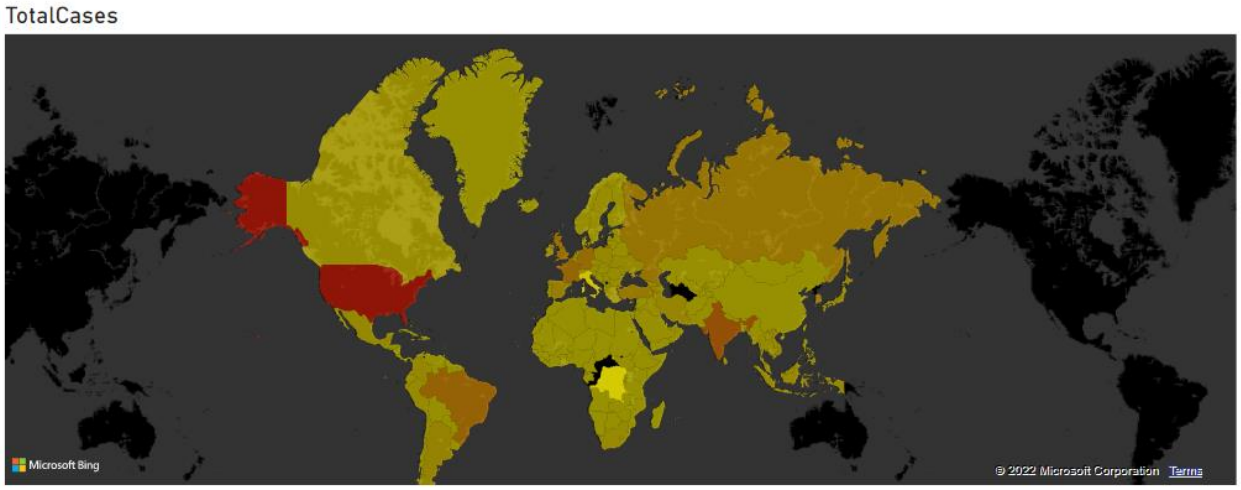
Figure 5



The above is the Stacked column chart. It is same as stacked bar chart but only the difference is in orientation of bars. In the above visualization there are three column charts showing the information related to covid-19 cases with respective year. The first chart shows the distribution of reporting cases through all 3 years. According to the graph, in 2020 number of covid cases reported are 0.22bn and that figure came to 0.56bn including the cases from the previous year. similarly the second columns chart displays the information related to recovered cases until now and the third chart displays the data related to number cases died until now in the world.

Map Graph of Covid-19 Cases

Figure 6



The above visualization is a map showing the information related number of covid cases that are reporting in the country. The color shows the intensity of cases reporting. The color varies from light yellow to dark red as intensity of cases increasing. According to the above map we can see that USA and India are in red color as number of cases reporting there is highest in the whole world.

Score Cards of Covid Cases of US

Figure 7

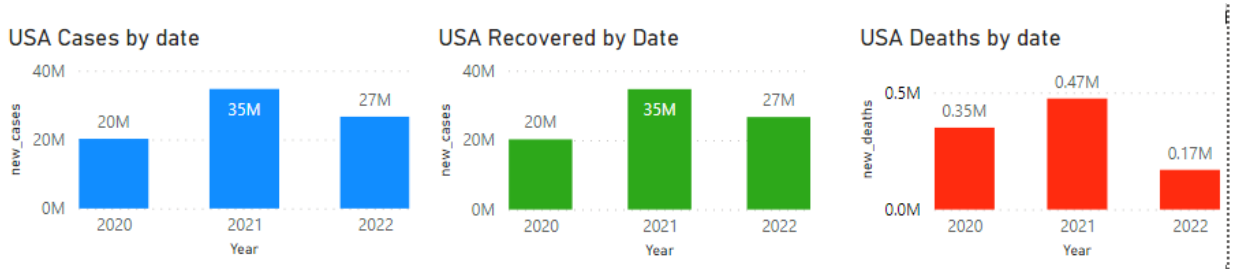
Total Effected Coutries	US Cases	US Recovered	US Deaths	Active Cases
1	83M	81M	1M	1M

There is a button in the 1st dashboard for selecting USA. When you click that button all dashboard will be changed to show data related to only USA. In that USA dashboard, the above is one of the visualization. In figure 3 we gave background about the type of visualization. Here we will just review the visualizations showing figures. According to the dataset, the total number

of cases reported in US is 83M until now. Among them 81M cases are recovered and 1M people died with virus. 1M cases are active as now in the US.

Bar Chart of Covid Cases in US

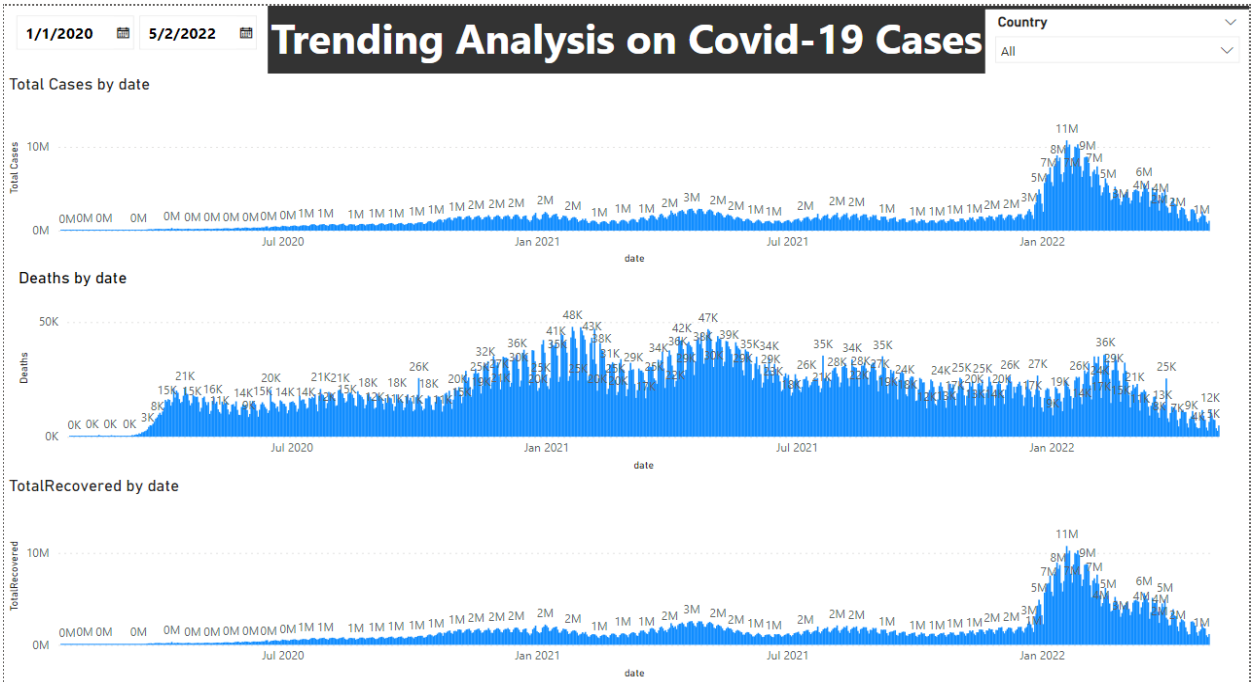
Figure 8



The above visualization is about the number of cases reporting in the cases on year basis. According to the graph showing above total 20M cases are reported in US in 2020, 35M cases in 2021 and 27M cases in 2022. Similarly, second graph shows number of cases recovered in US and third graph shows information related number of deaths happened in the US due to virus. In 2020, 0.5M people died, 0.47M people died in 2021 and 0.17M people are died in 2022.

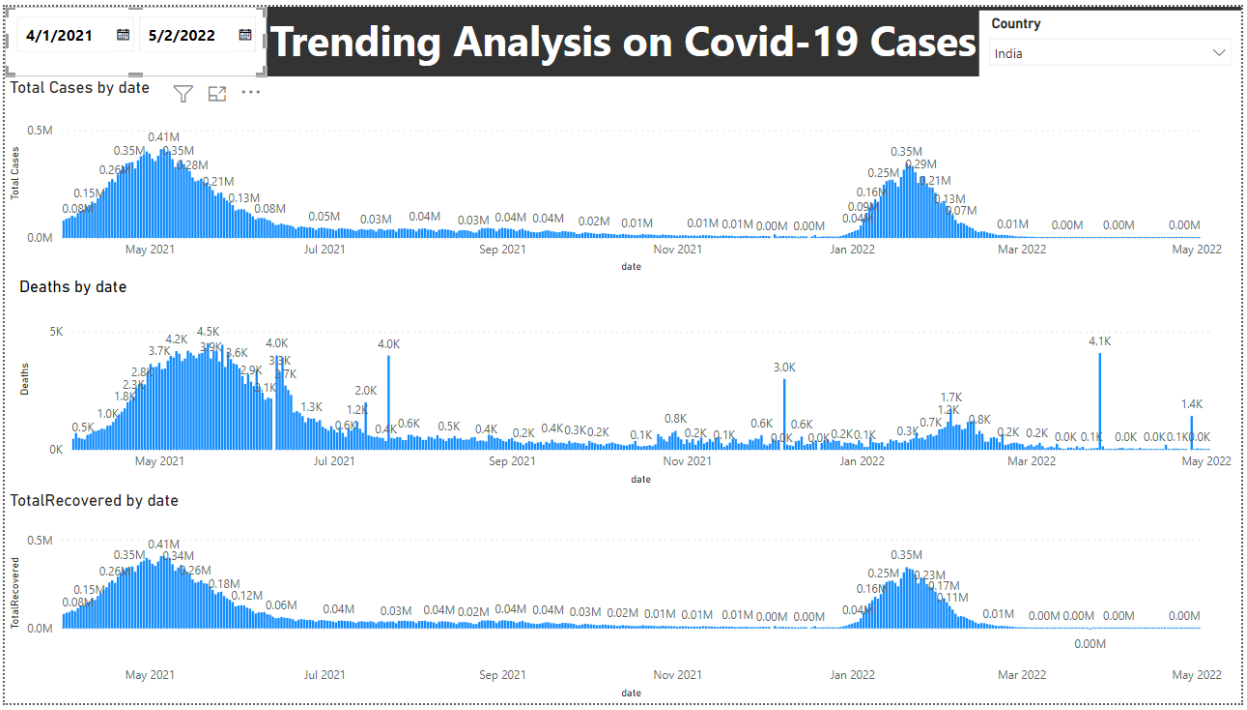
Column Charts of Trend Analysis

Figure 9



The above is the dashboard in the project which shows the trending of cases in the world. This dashboard is the interactive dashboard as it contains two slicers. Slicers means filters on dashboards where we can filter country and dates. The top left slicer for selecting the dates from which date to which date cases need to be plotted and top right slicer for selecting the country to see the cases. This dashboard contains three graphs showing number of cases reporting, number of death cases, number of cases recovered with respective date. Now I will show you the example by selecting India from April 2021 to May 2022.

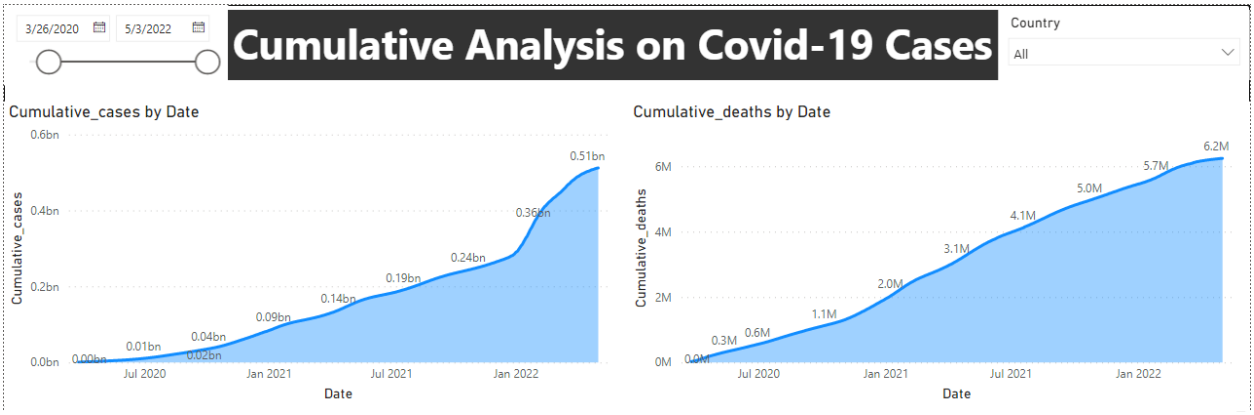
Figure 10



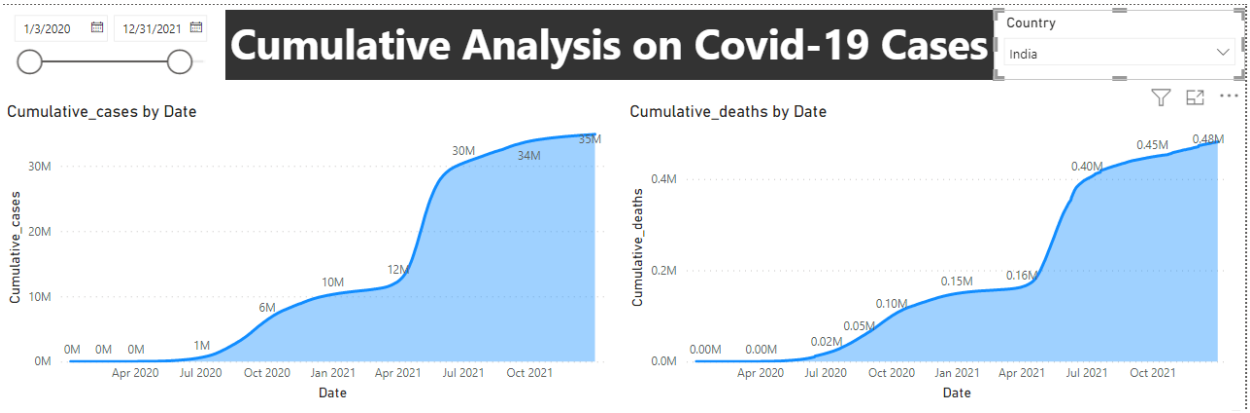
As shown in the above figure 10, date is selected from April 1 2021 to May 2 2022 and selected India from date and country slicer. Then you can see the distribution of cases in India from April 2021 to May 2022. As shown in the figure, cases are went to peak in mid may like 400K cases per day and 4.5K deaths per day. If you clearly observe graph of total cases and recovered cases there slit shift in distribution. Again in January of 2022 cases are raised and again decreased by march. The distribution is like normal distribution.

Area Chart For Cumulative Covid Cases

Figure 11



Area char is combination of both bar chart and line chart. When there is group of labeled data this graph is useful to plot all the groups to have numerical comparation among all of them. The above visualization shows the cumulative cases and cumulative deaths reporting in the world. Cumulative means aggregating all the values of that field to selected date. In this dashboard we kept an option to select date and to select a country as filter or slicers. According to the above figure 11, we can see a sudden raise in cases in 2022 Jan. In addition to that we can small jumps in the number of cases in both total cumulative cases and cumulative deaths. For example we will select a country India and date from Jan 2020 to Dec 2021 and will see the values.



In the above figure showing data related to India from Jan 2020 to Dec 2021. According to the figure there is two jumps in first graph. One is in July 2020 and another is in April 2021. In 2021, cases are jumped from 12M to 30M in July that is within 4 months. The same jumps are happened in deaths also.

Table Showing More Details

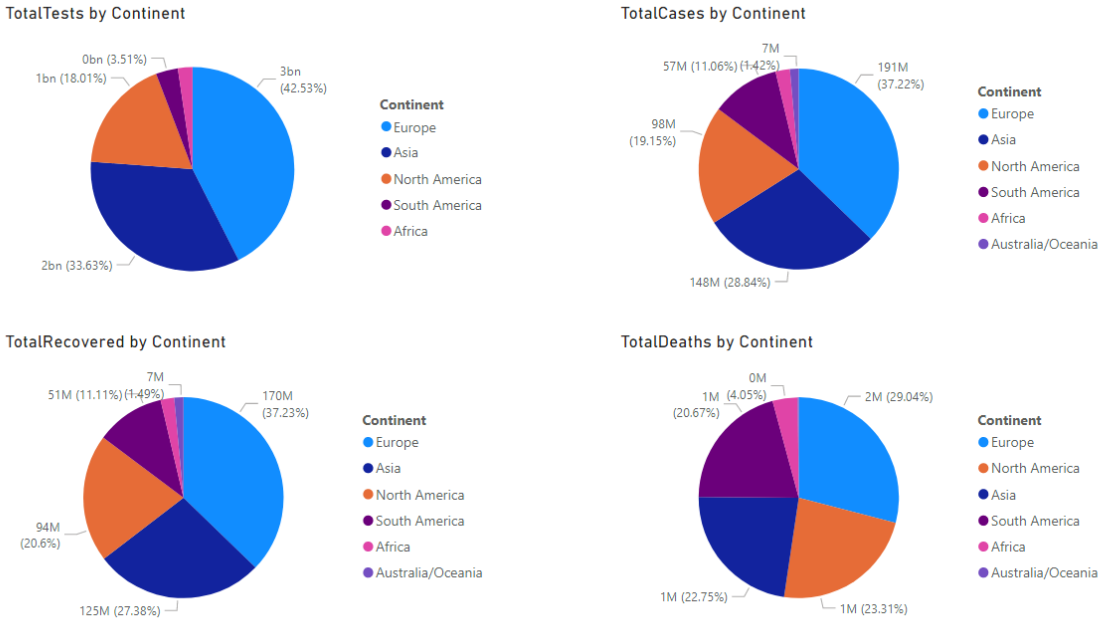
Figure 12

Date_reported	WHO_region	Country_code	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
5/19/2020	SEARO	IN	India	4970	101139	134	3163
5/20/2020	SEARO	IN	India	5611	106750	140	3303
5/21/2020	SEARO	IN	India	5609	112359	132	3435
5/22/2020	SEARO	IN	India	6088	118447	148	3583
5/23/2020	SEARO	IN	India	6654	125101	137	3720
5/24/2020	SEARO	IN	India	6767	131868	147	3867
5/25/2020	SEARO	IN	India	6977	138845	154	4021
5/26/2020	SEARO	IN	India	6535	145380	146	4167
Total				**13205926**	**2189176762**	**168436**	**32681239**

In addition to this there is another visualization in this dashboard which is a table. In this table, columns are Date_reported, WHO_region, Country Code, Country, New_cases, cumulative_cases, New_Cases, Cumulative Cases. It shows the number of new_cases reported on that day, cumulative cases (sum of all cases reported until that day), new deaths happened and cumulative deaths. In the above table information showing data about India as we selected India and May 2020. In addition to this, this table shows the total count of all the numerical columns listed in the table.

Pie Charts for Covid cases-Continent

Figure 13



The above visualizations is pie charts showing the share of cases by each continent. From the above visualization, we can see four pie charts representing total tests, total cases, total recovered cases and total deaths. As per pie charts, Europe is in the 1st place with highest number of tests and with highest number of cases and deaths. The next highest is continent is Asia.

Score Cards of Global Vaccinations

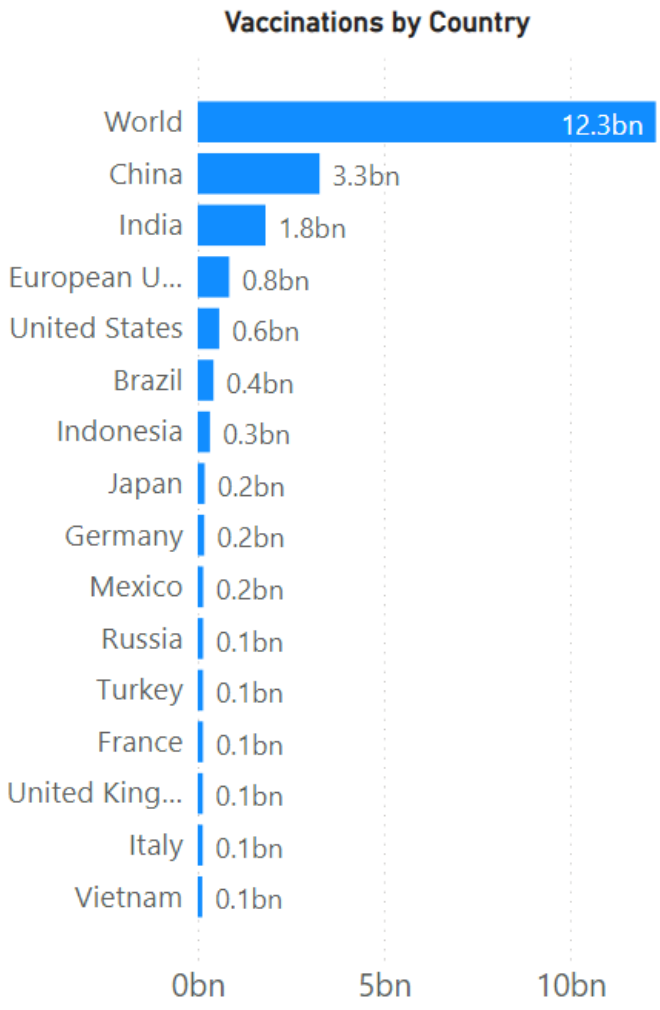
Figure 14

The above score cards shows the statistics about vaccination in the world. According the dataset, In 235 countries vaccine drives are carried out until now. Total number of vaccination

including 1st dose, 2nd dose and booster does until now is around 12bn. In that 5bn people are first dose vaccinations, 5bn are second dose vaccinations and 2bn are booster shot vaccinations.

Bar Chart of Global Vaccinations Country Wise

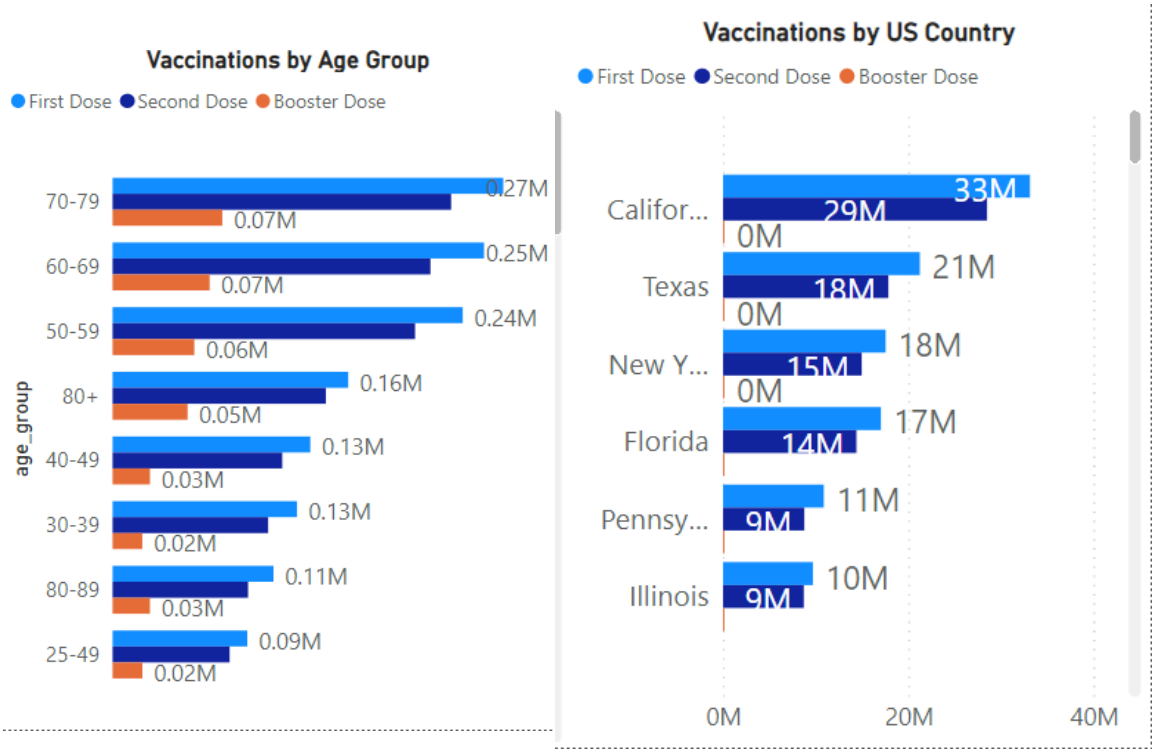
Figure 15



The above graph shows the information related to number of vaccinations in each country. The first bar shows total vaccinations in the world and remaining shows about each country. According to the graph we can say that china is the top 1 country with highest vaccination of 3.3bn. after that India stood second place with 1.8bn doses until now.

Bar Chart for Age Group and US State

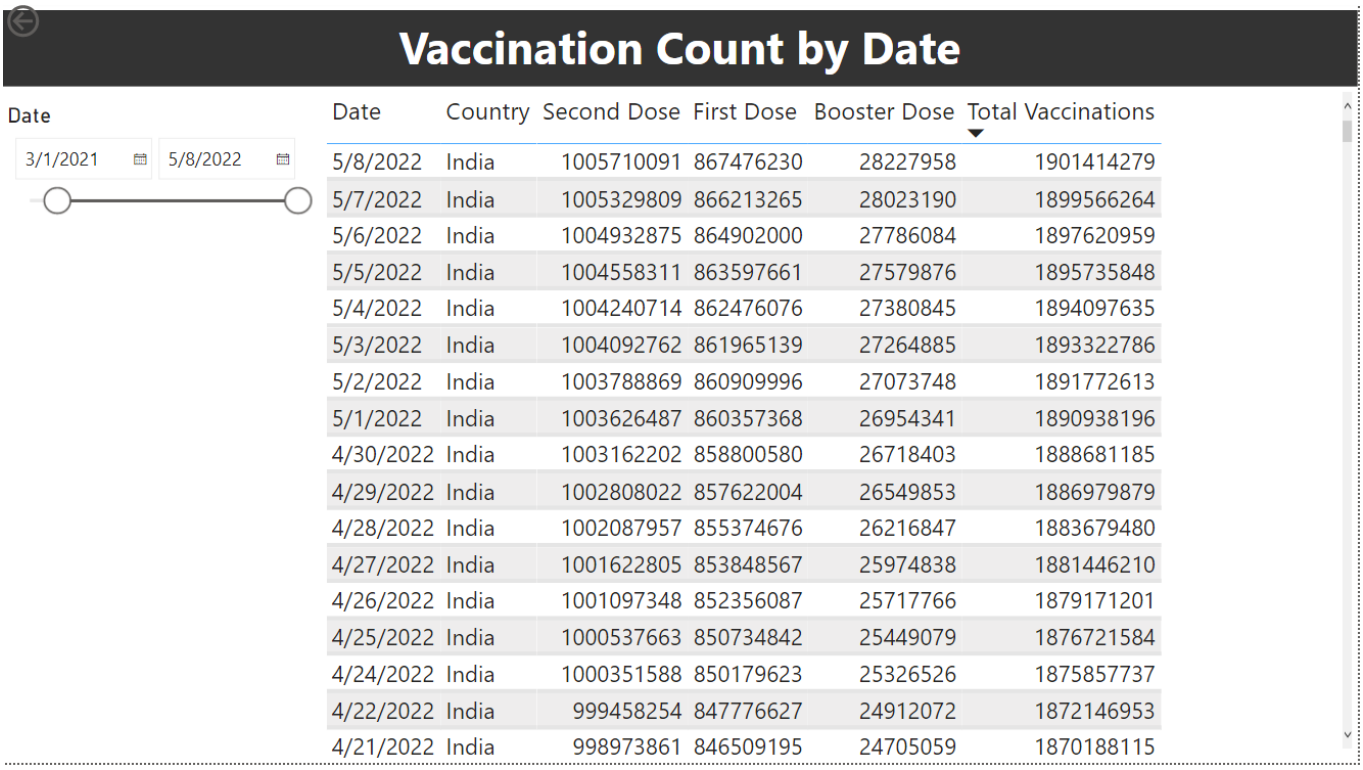
Figure 16



The above two bar charts represents the status of vaccination with respective age group and states in US. In first graph, the y-axis represents the age group like 25-49, 80-89 and etc and x-axis shows the number of vaccinations. The three stacks in each group represents the first dose, second dose and booster dose vaccinations in that age group. As shown in the graph the age group 70-79 stood in the 1st place with highest number of vaccinations. Similarly, the second graph shows the vaccinations in each state of the US. California is the 1st highest state with highest number of vaccinations and Texas stood next to California as 2nd in the list.

Table of Detailed Vaccinations

Figure 17



Vaccination Count by Date

Date

3/1/2021 5/8/2022

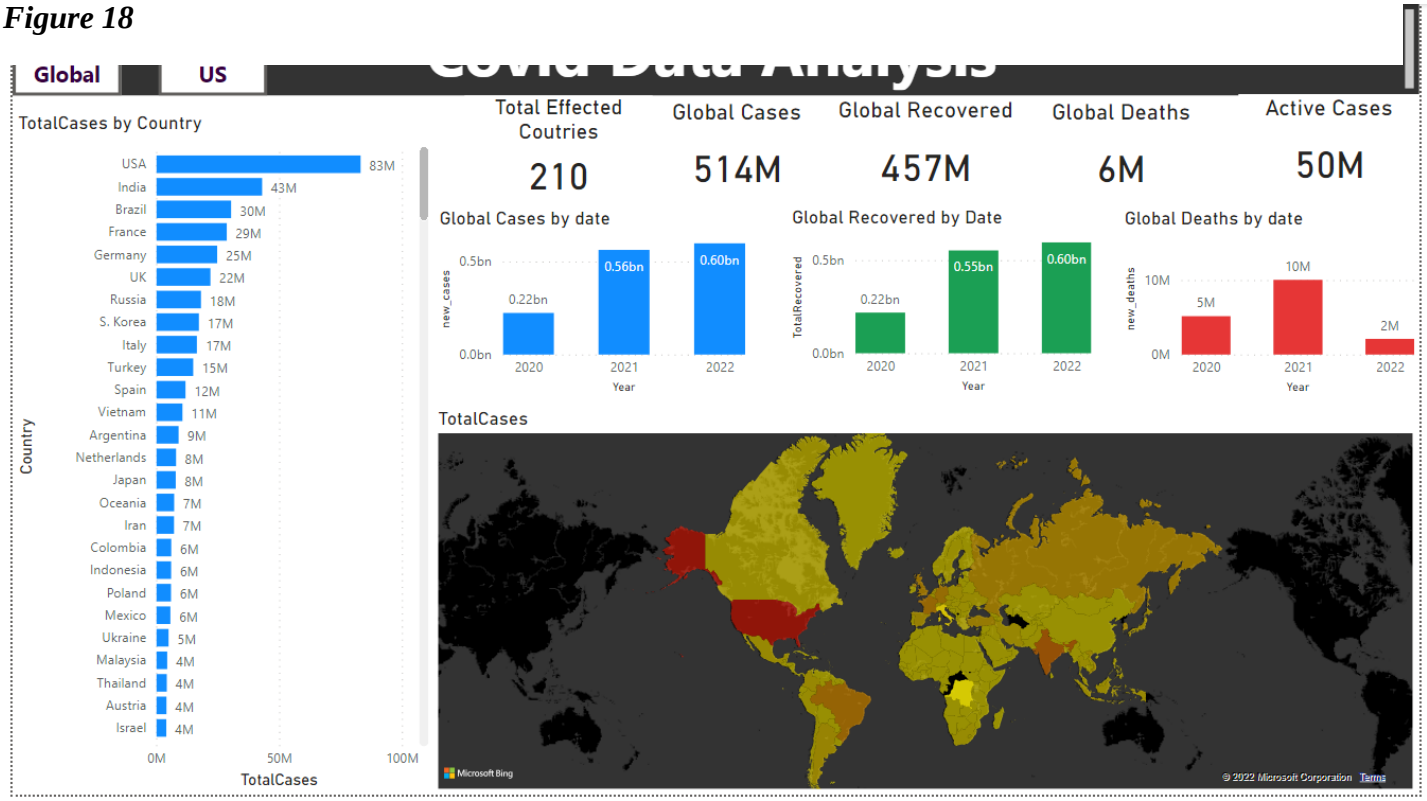
Date	Country	Second Dose	First Dose	Booster Dose	Total Vaccinations
5/8/2022	India	1005710091	867476230	28227958	1901414279
5/7/2022	India	1005329809	866213265	28023190	1899566264
5/6/2022	India	1004932875	864902000	27786084	1897620959
5/5/2022	India	1004558311	863597661	27579876	1895735848
5/4/2022	India	1004240714	862476076	27380845	1894097635
5/3/2022	India	1004092762	861965139	27264885	1893322786
5/2/2022	India	1003788869	860909996	27073748	1891772613
5/1/2022	India	1003626487	860357368	26954341	1890938196
4/30/2022	India	1003162202	858800580	26718403	1888681185
4/29/2022	India	1002808022	857622004	26549853	1886979879
4/28/2022	India	1002087957	855374676	26216847	1883679480
4/27/2022	India	1001622805	853848567	25974838	1881446210
4/26/2022	India	1001097348	852356087	25717766	1879171201
4/25/2022	India	1000537663	850734842	25449079	1876721584
4/24/2022	India	1000351588	850179623	25326526	1875857737
4/22/2022	India	999458254	847776627	24912072	1872146953
4/21/2022	India	998973861	846509195	24705059	1870188115

The above table is the details related to vaccinations of a particular country. There is slicer to select dates. This visualization is linked with the other dashboard visualization as shown in figure 17. If you want to see details of a country then you can click on that bar, right click on that and select drill through. Then you will be redirected to next visualization of figure 14.

Dashboards

Main Page (Dash Board -1)

Figure 18



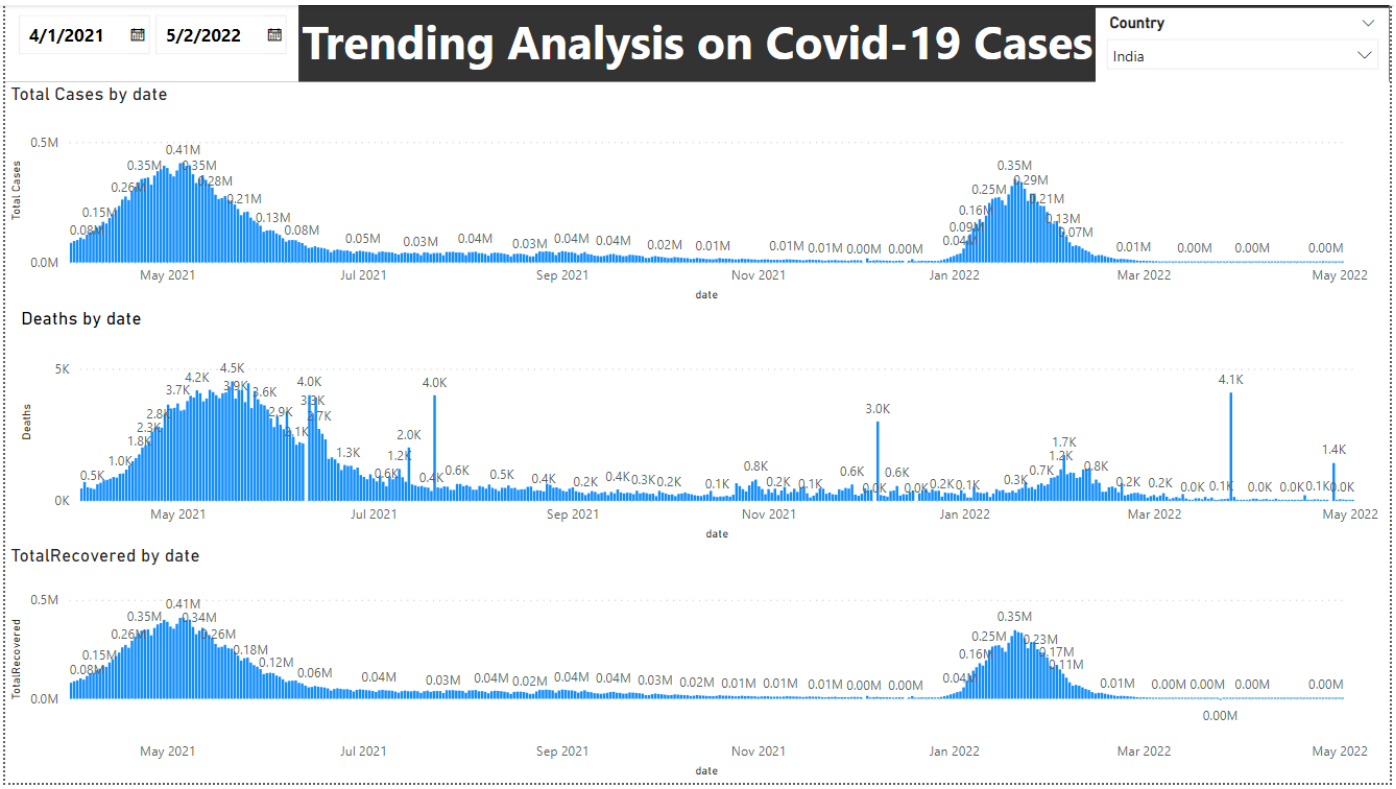
This is a first dashboard which gives a basic information related to covid cases in the world. This contains one text box, two buttons, one bar chart, third chart mini bar charts, five score cards, one map.

By default dashboard shows information related to the global. If you click on US named button then dashboard shows same content for only country US. If you want to see global data again then click on Global named button. Then dashboard goes back to default view. We used bookmarks for this functionality. Bookmarks option present in the Power BI and each visualization will be identified as an object with a specified name. We will link the visibility of visualization to the bookmarks and bookmarks are attached as click functionality to buttons.

Using these bookmarks with objects we can create meaningful on-click functions in power bi as shown in above dashboard.

Trend Analysis (Dash Board-2)

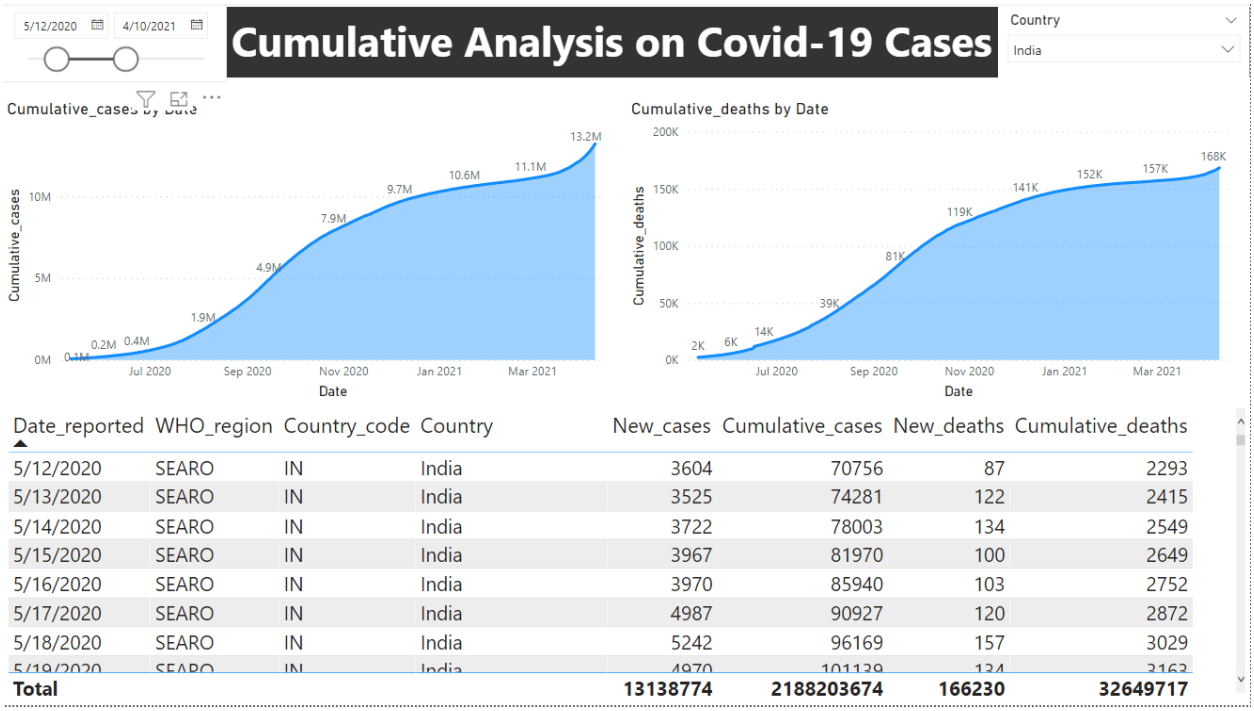
Figure 19



The above dash board is second dashboard showing about the trends of reporting cases, deaths cases, recovered cases. This dashboard contains three bar charts, two slicers (filters) and one text box for heading.

One slicer is useful to filter dates to see the trend from one date to another date. The second slicer is for selecting country. Using these two slicers you can see trend of cases in a particular period of a country.

Cumulative Data (Dashboard -3)

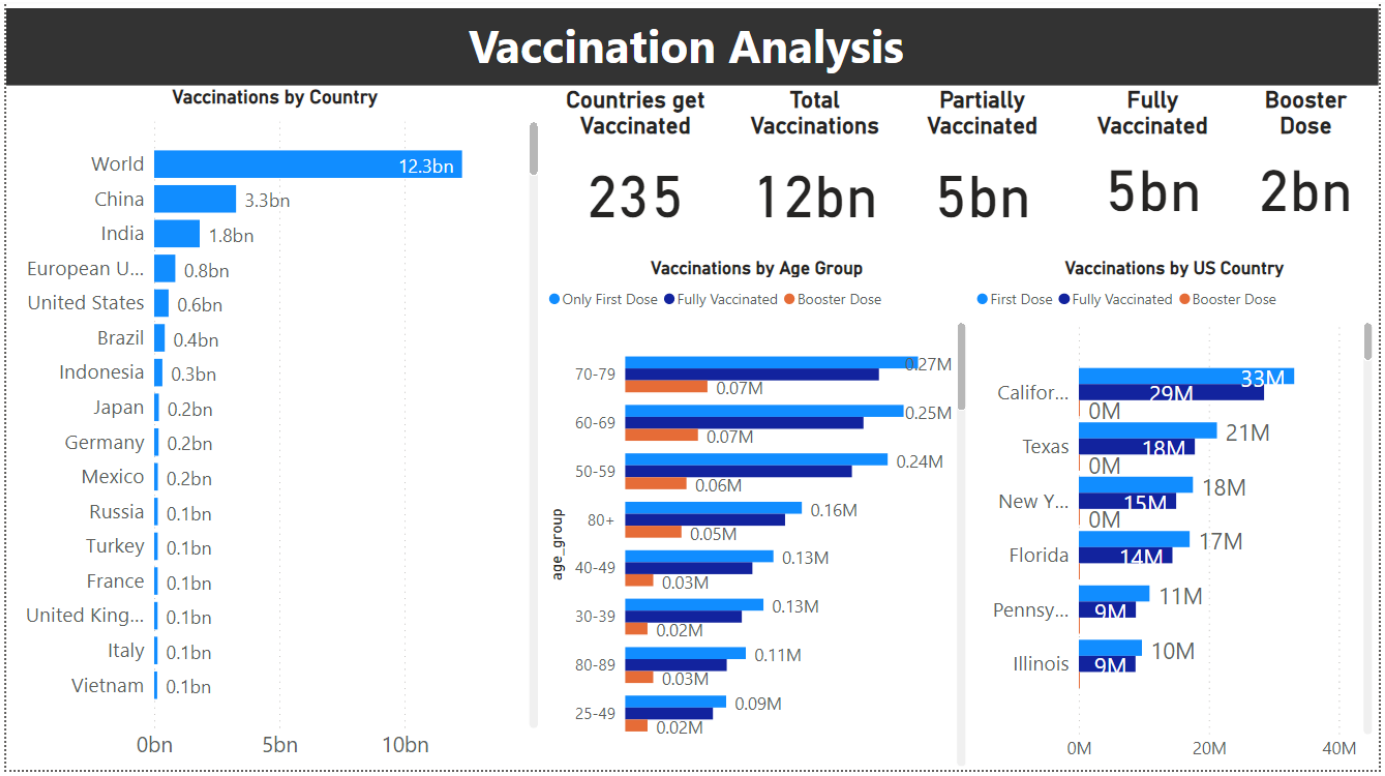


Date_reported	WHO_region	Country_code	Country	New_cases	Cumulative_cases	New_deaths	Cumulative_deaths
5/12/2020	SEARO	IN	India	3604	70756	87	2293
5/13/2020	SEARO	IN	India	3525	74281	122	2415
5/14/2020	SEARO	IN	India	3722	78003	134	2549
5/15/2020	SEARO	IN	India	3967	81970	100	2649
5/16/2020	SEARO	IN	India	3970	85940	103	2752
5/17/2020	SEARO	IN	India	4987	90927	120	2872
5/18/2020	SEARO	IN	India	5242	96169	157	3029
5/19/2020	SEARO	IN	India	4970	101139	134	3163
Total				**13138774**	**2188203674**	**166230**	**32649717**

This is a dashboard reflects the cumulative data about the number cases and deaths happening in a country. This dashboard contains one text box, two slicers, two area charts, one table. Text box is for heading of dashboard. Two slicers are for selecting dates from which date to which date you want to see and other slicer to select the country. One area chart for reporting cases and other area chart for deaths reporting. Table is for detailed information date wise. By this dashboard you can see the jump in cases reporting and deaths in a country.

Vaccination (Dash Board-4)

Figure 20



This dash board is information about vaccination in the world. This dashboard contains one text box, one bar chart, five score cards, two bar stacker bar charts. Bar chart is for showing number of vaccinations done in the whole world and in a country. Five score cards shows statistics about vaccinations. One stacked bar chart shows about the vaccinations in the US by three groups, one is total vaccinations second is only single dose vaccinated and third is booster shot. The second stacked bar chart shows the vaccinations by group wise.

Visualization on Localhost Using Python

In Power BI, there is also a option of integrating python with any step of visualizations. We can use python for preparing the data, manipulating data, and also can create the custom visualization apart from the built-in charts. In this, we are creating a bubble chart which is not there in Power BI using Python. For that we are considering totalcases, countries, totaldeaths,

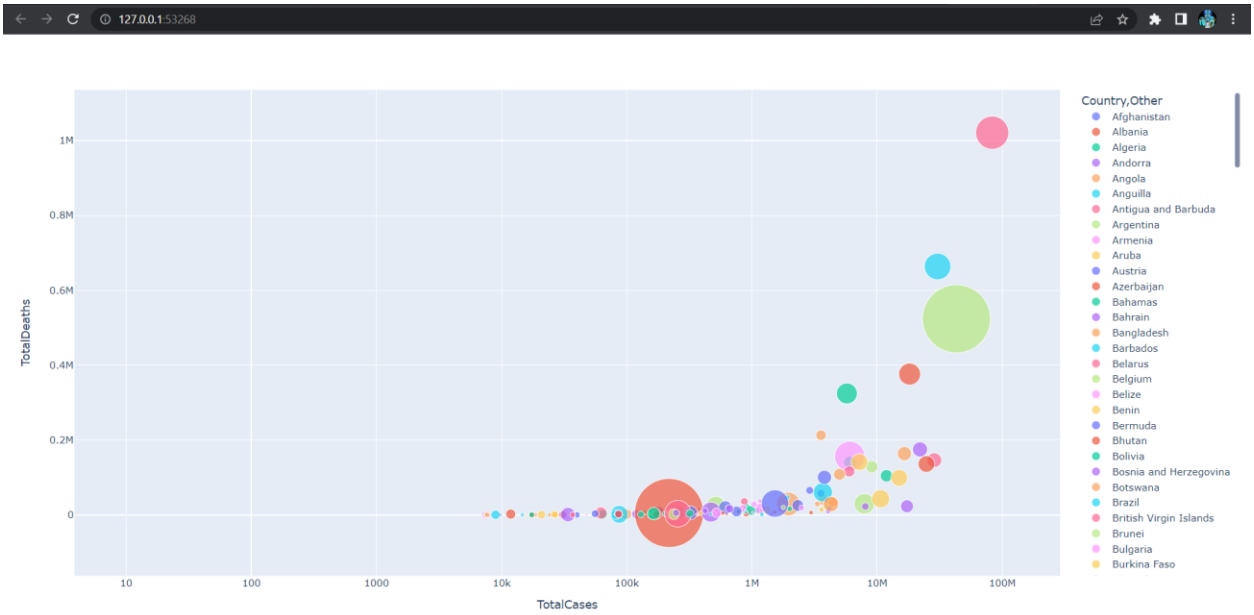
population from the table. We want to create a visualization which shows the deaths in totalcases of each country with respective population of that country.

For that, we need to write a python code in the python script block of Power BI. The below is the code return in the script block.

```
# The following code to create a dataframe and remove duplicated rows is always executed and acts as a preamble for your script:
# dataset = pandas.DataFrame(Country,Other, TotalCases, TotalDeaths, Population)
# dataset = dataset.drop_duplicates()
# Paste or type your script code here:
import plotly.express as px
#df = px.data.gapminder()
dataset.dropna(inplace=True)
fig = px.scatter(dataset,x="TotalCases",y="TotalDeaths",
          size="Population", color="Country,Other",
             hover_name="Country,Other", log_x=True, size_max=60)
fig.show()
```

Using the above code a graph is created as shown in the below figure 21. That is a interactive dashboard where you can click on a bubble to see the information. Size indicates the population of that country, x-axis is Total Cases and y-axis is Total Deaths. From the graph we can say that even with highest population China deaths are very less when compared with other countries. USA is having highest deaths rates among all even with less population than China. So, using this graph we can compare three parameters at a time.

Figure 21


127.0.0.1:53268
TotalDeaths
1M
0.8M
0.6M
0.4M
0.2M
0
10
100
1000
10k
100k
1M
10M
100M
TotalCases
Country,Other
Afghanistan
Albania
Algeria
Andorra
Angola
Anguilla
Antigua and Barbuda
Argentina
Armenia
Aruba
Austria
Azerbaijan
Bahamas
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belize
Benin
Bermuda
Bhutan
Bolivia
Bosnia and Herzegovina
Botswana
Brazil
British Virgin Islands
Brunei
Bulgaria
Burkina Faso

Conclusion

We implemented variety of visualizations on covid-19 and vaccination data collected from two different data source. Covid-19 data is available publicly published by many organizations. We succeeded in getting significant conclusions from the dataset. Those conclusions are useful to make decisions to face covid-19. These analysis of dashboards are useful to understand the current situation of covid-19 in any area. These statistics are useful to understand the data with better insights. These are conclusion we made from our analysis.

- USA is the country with highest number of cases recorded in all the three years.
- India is the second highest number of cases recorded in the world.
- We can see a multimodal distribution of number of cases of recorded, deaths, and recovered.
- In US california is the state with highest in number of cases recorded in any year.
- When we compare Population, Cases reporting, Deaths of each country we can see that countries like China with highest Population is with less cases reporting and less deaths when we compare with the other countries like USA and India.

Until now:

- Total No. of cases reported until now is 517 million among them 6 Million people died with this virus.
- With respective Continents, Europe is with the highest number of cases reported with 42% share.
- Total vaccinations are done throughout the world is 12 billion.

References

Comba, J. L. (2020). Data visualization for the understanding of COVID-19. *Computing in Science & Engineering*, *22*(6), 81-86.

Khanam, F., Nowrin, I., & Mondal, M. R. H. (2020). Data visualization and analyzation of COVID-19. *Journal of Scientific Research and Reports*, *26*(3), 42-52.