Analysis on Situation

Of Covid-19

May 2021

Prepared by

Pavitra Rana

Primary focus: Find the situation of India
Secondary focus: Find Active cases with cared rate and death rate

This project has been done under the guidance of self-learning. This report reflects the views only for the educational purpose and self-assessment which cannot be held responsible for any use which may be made of the information contained therein.

Table of Content

Introduction:

The project report title Analysis on situation of Covid-19 deals with the process of Collection, Storing, organizing, Visualization, Analyzing, and sharing of the data. Coronavirus disease (COVID-19) is an infectious disease caused by a newly discovered coronavirus. Most people infected with the COVID-19 virus will experience mild to moderate respiratory illness and recover without requiring special treatment. Older people, and those with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more likely to develop serious illness.

An analysis is a process of dividing a whole into its separate components for individual examination. It is a very significant aspect in terms of statistical calculation and analytical observation. It is based on Data-driven decision-making for a business strategy that is guided by facts. Facts help to understand the situation in an analytical manner and giving insight to solve the problem. The analysis performs with the sample data by identifying relationships and patterns to find the solution and analyzes the situation. There is a Data Life cycle where analysis has to go through a list of stages following are:

- Planning
- Capture
- Manage
- Analyze
- Archive
- Destroy

There are different stages to be followed to perform the analysis which is as:

- Ask
- Prepare
- Process
- Analyze
- Share
- Act

This report gives an overview of the qualitative analysis and quantitative analysis. The qualitative analysis aims to illustrate and complement the results of the statistical analyses. They are aimed to illustrate and complement the results of the statistical analyses concerning the benefits and their dimension.

- Find out the situation of India where people are infected with the virus;

- Explore the state-wise data of maximum people contacted with the virus;
- Identify the cured and deaths from the total cases with the help of visualization;

To answer these questions, we followed a list of stages of analysis to perform it in a structured manner. The analysis starts with the asking phase where relevant questions being asked to define the problem and plan accordingly. The further processes are for cleaning and storing the raw data to analyze with statistical calculation and analytical thinking to find the optimum solution of the problem. If the analysis is performed correctly and the findings are clear then it can be a move to further process. Moreover, this approach leads to the solution to the problem and concluded by deciding according to the analysis outcomes.

Objective:

The objective of the analysis on the situation of covid-19 is referred to know the situation as:

i. To know the situation of India.
ii. Find the rate of infected people.
iii. To know how covid-19 affected people at a maximum rate.
iv. Know the rate of deaths & cured.
v. How deadly is this virus?

Methodology

This report presents the findings of the qualitative components and quantitative components that perform by collecting a list of raw data of covid-19. The list consists of available columns which are states, total confirmed, cure, deaths, etc. A quality and numeric questions being asked where qualitative questions help in asking relevant questions to understand the character and quality of the data. And Quantitative calculations helps in finding the numeric solutions.

Research Design

The research has been done by following an organized and systematic procedure. An analysis is a very critical process to understand and execute for findings. Before findings, there is a need for extensive research. By following every step or process is very important to compete as per the requirements of the analysis. During the research, the process has been designed with the help of some tools to strategize and complete the project. This report is based on Data-driven decision-making and followed the different stages of the data life cycle to analyze in an organized manner. The following stages are:

- Planning: A Business or organization decides what sort of data needed. And who will be managed and responsible for its optimal outcomes?

- Capture: Data for further analysis is collected from different sources.

- Manage: How and where its data stored using tools to keep it safe and secure.

- Analyze: The gathered Data used to solve the problem and make decisions with the help of various analysis tools.

- Archive: Storing the data in the place where it's still available but may not be used again.

- Destroy: Use erasure tools to destroy data to protect the private files of the organization.

The report consists of gathered data in two ways:

i. Primary data

ii. Secondary data

The steps that are used in this project start from collecting, organizing, storing, manipulating, analyzing, sharing to action.

However, every process includes a dedicated process to go to the next or further steps. Collection of data has been done from primary and secondary sources as well. The next organizing, cleaning, and storing has been done with the help of tools and some codings. Analysis can be done after cleaning data to that make sure there are no errors in the data. Tools are helpful that make the task easy and efficient. It saves time as well as unwanted efforts for lengthy tasks. Analysis performed with the help of some coding to find the problem and get the solutions according to the facts. And there is also a need for analytical and creative skills to deep dive into the problem to get the best possible solution. Therefore, the final stage can be reached to share or communicate with the stakeholders or supervisors with the help of some visualization tools to take actions

Data Collection

Data are collected from both primary and secondary sources and the data is divided into two parts. The qualitative and quantitative parts were carried out with some sort of questions to ask and open source platforms. The data collected from online surveys which were pre-performed in

which there were lists of questionnaires and a list of columns with the number of cases. Besides a series of closed survey questions, the survey questionnaire is about effects and changes and asking from the data about the maximum confirmed cases, death cases, and cured cases that respondents to describe the outcomes and effects (immediate or long term) of their course situation.

Data Analysis

The data were analyzed based on cleaning data, systematic coding, and visualization. This type of analysis consists of cleaning the data first to perform correctly without any null values and systematic coding to identify (practically and theoretically) relevant patterns. Furthermore, the visualization segment helps to find the patterns clearly and easily to analyze effectively.

Analysis & Findings

Analysis

The part of the analysis carried out by following step by step process of the data life cycle where a systematic approach used as:

i. Ask-

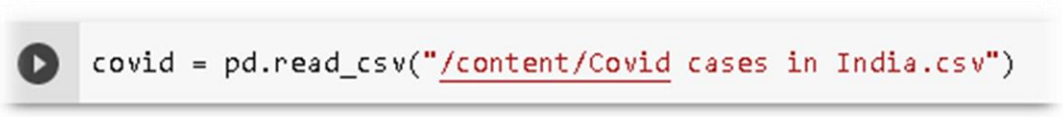
The First step followed by asking relevant questions and defining the problem:

- What is the infection rate of the virus in India?
- What is the death rate in India?
- Are people recovering from this virus? If yes then what is the rate of Cured in India?

ii. Prepare-

The second step is followed by collecting relevant data and storing it. Collected Data has been stored in the cloud and can be retrieved or accessible in the required time. In this step data collected from primary and secondary sources to perform the analysis.

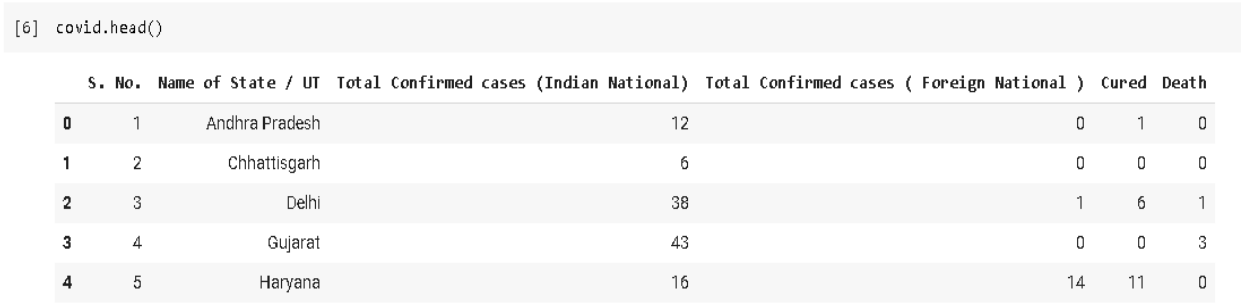
- ❖ To read the Covid_19 data which is given in CSV format using pandas library of python by adding the location path as:



```
covid = pd.read_csv("/content/Covid cases in India.csv")
```

Fig.1

❖ To check the data is reading successfully and also the first five rows of data as:



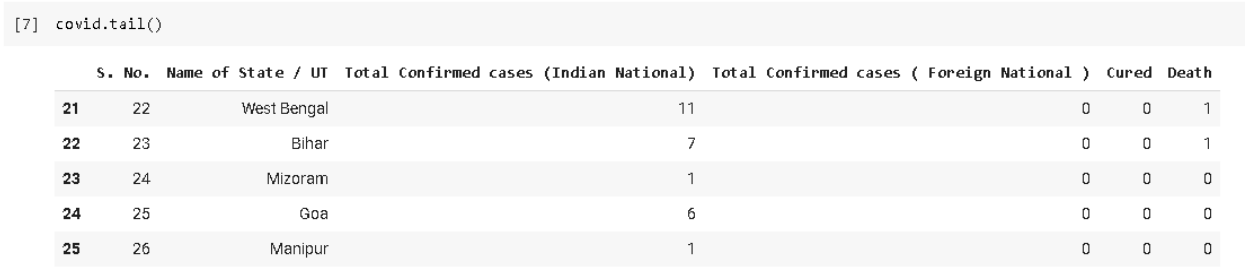
```
[6] covid.head()
```

	S. No.	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death
0	1	Andhra Pradesh	12	0	1	0
1	2	Chhattisgarh	6	0	0	0
2	3	Delhi	38	1	6	1
3	4	Gujarat	43	0	0	3
4	5	Haryana	16	14	11	0

Fig.2

Outcome- The data shows the different columns/attributes and the first five data in sequence order with some values. The columns describe as S.NO, Name of State/UT, Total Confirmed cases (Indian National), Total Confirmed cases (Foreign National), Cured and Death.

❖ To check the last five rows of data as:



```
[7] covid.tail()
```

	S. No.	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death
21	22	West Bengal	11	0	0	1
22	23	Bihar	7	0	0	1
23	24	Mizoram	1	0	0	0
24	25	Goa	6	0	0	0
25	26	Manipur	1	0	0	0

Fig.3

Outcome- The data shows the different columns/attributes and the Last five data in Sequence order with some values. The columns describe as S.NO, Name of State/UT, Total Confirmed cases (Indian National), Total Confirmed cases (Foreign National), Cured and Death.

iii. Process –

The next step is processing the data to clean and check the information. In this step, the information is cleaned by replacing the Null, Nan, na, etc or missing values for further process.

- To check the available columns in the data as:

```
[8]  covid.columns

     Index(['S. No.', 'Name of State / UT',
            'Total Confirmed cases (Indian National)',
            'Total Confirmed cases ( Foreign National )', 'Cured', 'Death'],
           dtype='object')
```

Fig.4

Outcome- The result shows the total number of columns/attributes available in the data. The columns are S.NO, Name of State/UT, Total Confirmed cases (Indian National), Total Confirmed cases (Foreign National), Cured and Death.
The additional information provided as the type of the data is an object.

- For Null values exist in the data set as:

```
[9]  covid.isnull()
```

	S. No.	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
5	False	False	False	False	False	False

Fig.5

Outcome- The result shows the values as FALSE. This Value depicts that there is no Null values. There are true and false values present in the data in which True depicts as Null value and False depicts as Non- Null value.

- To check for any Null values that exist in the dataset as:

```
[11] covid.isnull().any()

S. No.                                          False
Name of State / UT                              False
Total Confirmed cases (Indian National)         False
Total Confirmed cases ( Foreign National )      False
Cured                                           False
Death                                           False
dtype: bool
```

Fig.6

Outcome- This is a result of any Null values that exist in the data set. Shows the values as false which means no Null values present in all columns/attributes.

- To check for the total number of Null values that exist in the datasets as**:**

```
[12] covid.isnull().sum()

S. No.                                          0
Name of State / UT                              0
Total Confirmed cases (Indian National)         0
Total Confirmed cases ( Foreign National )      0
Cured                                           0
Death                                           0
dtype: int64
```

Fig.7

Outcome- This shows the outcome of the total number of Null values present in every column of the data set which is 0. There is additional information provided as the data type id int64.

- Check for Null values in a particular column (Name of State/UT) as:

```
[17] covid[covid['Name of State / UT'].isnull()]
```

	S. No.	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death

Fig.8

Outcome- For a specific or a single column, the Null values present as non-NaN as shown in the figure.

- Check the Shape of the dataset (Rows X Columns) as

```
[19] covid.shape

(26, 6)
```

Fig.9

Outcome- This represents the total number of rows and columns. The Total number of rows is 26 and the total number of columns is 6.

iv. Analyze-

The Analyze step is very crucial, where the cleaned data being analyzed to find patterns, relationships, and trends.

- ❖ Removing the column called 'S.No.' as for sequencing by numbering is available in the data set.

```
[ ] covid.drop(['S. No.'],axis=1,inplace=True)
```

```
covid
```

	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death
0	Andhra Pradesh	12	0	1	0
1	Chhattisgarh	6	0	0	0
2	Delhi	38	1	6	1
3	Gujarat	43	0	0	3
4	Haryana	16	14	11	0
5	Himachal Pradesh	4	0	0	1

Fig.10

Outcome: The results show removed the column S.No. successfully.

- ❖ Find out the Total number of cases of Covid-19 till now in Indi**a**

```
[ ] covid['Total_Cases']=covid['Total Confirmed cases (Indian National)']+covid['Total Confirmed cases ( Foreign National )']
```

```
covid
```

	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death	Total Cases
0	Andhra Pradesh	12	0	1	0	12
1	Chhattisgarh	6	0	0	0	6
2	Delhi	38	1	6	1	39
3	Gujarat	43	0	0	3	43
4	Haryana	16	14	11	0	30
5	Himachal Pradesh	4	0	0	1	4
6	Karnataka	20	0	3	2	20
7	Kerala	131	7	11	0	138
8	Madhya Pradesh	23	0	0	1	23

Fig.11

Outcome: The result of Total cases retrieved by sum up the Total Confirmed cases (Indian National) and Total Confirmed cases (Foreign National).

❖ Find out the Active cases of Covid-19 till now in India

```
[ ] covid['Active_Cases']=covid['Total_Cases']-(covid['Death']+covid['Cured'])
```

```
[ ] covid
```

	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death	Total Cases	Active Cases
0	Andhra Pradesh	12	0	1	0	12	11
1	Chhattisgarh	6	0	0	0	6	6
2	Delhi	38	1	6	1	39	32
3	Gujarat	43	0	0	3	43	40
4	Haryana	16	14	11	0	30	19
5	Himachal Pradesh	4	0	0	1	4	3
6	Karnataka	20	0	3	2	20	15
7	Kerala	131	7	11	0	138	127
8	Madhya Pradesh	23	0	0	1	23	22

Fig.12

Outcome: The result of Active cases retrieved by sum up the Death and Cured and subtracted by Total cases. Active cases showed all the state-wise numbers in the dataset.

❖ Visualize the maximum values or cases of all available columns

```
covid.style.background_gradient(cmap='Reds')
```

	Name of State / UT	Total Confirmed cases (Indian National)	Total Confirmed cases (Foreign National)	Cured	Death	Total Cases	Active Cases
0	Andhra Pradesh	12	0	1	0	12	11
1	Chhattisgarh	6	0	0	0	6	6
2	Delhi	38	1	6	1	39	32
3	Gujarat	43	0	0	3	43	40
4	Haryana	16	14	11	0	30	19
5	Himachal Pradesh	4	0	0	1	4	3
6	Karnataka	20	0	3	2	20	15
7	Kerala	131	7	11	0	138	127
8	Madhya Pradesh	23	0	0	1	23	22
9	Maharashtra	144	3	15	4	147	128
10	Odisha	3	0	0	0	3	3
11	Puducherry	1	0	0	0	1	1
12	Punjab	29	0	0	1	29	28
13	Rajasthan	41	2	3	0	43	40
14	Tamil Nadu	32	3	1	1	35	33
15	Telengana	34	11	1	0	45	44
16	Chandigarh	7	0	0	0	7	7
17	Jammu and Kashmir	18	0	1	1	18	16
18	Ladakh	13	0	0	0	13	13
19	Uttar Pradesh	42	1	11	0	43	32
20	Uttarakhand	4	0	0	0	4	4
21	West Bengal	11	0	0	1	11	10
22	Bihar	7	0	0	1	7	6
23	Mizoram	1	0	0	0	1	1
24	Goa	6	0	0	0	6	6
25	Manipur	1	0	0	0	1	1

Fig.13

Outcome: The chart of the data shown in Fig.9 visualizes the maximum values as dark red and minimum values followed by light red. Data shows the maximum values confirmed cases in Kerala and Maharashtra where cured are 11 and 15. However, Total cases and Active cases are maximum in Maharashtra and Kerala.

- ❖ Find out the Active cases in ascending order or maximum order state wise of Covid-19 till now in India

```
[ ] Total_Active_Cases=covid.groupby('Name of State / UT')['Total_Cases'].sum().sort_values(ascending=False).to_frame()
```

```
[ ] Total_Active_Cases
```

Name of State / UT	Total Cases
Maharashtra	147
Kerala	138
Telengana	45
Uttar Pradesh	43
Rajasthan	43
Gujarat	43
Delhi	39
Tamil Nadu	35
Haryana	30
Punjab	29
Madhya Pradesh	23
Karnataka	20
Jammu and Kashmir	18
Ladakh	13
Andhra Pradesh	12
West Bengal	11
Chandigarh	7
Bihar	7
Goa	6
Chhattisgarh	6
Uttarakhand	4
Himachal Pradesh	4
Odisha	3
Manipur	1
Mizoram	1
Puducherry	1

Fig.14

Outcome: The result arranged the data in ascending order and shows the maximum number of Total cases state-wise.

- ❖ Visualize the maximum values of Total cases state wise

```
[ ]  Total_Active_Cases.style.background_gradient(cmap='Reds')
```

Name of State / UT	Total Cases
Maharashtra	147
Kerala	138
Telengana	45
Uttar Pradesh	43
Rajasthan	43
Gujarat	43
Delhi	39
Tamil Nadu	35
Haryana	30
Punjab	29
Madhya Pradesh	23
Karnataka	20
Jammu and Kashmir	18
Ladakh	13
Andhra Pradesh	12
West Bengal	11
Chandigarh	7
Bihar	7
Goa	6
Chhattisgarh	6
Uttarakhand	4
Himachal Pradesh	4
Odisha	3
Manipur	1
Mizoram	1
Puducherry	1

Fig.15

Outcome: The result arranged the data in ascending order and shows the maximum number of Total cases state-wise with gradient background to visualize critical situation in color red.

❖ Visualize the maximum values of Total cases state wise in the Bar graph

```
[30] #Pandas vis
     covid.plot(kind='bar',x='Name of State / UT',y='Total_Cases')
     plt.show()
```

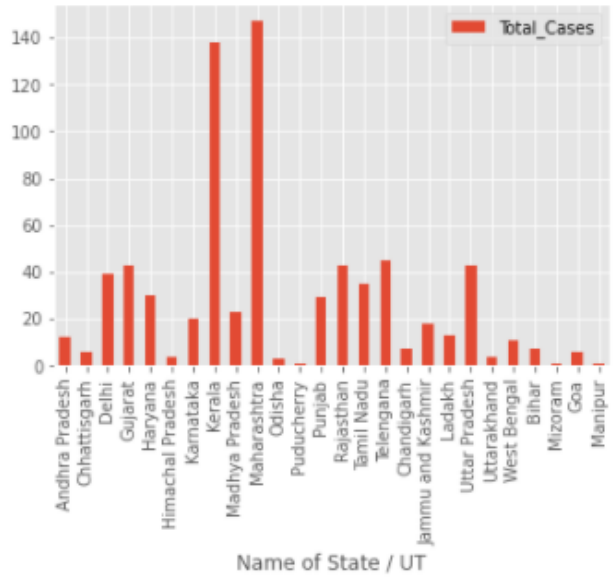


Fig.16

Outcome: The result arranged the data in a bar graph to visualize the maximum cases among the states. The maximum values shown in the Fig.16 of Kerala and Maharashtra.

v. Share-

The next step is to share where the analyzed data have to share with shareholders and communicate by visualizing and presenting the matter. Communication is very important between the shareholders to interpreting the analysis results. So, that they can take decisions for further process.

vi. Act-

The final step of the analysis is act or action. In this step, the action can be taken by using the analysis results. The decision will be taken after discussion according to the presentation or visualization of the analysis and findings.

Findings

According to the analysis the findings following are:

- The total number of cases is high in Kerala and Maharashtra was the lowest in Manipur.
- The cured rate is low but maximum cured among Haryana, Kerala, and Maharashtra.
- The death rate is high in Gujrat and lowest cases in Manipur
- According to data, the situation of Kerala and Maharashtra is critical as there is a high number of cases in total and the cured rate is also very low.

Conclusions

According to the analysis on Covid 19 situation, the quantitative analysis shows, the data results of total cases, active cases, death rate, and cured rate in terms to know the situation. To know the situation of the Covid 19 the analysis includes some systemically coding and visualization with the help of python libraries.
The qualitative analysis was based on some research questions as
1.What is the situation in India?
2.How many are infected with the virus or active cases?
3.What are the death rate and cured rate?

The analysis shows the critical situation in some states of India that need a crucial action of a plan to decrease the strength of the virus and prevent it with possible alternatives.

Based on the qualitative and quantitative analysis we can draw several conclusions:

- According to data, the situation of Kerala and Maharashtra is critical as there is a high number of cases in total and the cured rate is also very low.
- The cured rate is low but the maximum cured among all given states are Haryana, Kerala, and Maharashtra which is not a good metric according to the total number of cases.
- The metrics of death rate is high in Gujrat and lowest cases in Manipur so it is low as compared to active cases and cured cases.
- For a long-term cure plan, there is a need for an action plan or some possible alternatives as soon as possible to prevent this deadly virus.

References

1. https://en.wikipedia.org/wiki/Analysis
2. https://en.wikipedia.org/wiki/Data_analysis
3. https://www.kaggle.com/datasets?search=covid+cases+in+india
4. https://www.w3schools.com/python/matplotlib_intro.asp
5. https://www.w3schools.com/sql/sql_intro.asp
6. https://www.w3schools.com/python/pandas/pandas_csv.asp
7. https://www.w3schools.com/python/numpy/numpy_getting_started.asp
8. https://www.w3schools.com/python/numpy/numpy_random_seaborn.asp
9. https://www.who.int/emergencies/diseases/novel-coronavirus-2019
10. https://colab.research.google.com/notebooks/intro.ipynb#recent=true
11. https://docs.python.org/3/library/
12. https://www.who.int/health-topics/coronavirus#tab=tab_1