

Factorial Analysis and visualization of covid-19 clinical data utilizing of R

Venu Paritala^{1*}, Harsha Thummala¹

¹Department of Biotechnology, Vignan Foundation for Science Technology and Research, India

ABSTRACT

Information visualization and examination plays a key wedge in high-throughput science. It is a fundamental apparatus for data exploration permitting to shed light on information structure and designs of intrigued. Visualization is also vital significance as a shape of communicating information to a wide gathering of people. Data investigations find important experiences that drive more intelligent and more viable commerce choices. Here, provided factorial analysis, statistical analysis and visualization of Covid 19 data. Factor analysis is a technique that is used to reduce a large number of variables into fewer numbers of factors. This technique extracts maximum common variance from all variables and puts them into a common score. As an index of all variables, we can use this score for further analysis. The analysis also proves results out the mean, median and quality of each variable. We exhibit a rundown of R's plotting frameworks they are utilized to imagine and understand crude like corona cases in world.

Keywords: Bioinformatics, Data analysis, Programming, R, Visualization, covid-19

INTRODUCTION

Visualization plays an fundamental part in high-throughput biology. It could be a implies to investigate, get it, and communicate data in a way graphic factual properties can regularly hardly compete with. Whereas these contrasts gotten to be clear upon assessment of the relapse coefficients, the x-y diffuse plots (ggplot2, ggfortify packages , see underneath) outline the striking differences between these information which as it were the primary data set can be demonstrated by factorize and summarize .R analytic is a free, open-source program utilized for all sorts of data science . It exceeds expectations in statistical data mining and modeling and has exceptionally powerful graphics offices to uncover particular designs in information. R is an open-source venture that has been received in different areas in scholarly inquire about and industry. In initiation visualization utilizing R, we are going give an overview of R's graphical frameworks and how they are connected to analyze and visualization of covid19 clinical data. A companion vignette complements this work with additional visualizations and provides the detailed code allowing readers to reproduce the figures and experiment with the plotting functions described hereafter. The data sets that are already read in R. Here, we shall

be using the covid-19 clinical data set that come built-in R .While using any external data source; we can use the read command to load the files (Excel, CSV, HTML and text files etc.)This data set is also available at WHO. To be start downloads the data set, both train and test file. The vignette too gives common assets almost R graphics and visualization. The detailed code is accessed by installing ggplot2, ggfortify and psych as described on the package landing page and typing:

Library (ggplot2)

Library (ggfortify)

Library (tidyverse)

this is "ggplot2" Version 3.3.3

To get start visit

L <https://ggplot2.tidyverse.org>, <https://github.com/tidyverse/ggplot2>

This is "ggfortify" Version 0.4.11

To get start visit

*Corresponding to: Venu Paritala, Dept of Biotechnology, Vignan Foundation for Science Technology and Research, India; E-mail: venuparitala@gmail.com

Received date: July 06, 2021; Accepted date: October 09, 2021; Published date: October 18, 2021

Citation: Paritala V and Thummala H (2021) Factorial Analysis and visualization of covid-19 clinical data utilizing of R. J Clin Exp Cardiol 12:p14

Copyright: © 2021 Paritala V et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

L <https://ggplot2.tidyverse.org>, <https://github.com/tidyverse/ggplot2>

This is "ggfortify" Version 1.3.0

To get start visit

<https://tidyverse.tidyverse.org>, <https://github.com/tidyverse/tidyverse>

There are thousands of add-on bundles on the Comprehensive R File Organize (CRAN) servers, transparently and freely contributed by designers around the world. These bundles can be effectively introduced and utilized to complement R's out-of-the box usefulness, they offer novel factual functionality, extended back visualization (counting intuitively visualziation) as well as foundation committed to the examination and visualization of organic information.As of November 2020, more than 16,000 packages is available in R.R for the analysis of covid19 data. In this manuscript, we will focus on the analyses and visualization capabilities of R itself (Section 3 and R add-on packages. In this manuscript to analyses and visualization of covid-19 cases data. These analysis and visualization using different kind of in R. (1) ggplot2 is a framework for decoratively making design, based on the Language structure of Design. Ggplot2 how to outline factors to aesthetics, what graphical primitives to use. (2)ggfortify: Information Visualization Instruments for Factual Investigation Comes about. Bound together plotting devices for insights commonly utilized, such as GLM, time arrangement, PCA families, clustering and survival investigation. The package offers a single plotting interface for these examination comes about and plots in a bound together style using ggplot2. (3) Tidy verse is a collection of fundamental R packages for information science. The packages under the tidy verse umbrella offer assistance us in performing and connection with the information. There are an entire have of things you'll do along with your information, such as sub setting, changing, visualizing.

As the corona virus far reaching rapidly clears over the world, it is activating a noteworthy degree of fear, push and concern inside the people at gigantic and among certain bunches in particular, such as more prepared grown-ups, care providers and people with fundamental prosperity conditions. A plenty of work has been done on visualizing COVID-19 information since the episode of the widespread. Those visualizations can be basically classified into two bunches: visualization with zero or less interactivity speaks to the primary gather, and complex intelligently visualization methods and devices speak to the moment. Active visualization [1, 2] employments essential charts such as bar and line charts for speaking to traits of COVID-19 data set. The utilized visualization and examination procedures for overseeing particular data with live upgrading design as takes after: (1) Interactive direction chart for number of cases and deaths, (2) Sunburst was utilized for breaking the effect of infection by nations, (3)Searchable line chart for number of affirmed cases and recuperated cases, (4)World outline charts for number of cases and passing.

METHODS

Before read the dataset first set the directories and instal and load the pakages. Next read the dataset using fallowing commands given below.

```
library(ggplot2)
```

```
library(ggfortify)
```

```
library(tidyverse)
```

```
# read the dataset
```

```
train<-read.csv("covid19.csv",header = T)
```

```
test<-read.csv("covid19.csv",header = T)
```

Data Analysis

We are going to analysis stat with test datatype. Names () the command used for helps us in checking out all the variables in the data set. Familiarization with all the Variables/Column Names. The result provides [Alarithem 1]. **str()** This helps in understanding the structure of the data set, data type of each attribute and number of rows and columns present in the data. **Summary ()** is one of the most important functions that help in summarizing each attribute in the data set. It gives a set of descriptive statistics, depending on the type of variable, In case of a Numerical Variable -> Gives Mean, Median, Mode, Range and Quartiles. In case of a Factor Variable -> gives a table with the frequencies. In case of Factor + Numerical Variables -> Gives the number of missing values. In case of character variables -> gives the length and the class. In case we just need the summary statistic for a particular variable in the data set.

```
test<-read.csv("covid19.csv",header = T)
```

```
name(test)
```

```
[1] "i..Name"
```

```
[2] "WHO.Region"
```

```
[3] "Cases...cumulative.total"
```

```
[4] "Cases...cumulative.total.per.100000.population"
```

```
[5] "Cases...newly.reported.in.last.7.days"
```

```
[6] "Cases...newly.reported.in.last.7.days.per.100000.population"
```

```
[7] "Cases...newly.reported.in.last.24.hours"
```

```
[8] "Deaths...cumulative.total"
```

```
[9] "Deaths...cumulative.total.per.100000.population"
```

```
[10] "Deaths...newly.reported.in.last.7.days"
```

```
[11] "Deaths...newly.reported.in.last.7.days.per.100000.population"
```

```
[12] "Deaths...newly.reported.in.last.24.hours"
```

```
[13] "Transmission.Classification"
```

```
str(test)
```

```
'data.frame': 237 obs. of 13 variables:
```

\$ i..Name : chr "Global" "United States of America" "Brazil" "India" ...
 \$ WHO.Region : chr "" "Americas" "Americas" "South-East Asia" ...
 \$ Cases...cumulative.total : int 133552774 30615849 13193205 13060542 4862115 4623984 4370325 3717602 3689866 3336637 ...
 \$ Cases...cumulative.total.per.100000.population : num 1711 9249 6207 946 7476 ...
 \$ Cases...newly.reported.in.last.7.days : int 4293881 451664 444458 757411 241758 60928 20055 110519 331878 36921 ...
 \$ Cases...newly.reported.in.last.7.days.per.100000.population : num 55 136.4 209.1 54.9 371.7 ...
 \$ Cases...newly.reported.in.last.24.hours : int 733397 74849 92625 131968 23761 9150 3030 17209 55941 5279 ...
 \$ Deaths...cumulative.total : int 2894295 553801 340776 167642 97430 102247 126980 112861 33201 76179 ...
 \$ Deaths...cumulative.total.per.100000.population : num 37.1 167.3 160.3 12.2 149.8 ...
 \$ Deaths...newly.reported.in.last.7.days : int 73187 6857 19261 4246 2064 2614 216 3014 1488 276 ...
 \$ Deaths...newly.reported.in.last.7.days.per.100000.population: num 0.938 2.07 9.06 0.31 3.17 ...
 \$ Deaths...newly.reported.in.last.24.hours : int 12924 873 3829 780 331 402 53 487 258 25 ...
 \$ Transmission.Classification : chr "" "Community transmission" "Community transmission" "Clusters of cases" ...

Summary(test)

i..Name WHO.Region

Length:237 Length:237

Class :character Class :character

Mode :character Mode :character

Cases...cumulative.total

Min. : 0

1st Qu.: 3028

Median : 29385

Mean : 1127028

3rd Qu.: 242848

Max. :133552774

Cases...cumulative.total.per.100000.population

Min. : 0.0

1st Qu.: 145.8

Median : 1359.1

Mean : 2759.5

3rd Qu.: 4874.0

Max. :16060.3

NA's :1

Cases...newly.reported.in.last.7.days

Min. : 0

1st Qu.: 29

Median : 680

Mean : 36235

3rd Qu.: 7239

Max. :4293881

Cases...newly.reported.in.last.7.days.per.100000.population

Min. : 0.00 1st Qu.: 1.77 Median : 18.80 Mean : 90.88 3rd Qu.: 123.94 Max. :1211.51 NA's :1 Cases...newly.reported.in.last.24.hours

Min. : 0

1st Qu.: 0

Median : 64

Mean : 6189

3rd Qu.: 1115

Max. :733397

Deaths...cumulative.total

Min. : 0

1st Qu.: 29

Median : 450

Mean : 24424

3rd Qu.: 4108

Max. :2894295

Deaths...cumulative.total.per.100000.population

Min. : 0.000

1st Qu.: 1.635

Median : 15.355

Mean : 45.322

3rd Qu.: 74.315

Max. :279.010

NA's :1

Deaths...newly.reported.in.last.7.days

Min. : 0.0

1st Qu.: 0.0

Median : 8.0

Mean : 617.6

3rd Qu.: 98.0
 Max. :73187.0
 Deaths...newly.reported.in.last.7.days.per.100000.population
 Min. : 0.000
 1st Qu.: 0.000
 Median : 0.235
 Mean : 1.496
 3rd Qu.: 1.403
 Max. :19.130
 NA's :1
 Deaths...newly.reported.in.last.24.hours
 Min. : 0.0
 1st Qu.: 0.0
 Median : 1.0
 Mean : 109.1
 3rd Qu.: 16.0
 Max. :12924.0

Transmission.Classification

Length:237

Class :character

Mode :character

The dataset having the 237 observations of 13 variables. Text data type is known as Strings, Strings can contain numbers and / or characters. A string can also contain or consist of numbers. For instance, '1234' could be stored as a string, as could '10.23'. However strings that contain numbers cannot be used for mathematical operations. The foremost common dtype. Will be relegated to your column on the off chance that column has blended sorts (numbers and strings). Int64 int Numeric characters. 64 alludes to the memory designated to hold this character. There are several kinds of mean in mathematics, especially in statistics. For a data set, the arithmetic mean, also known as average or arithmetic average, is a central value of a finite set of numbers. The covid cases reported in countries in cumulative total Minimum reports is 0, Maximum. : 133552774.the covid19 Cases newly reported in last 7 days maximum approximately 4293881 and Cases newly reported in last 24 hours maximum approximately 733397. Covid19 death newly reported in last 7 days maximum approximately 73187.0 and death newly reported in last 24 hours maximum approximately 12924.07.The death reported in countries in cumulative total Minimum reports is 0, Maximum is 2894295.The data set of each column heading present in the (algorithm 1).The data set analyse the death and cases reported to calculated the mean and median and find out the averages. The data analysis part summarize the data is main role .the data-set contain the Cases newly reported in last 7 days per 100000.population comparison of 100000 of population 90%

of the population causes. The Deaths newly reported in last 7 days per100000. Population comparison of 100000 populations 1.46 population causes death report. According the icmr survey result mach my data analysis report. Cases reported every 100000 of population in 2% and death report also 1.46% hence out of these result every 100000 reported covid19 at 3.146%.

Factorial analysis

Factor analysis is a technique that is used to reduce a large number of variables into fewer numbers of factors. This technique extracts maximum common variance from all variables and puts them into a common score. As an index of all variables, we can use this score for further analysis. Exploratory factor analysis (EFA) is generally used to discover the factor structure of a measure and to examine its internal reliability. EFA is often recommended when researchers have no hypotheses about the nature of the underlying factor structure of their measure. The aim of this work is to identify main impact factors affecting variations in the covid19 data set. To study such a complex system, an objective method combining various approaches (statistics, R, GIS, descriptive analysis and graphical plotting) was performed. Methodology of the research includes following clusters, R programming language for writing codes, statistical analysis, and mathematical algorithms for data processing, analysis and visualizing diagrams.

```
x<-factorial(test$Cases...newly.reported.in.last. 24.hours)
```

```
x
```

- [1] Inf Inf Inf
- [4] Inf Inf Inf
- [7] Inf Inf Inf
- [10] Inf Inf Inf
- [13] Inf Inf Inf
- [16] Inf Inf Inf
- [19] Inf Inf Inf
- [22] Inf Inf Inf
- [25] Inf 2.000000e+00 Inf
- [28] Inf Inf Inf
- [31] Inf Inf Inf
- [34] Inf Inf Inf
- [37] 5.040000e+03 Inf Inf
- [40] Inf Inf Inf
- [43] Inf Inf Inf
- [46] Inf Inf Inf
- [49] Inf 1.000000e+00 Inf
- [52] Inf Inf Inf
- [55] Inf Inf Inf
- [58] Inf Inf Inf

[61] Inf Inf Inf [178] 1.000000e+00 3.628800e+05 1.000000e+00
 [64] Inf Inf Inf [181] 3.628800e+06 3.991680e+07 1.000000e+00
 [67] Inf 1.000000e+00 Inf [184] 1.000000e+00 7.200000e+02 3.628800e+06
 [70] Inf Inf Inf [187] 1.000000e+00 4.032000e+04 1.000000e+00
 [73] Inf Inf Inf [190] 1.650796e+136 1.000000e+00 6.227021e+09
 [76] Inf Inf 3.945524e+124 [193] 2.400000e+01 1.000000e+00 1.000000e+00
 [79] Inf Inf 3.628800e+05 [196] 1.000000e+00 1.405006e+51 1.000000e+00
 [82] Inf Inf Inf [199] 1.000000e+00 1.000000e+00 1.000000e+00
 [85] 1.974507e+182 Inf Inf [202] 1.000000e+00 7.200000e+02 1.000000e+00
 [88] Inf Inf 8.222839e+33 [205] 1.000000e+00 1.000000e+00 1.000000e+00
 [91] Inf Inf Inf [208] 1.000000e+00 1.000000e+00 1.000000e+00
 [94] 1.000000e+00 9.332622e+157 8.627210e+264 [211] 1.000000e+00 1.000000e+00 1.000000e+00
 [97] Inf Inf Inf [214] 1.000000e+00 1.000000e+00 1.000000e+00
 [100] Inf 1.033315e+40 1.000000e+00 [217] 1.000000e+00 1.000000e+00 1.000000e+00
 [103] Inf 5.109094e+19 1.000000e+00 [220] 1.000000e+00 1.000000e+00 1.000000e+00
 [106] 1.885495e+111 Inf 1.307674e+12 [223] 1.000000e+00 1.000000e+00 1.000000e+00
 [109] 1.032998e+148 1.000000e+00 Inf [226] 1.000000e+00 1.000000e+00 1.000000e+00
 [112] 2.092279e+13 4.470115e+105 2.480036e+96 [229] 1.000000e+00 1.000000e+00 1.000000e+00
 [115] 2.432902e+18 1.000000e+00 Inf [232] 1.000000e+00 1.000000e+00 1.000000e+00
 [118] 1.269640e+73 7.200000e+02 1.268869e+89 [235] 1.000000e+00 1.000000e+00 1.000000e+00
 [121] Inf 2.231193e+184 1.762953e+180 x<-factorial(test\$Deaths...newly.reported.in.last.24.hours)
 [124] 9.426890e+153 1.000000e+00 1.000000e+00 x
 [127] 9.875044e+202 2.006344e+269 2.400000e+01 [1] Inf Inf Inf
 [130] 1.000000e+00 8.159153e+47 1.000000e+00 [4] Inf Inf Inf
 [133] 2.000000e+00 1.000000e+00 1.000000e+00 [7] 4.274883e+69 Inf Inf
 [136] 8.222839e+33 6.227021e+09 Inf [10] 1.551121e+25 Inf Inf
 [139] 4.032000e+04 5.502622e+57 Inf [13] Inf Inf Inf
 [142] 1.000000e+00 2.556324e+258 Inf [16] Inf Inf Inf
 [145] 5.797126e+120 1.000000e+00 Inf [19] 8.627210e+264 3.146997e+85 2.004402e+291
 [148] 1.088887e+28 1.000000e+00 3.307885e+107 [22] 1.216451e+17 Inf 2.631308e+35
 [151] Inf 6.041526e+52 2.092279e+13 [25] Inf 3.991680e+07 2.952328e+38
 [154] 5.109094e+19 1.000000e+00 1.229694e+289 [28] 1.000000e+00 3.628800e+06 8.320987e+81
 [157] 2.000000e+00 1.000000e+00 1.000000e+00 [31] 3.628800e+05 Inf 9.426890e+153
 [160] 1.200000e+02 1.000000e+00 1.000000e+00 [34] 3.307885e+107 9.916779e+149 3.345253e+49
 [163] 1.000000e+00 2.422710e+130 1.000000e+00 [37] 1.000000e+00 3.719933e+41 7.200000e+02
 [166] 3.048883e+29 3.628800e+06 3.048883e+29 [40] 8.683318e+36 1.376375e+43 6.000000e+00
 [169] 4.032000e+04 1.000000e+00 1.000000e+00 [43] 3.628800e+05 1.650796e+136 3.659043e+232
 [172] 1.000000e+00 7.200000e+02 1.000000e+00 [46] 2.000000e+00 2.400000e+01 2.350561e+78
 [175] 1.216451e+17 1.000000e+00 1.551119e+66 [49] 3.628800e+06 1.000000e+00 3.048883e+29

[52] 3.041409e+64 4.470115e+105 3.628800e+05
 [55] 8.222839e+33 1.000000e+00 1.216451e+17
 [58] 6.082819e+62 3.628800e+05 2.400000e+01
 [61] 1.200000e+02 2.039788e+46 1.000000e+00
 [64] 3.146997e+85 6.227021e+09 2.432902e+18
 [67] 1.000000e+00 1.000000e+00 8.683318e+36
 [70] 1.124001e+21 1.088887e+28 3.556874e+14
 [73] 2.000000e+00 3.314240e+126 4.790016e+08
 [76] 4.790016e+08 1.088887e+28 1.000000e+00
 [79] 1.000000e+00 2.092279e+13 1.000000e+00
 [82] 1.033315e+40 7.200000e+02 8.159153e+47
 [85] 6.000000e+00 2.092279e+13 2.000000e+00
 [88] 7.200000e+02 3.556874e+14 1.000000e+00
 [91] 1.000000e+00 4.790016e+08 2.000000e+00
 [94] 1.000000e+00 1.000000e+00 1.000000e+00
 [97] 1.000000e+00 1.000000e+00 1.000000e+00
 [100] 2.400000e+01 6.000000e+00 6.000000e+00
 [103] 1.000000e+00 1.000000e+00 1.000000e+00
 [106] 1.200000e+02 2.000000e+00 3.628800e+05
 [109] 1.000000e+00 1.000000e+00 1.307674e+12
 [112] 1.000000e+00 6.000000e+00 1.000000e+00
 [115] 1.000000e+00 1.000000e+00 1.000000e+00
 [118] 1.000000e+00 1.000000e+00 1.000000e+00
 [121] 7.200000e+02 1.000000e+00 1.000000e+00
 [124] 2.000000e+00 1.000000e+00 1.000000e+00
 [127] 3.628800e+06 1.000000e+00 1.000000e+00
 [130] 1.000000e+00 1.000000e+00 1.000000e+00
 [133] 1.000000e+00 1.000000e+00 1.000000e+00
 [136] 2.000000e+00 1.000000e+00 1.000000e+00
 [139] 1.000000e+00 1.000000e+00 8.841762e+30
 [142] 1.000000e+00 2.000000e+00 2.000000e+00
 [145] 6.000000e+00 1.000000e+00 2.000000e+00
 [148] 1.000000e+00 1.000000e+00 1.000000e+00
 [151] 2.000000e+00 1.000000e+00 1.000000e+00
 [154] 1.000000e+00 1.000000e+00 1.000000e+00
 [157] 1.000000e+00 1.000000e+00 1.000000e+00
 [160] 1.000000e+00 1.000000e+00 1.000000e+00
 [163] 1.000000e+00 6.402374e+15 1.000000e+00
 [166] 1.000000e+00 1.000000e+00 1.000000e+00
 [169] 1.000000e+00 1.000000e+00 1.000000e+00
 [172] 1.000000e+00 1.000000e+00 1.000000e+00
 [175] 1.000000e+00 1.000000e+00 1.000000e+00
 [178] 1.000000e+00 1.000000e+00 1.000000e+00
 [181] 1.000000e+00 1.000000e+00 1.000000e+00
 [184] 1.000000e+00 1.000000e+00 1.000000e+00
 [187] 1.000000e+00 1.000000e+00 1.000000e+00
 [190] 1.000000e+00 1.000000e+00 1.000000e+00
 [193] 1.000000e+00 1.000000e+00 1.000000e+00
 [196] 1.000000e+00 1.000000e+00 1.000000e+00
 [199] 1.000000e+00 1.000000e+00 1.000000e+00
 [202] 1.000000e+00 1.000000e+00 1.000000e+00
 [205] 1.000000e+00 1.000000e+00 1.000000e+00
 [208] 1.000000e+00 1.000000e+00 1.000000e+00
 [211] 1.000000e+00 1.000000e+00 1.000000e+00
 [214] 1.000000e+00 1.000000e+00 1.000000e+00
 [217] 1.000000e+00 1.000000e+00 1.000000e+00
 [220] 1.000000e+00 1.000000e+00 1.000000e+00
 [223] 1.000000e+00 1.000000e+00 1.000000e+00
 [226] 1.000000e+00 1.000000e+00 1.000000e+00
 [229] 1.000000e+00 1.000000e+00 1.000000e+00
 [232] 1.000000e+00 1.000000e+00 1.000000e+00
 [235] 1.000000e+00 1.000000e+00 1.000000e+00

R graphical systems

There exist two illustrations frameworks in R, to be specific the traditional graphical offices and the network illustrations framework. The previous executes a write and paper show framework where person graphical things (either graphical primitives such as lines, focuses, etc. or whole plots) are plotted on beat or each other. This plotting framework is quick, with generally little flexibility. It offers the default, most coordinate and basic bolster for plotting. The network framework, on the other hand, offers graphical primitive objects that can be made and controlled autonomously of the plot. Most clients don't utilize grid directly to create total plots but or maybe depend on one of the sophisticated, high-level interfacing that effectively deliver complex figures, specifically gg fortify [14] and ggplot2 [15]. The ggplot2 bundle actualizes Wilkinson's Language structure of Graphics [16] to outline information to geometric objects (focuses, lines, etc.) and stylish qualities (color, shape, measure, etc.) and optional factual changes to create figures. Figure 1 display four cases plots produced with the different input data using the three aforementioned systems. cases plots ((A).cases cumulative total/Death cumulative total (B). cases newly reported in last 7 days/Death newly reported in last 7 days (C). cases newly reported in last 24 hours/Death newly

reported in last 24 hours (D). Countries region/mode of Transmission classification), it displaying the relation between the (average) log2 fold-change and the average expression intensity of features between groups. In that graphs contain the r factor at 7.5 at each on the (A) cases newly death report 2% of cumulative total of active and newly reported cases. The countries region mostly European, America and Asia region is sporadic cases of transmission causes. African countries causes in community transmission.

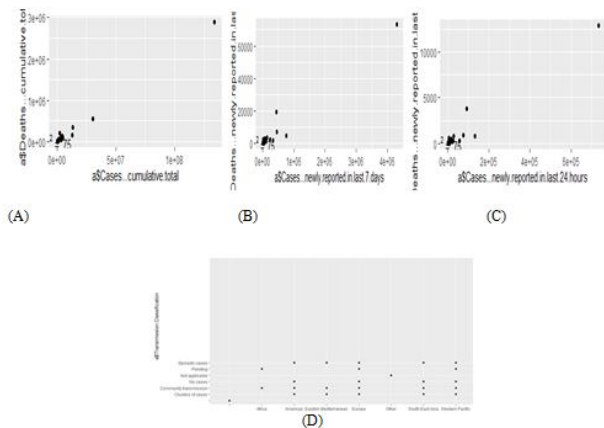


Figure1: Scatter Plot (A) Cases cumulative total/Death cumulative total; (B) Cases newly reported in last 7 days/Death newly reported in last 7 days; (C) Cases newly reported in last 24 hours/Death newly reported in last 24 hours; (D) Countries region/mode of Transmission classification

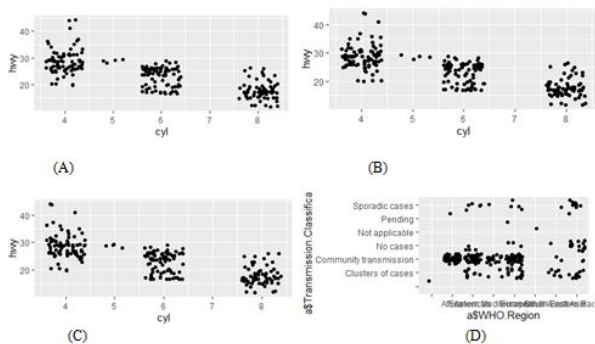


Figure2: Dotted Plot (A) Cases cumulative total/Death cumulative total; (B) Cases newly reported in last 7 days/Death newly reported in last 7 days; (C) Cases newly reported in last 24 hours/Death newly reported in last 24 hours; (D) Countries region/mode of Transmission classification

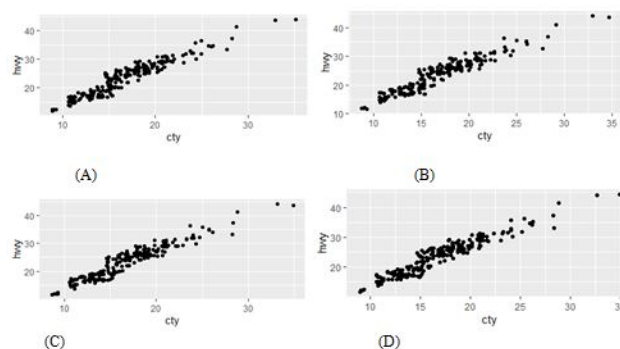


Figure3: Jitter Plot (A) Cases cumulative total/Death cumulative total; (B) Cases newly reported in last 7 days/Death newly reported in last 7 days; (C) Cases newly reported in last 24 hours/Death newly reported in last 24 hours; (D) Countries region/mode of Transmission classification

Plotting with R

This area presents a really brief presentation to R’s basic plotting usefulness. There exist various R tutorials that moreover present R’s plotting frameworks and syntax.

Good plotting reference fabric for the conventional plotting framework and network can be found in R Design by Paul Murrell [3]. Deepayan Sparker’s book Cross section. Multivariate Information Visualization with R [19] gives an in-depth and thorough reference to cross section (<http://lattice.r-forge.r-project.org/>). The ggplot2 book by Hadley Wickham [20] portrays theoretical principles of the bundle but is marginally obsolete in terms of syntax; the on-line documentation provides the latest reference. Finally, the Graphs chapter in the on-line Cookbook for R by Winston Chang or his R Graphics Cookbook are excellent practical guides for plotting with R.

Scatter plots, i.e. the representation of a set of (x, y) arrange in a 2D chart can effectively be created with the plot function. Usually the work that was utilized to form the top cleared out covid cases plots in figure 1. This work can moreover be used to deliver well of lava plots, by plotting the log2 fold-changes of proteins against $-\log_{10}(p)$ (where p is the p-value or, superior, the p-value after alteration for numerous comparisons) to highlight proteins with huge impact estimate (toward two sides of the chart) and little balanced p-values (toward the top corners of the chart). Group Comparison Plots (with type jitter and dabled Plot) for reported covid cases information (see cases within the companion vignette).

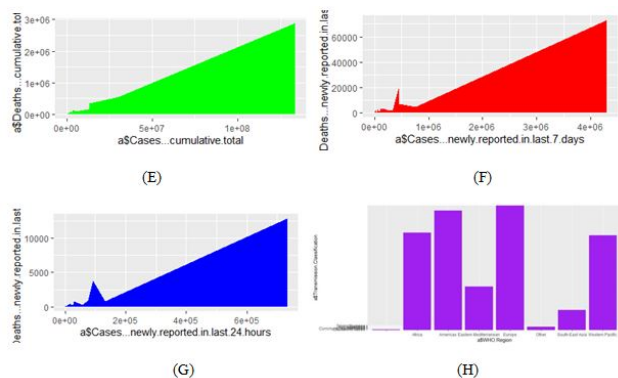


Figure4: Area Plot (E) Cases cumulative total/Death cumulative total; (F) Cases newly reported in last 7 days/Death newly reported in last 7 days; (G) Cases newly reported in last 24 hours/Death newly reported in last 24 hours; (H) Countries region/mode of Transmission classification

Graph: A In the first graph on x-axis $0e+00$ and e cases are in starting point .In between $0e+00$ and $2e+05$ the death r on y-axis $0e+00$ the cases are in starting stage .Graph:-B In the graph on x-axis $0e+00$ and y-axis 0 the case are in starting point In between $0e+00$ and $1e+06$ the ca Graphic In the graph $0e+00$ coming to next $5e+07$ on x-axis $1e+06$ the cumulative cases increase at the time death rate increase Coming to next on x-axis $1e+08$ on y-axis $3e+06$ the death rate and cumulative cases are in peak stage ses are slightly increased and death tare also reported Coming to next day the death rate and cases are also newly reported The highest death rate is reported in $4e+06$ on x-axis cases increase and no:of death rate also increase rate also increase.

Covid-19 cases data

The WHO corona virus (COVID-19) dashboard presents official daily counts of COVID-19 cases, deaths and vaccine utilization reported by countries, territories and areas. Through this dashboard, we aim to provide a frequently updated data dissemination and data exploration resource, while linking users to other useful and informative resources. Caution must be taken when interpreting all data presented, and differences between information products published by WHO, national public health authorities, and other sources using different inclusion criteria and different data cut-off times are to be expected. While steps are taken to ensure accuracy and reliability, all data are subject to continuous verification and change. All counts are subject to variations in case detection, definitions, laboratory testing, vaccination strategy, and reporting strategies. This data-set has daily level information on the number of affected cases, deaths and recovery from 2019 novel corona-virus. Please note that this is a time series data and so the number of cases on any given day is the cumulative number. Our complete COVID-19 data-set is a collection of the COVID-19 data maintained by Our World in Data. It is updated daily and includes data on confirmed cases, deaths, and testing.

The covid19.analytics is a R package allows users to obtain live worldwide data from the novel Corona-virus Disease originally

reported in 2019, CoViD-19. One of the main goals of this package is to make the latest data about the covid19 pandemic promptly available to researchers and the scientific community. The covid19 analytics is a package also provides basic analysis tools and functions to investigate these datasets. If you have data structured in a data. Frame organized as described above, then most of the functions provided by the covid19.analytics package for analyzing Time Series data will work with your data. In this way it is possible to add new data sets to the ones that can be loaded using the repositories predefined in this package and extend the analysis capabilities to these new datasets.

Be sure also to check the compatibility of these datasets using the Data Integrity and Consistency Checks functions described in the following section.

Statistical analysis for Quality control

Quality control is regularly a visualization-intensive errand, where specific measurements of intrigued are plotted for either numerous tests in an test, all highlights in a single try or several covid-19 cases runs are plotted to highlight efficient exceptions or dubious designs. These quality plots regularly apply essential plotting capacities, such as those briefly presented in Section 3, or their ggfortify and ggplot2 reciprocals. For example, the tidyverse function from the ggplot2 package implements the histogram of mass delta distributions described in Foster et al. [18] (Figure. 5). Information quality control is the method of controlling the utilization of information for an application or a prepare. This prepare is performed both sometime recently and after a Information Quality Affirmation (QA) handle, which comprises of revelation of information irregularity and adjustment. To survey quality, one must degree “the state of completeness, legitimacy, consistency, timeliness and precision that produces information fitting for a particular use”. Once surveyed, data improvement (cleansing) forms can be put in put to preserve tall information quality, as measured with continuous observing. The method distinguishes holes in information creating visual representations of information.

The data-set are doing the quality control to remove the gaps and comparison analyst in the data set. It removes from the duplicate identifiers. It also performs the statistical elevation. The evaluation to represent visualize the graphically.

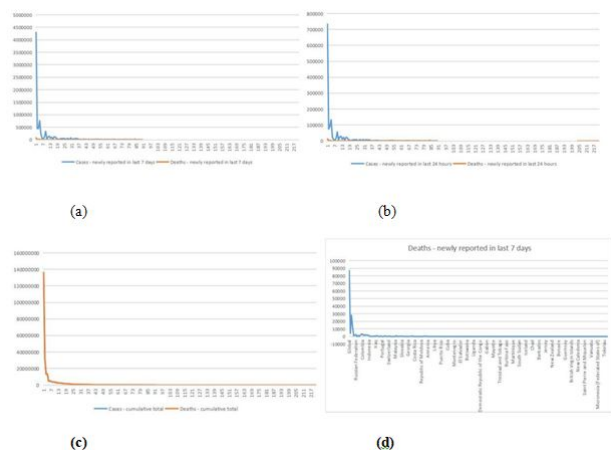


Figure5: Quality control: (a) Cases cumulative total/Death cumulative total; (b) Cases newly reported in last 7 days/Death newly reported in last 7 days; (c) Cases newly reported in last 24 hours/Death newly reported in last 24 hours; (d) Countries region/mode of Transmission classification

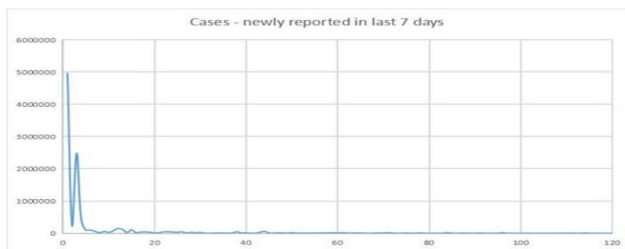


Figure6: Cases Reported.

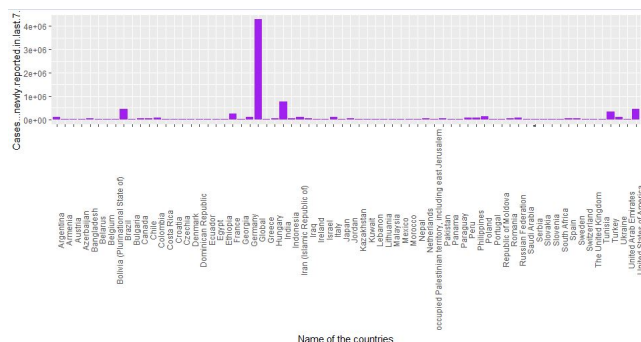


Figure7: Countries with COVID-19 Widespread.

CONCLUSION AND DISCUSSION

Cases newly reported in last 7 days per 100000.population comparison of 100000 of population 90% of the population causes. The Deaths newly reported in last 7 days per 100000 population comparison of 100000 populations 1.46 population causes death report. According the ICMR survey result match my data analysis report. Cases reported every 100000 of population in 2% and death report also 1.46% hence hence out of these result every 100000 reported covid19 at 3.146%.These analysis is used to know the death rate and positive rate in future.

REFERENCES

1. Anscombe, F. J., Graphs in statistical analysis. *Am. Stat.* 1973, 27, 17-21.
2. R Core Team 2014 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria 2014.
3. Murrell, P., R Graphics, Second Edition. Chapman & Hall/CRC Press, Boca Raton, FL, USA 2011.

4. Gatto, L., RforProteomics: Companion Package to the 'Using R and Bioconductor for Proteomics Data Analysis' Publication. R package version 1.5.5, 2015, Bioconductor, Seattle, USA.
5. Gentleman, R., Carey, V. J., Bates, D. M., Bolstad, B. et al., Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004, 5, R80.
6. Gatto, L., A current perspective on using R and Bioconductor for proteomics data analysis. *HUPO Conference 2014*, Madrid, Spain.
7. Gatto, L., Christoforou, A., Using R and Bioconductor for proteomics data analysis. *Biochim Biophys Acta.* 2014; 1844: 42-51.
8. Horn, T., Sandmann, T., Fischer, B., Axelsson, E. et al., Mapping of signaling networks through synthetic genetic interaction analysis by RNAi. *Nat Methods.* 2011; 8: 341-346.
9. Perez-Riverol, Y., Audain, E., Millan, A., Ramos, Y. et al., Isoelectric point optimization using peptide descriptors and support vector machines. *J Proteomics.* 2012; 75: 2269- 2274.
10. Kuhn, M., caret: Classification and Regression Training. R package version 6.0-41, 2015.
11. Karato, S., Riedel, M.R., Yuen, D.A., (2001). Rheological structure and deformation of subducted slabs in the mantle transition zone: implications for mantle circulation and deep earthquakes. *Phys Earth Planet In.* 127: 1-7.
12. Hirano, N., Ogawa, Y., Saito, K., Longlived early Cretaceous seamount volcanism in the Mariana Trench, Western Pacific Ocean. *Marine Geol.* 2003; 189: 371-379.
13. Curtis, A.C., Moyer, C.L., Mariana forearc serpentinite mud volcanoes harbor novel communities of extremophilic Archaea[J]. *Geomicrobiol J.* 2005; 30(5): 430-441.
14. shizuka, O., Hickey-Vargas, R., Arculus, R.J., Yogodzinski, G.M., Savov, I.P., Kusano, Y., McCarthy, A., Brandl, Ph., Sudo, M., Age of Izu-Bonin-Mariana arc basement. *Earth Planet Sci Lett.* 2018; 481: 80-90.
15. Pabst, S., Zack, Th., Savov, I.P., Ludwig, Th., Rost, D., Tonarini, S., Vicenzi, E.P., The fate of subducted oceanic slabs in the shallow mantle: Insights from boron isotopes and light element composition of metasomatized blueschists from the Mariana forearc. *Lithos.* 2012; 132: 162-179.
16. Yoshida, M., Trench dynamics: Effects of dynamically migrating trench on subducting slab morphology and characteristics of subduction zones systems. *Phys Earth Planet In.* 2017; 268: 35-53.
17. Faccenna, C., Di Giuseppe, E., Funicello, F., Lallemand S., van Hunen J., Control of seafloor aging on the migration of the Izu-Bonin-Mariana trench. *Earth Planet Sci Lett.* 2009; 288: 386-398.
18. Freymuth, H., Vils, F., Willbold, M., Taylor, R. N., Elliott, T., Molybdenum mobility and isotopic fractionation during subduction at the Mariana arc. *Earth Planet Sci Lett.* 2015; 432: 176-186