



NITTE
(Deemed to be University)

**NMAM INSTITUTE
OF TECHNOLOGY**

Nitte (DU) established under Section 3 of UGC Act 1956 | Accredited with 'A+' Grade by NAAC

Department of Computer Science & Engineering

Report on Mini Project

Covid – 19 Data Analysis

Course Code: CS1602-1

Course Name: Data Analysis using R Programmig

Semester: III

Section: C

Submitted To:

Dr. Anisha P Rodrigues
Assistant Professor [GD-I]
Department of Computer Science and
Engineering

Submitted By:

T Rohan Kini (NNM22CS187)

Date of submission:

20/11/2023

Signature of Course Instructor

ABSTRACT

A novel coronavirus (CoV) named '2019-nCoV' or '2019 novel coronavirus' or 'COVID-19' by the World Health Organization (WHO) is in charge of the current outbreak of pneumonia that began at the beginning of December 2019 near in Wuhan City, Hubei Province, China [1–4]. COVID-19 is a pathogenic virus. From the phylogenetic analysis carried out with obtainable full genome sequences, bats occur to be the COVID-19 virus reservoir, but the intermediate host(s) has not been detected till now.

Coronavirus disease (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus.

Most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, some will become seriously ill and require medical attention. Older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illness. Anyone can get sick with COVID-19 and become seriously ill or die at any age.

We use R Studio (or VS Code), a free software environment for statistical computing and graphics.

Let us explore and visualize public health (Confirmed, death and recovery cases) **Covid-19** data using **R**.

TABLE OF CONTENTS

Title Page	i
Abstract.....	ii
Table of Contents	iii
Introduction	4
Problem Statement.....	5
Objectives	6
Methodology.....	7
Implementation Details	8-16
Results	17-35
Conclusion and Future Scope.....	36-38
References	39

INTRODUCTION

A novel coronavirus (CoV) named '2019-nCoV' or '2019 novel coronavirus' or 'COVID-19' by the World Health Organization (WHO) is in charge of the current outbreak of pneumonia that began at the beginning of December 2019 near in Wuhan City, Hubei Province, China [[1](#)–[4](#)]. COVID-19 is a pathogenic virus.

The COVID-19 Data provides a daily summary of COVID-19 cases, deaths, recovered for 230+ countries, 760+ regions, and 12000+ administrative divisions of lower level.

The goal of COVID-19 Data also provide the research community with a unified dataset by collecting worldwide fine-grained case data, merged with exogenous variables helpful for a better understanding of COVID-19 Analysis.

In this report we are analyzing Death, Recovered and Confirmed cases on chosen countries i.e United States, Germany, Italy, South Korea. We also explore data based on date and population, using required packages.

We used Visual and descriptive analysis that help us to understand the data better. Hence we used graph to represent data using R programming.

PROBLEM STATEMENT

Using R programming implement the following operations

- a) Install and implement *ggplot2*, *ggthemes*, *tidyverse*, *cowplot*, *scales*, *glue*, *ggcorrplot* packages and attach necessary csv files of bulk raw data.
- b) Then converted all the empty columns to the value NA.
- c) Use *unique()* function to get a list of all the countries with different states so that we can check to see if the State column is null or not.
- d) Create new data frames for each type of cases and for each Country and merge those data frames into a more encompassing one.
- e) Reshape the data frames and also merge the confirmed, deaths, and recovered of each country into one data frame.
- f) Create a new data frames with all confirmed, deaths and recovered joined together and the dates as a single column that can be used for indexing.
- g) Plot a Graphs on Recovered, Deaths, and Confirmed against specific year on 4 chosen countries i.e USA, Germany, Italy and South Korea.



OBJECTIVES

Basically after implementing necessary library functions and attaching necessary csv files, our objectives are as follows:

- converted all the empty columns to the value NA
- We are getting a list of all the countries with different states so that we can check to see if the State column is NULL or not.
- We are creating new data frames for each type of cases and for each country, and to make it easier to work we are merging those data frames into a more encompassing one.
- Reshaping the data frames and also merging the confirmed, deaths, and recovered of each country into one data frame.
- Our main objective is to see which one of our chosen 4 countries have handled the data of virus in a way that can be generalized to everyone by plotting a graph. The targeted countries are:
 - ❖ United States
 - ❖ Germany
 - ❖ Italy
 - ❖ South Korea

METHODOLOGY

a) Package installing and loaded:

Necessary R packages ("**ggplot2**", "**ggthemes**", "**tidyverse**", "**cowplot**", "**scales**", "**glue**", "**ggcorrplot**") are installed and loaded using **library()**

To provide required functionality for data manipulation, exploration, transformation, visualization and Descriptive Analysis.

b) Data Collection:

We have got data source on Confirmed, Deaths, Recovered csv files from Kaggle site.

c) Data Cleaning and Preprocessing:

Perform data cleaning to handle missing values, Creating data, Removing empty rows and columns, Assigning NA or NULL to empty spaces, Remove Rows with Missing Values, Remove Duplicates, Modify Classes of Columns, Detect & Remove Outliers, Remove Spaces in Character Strings, Combine Categories.

d) Data Analysis:

We have generated Summery statistics, visualizations (such as plot_grid).

Observing a graph and applying statistical formule, we obtain "Number of cases which had an outcome", "percentage of cases that had an outcome", "Death rate", "Recovery rate", "Currently Active cases".

e) Visualization and Communication:

We used R data visualization packages to create clear and informative graphs, that effectively communicate the results of the analysis.

COVID-19 Data Analysis

- 1.) Load all Packages using library() function and link all necessary .csv files.

```
library(ggplot2)
library(ggthemes)
library(tidyverse)
library(cowplot)
library(scales)
library(glue)
library(ggcorrplot)

confirmed_file = "../R Project/time_series_covid19_confirmed_global.csv"
deaths_file = "../R Project/time_series_covid19_deaths_global.csv"
recovered_file = "../R Project/time_series_covid19_recovered_global.csv"
owid_file = "../R Project/owid-covid-data.csv"
```

- 2.) Read all csv files and converted all the empty columns to the value **NA** so it will be easier to work with later on.

```
confirmed_df = read.csv(confirmed_file, header = TRUE, na.strings = c("", " "))
deaths_df = read.csv(deaths_file, header = TRUE, na.strings = c("", " "))
recovered_df = read.csv(recovered_file, header = TRUE, na.strings = c("", " "))
owid_df = read.csv(owid_file, header = TRUE, na.strings = c("", " "), stringsAsFactors = F)

head(confirmed_df)
head(deaths_df)
head(recovered_df)
head(owid_df)
```

- 3.) From below code, Canada as a region appears multiple times, and that's due to the inclusion of states, there are other countries where the states are included as well to get a list of all the countries with different states we can check to see if the State column is null or not.

```
confirmed_df[confirmed_df$Country.Region == "Canada", ]
```

- 4.) Using unique() the countries with states specified all across the file ("**time_series_covid19_confirmed_global.csv**").

```
unique(confirmed_df[!is.na(confirmed_df$Province.State), "Country.Region"])
```

- 5.) Create data frames for each type of cases and for each country, and to make it easier to work with its better to merge those data frames into a more encompassing one.

Here we r analyzing confirmed, deaths and recovery cases beginning of Covid arrival of top 5 days from month January.

```
us_confirmed = confirmed_df[confirmed_df$Country.Region == "US", ]
```

```
us_deaths = deaths_df[deaths_df$Country.Region == "US", ]
```

```
us_recovered = recovered_df[recovered_df$Country.Region == "US", ]
```

```
germany_confirmed = confirmed_df[confirmed_df$Country.Region == "Germany", ]
```

```
germany_deaths = deaths_df[deaths_df$Country.Region == "Germany", ]
```

```
germany_recovered = recovered_df[recovered_df$Country.Region == "Germany", ]
```

```
italy_confirmed = confirmed_df[confirmed_df$Country.Region == "Italy", ]
```

```
italy_deaths = deaths_df[deaths_df$Country.Region == "Italy", ]
```

```
italy_recovered = recovered_df[recovered_df$Country.Region == "Italy", ]
```

```
sk_confirmed = confirmed_df[confirmed_df$Country.Region == "Korea, South", ]
```

```
sk_deaths = deaths_df[deaths_df$Country.Region == "Korea, South", ]
```

```
sk_recovered = recovered_df[recovered_df$Country.Region == "Korea, South", ]
```

```
us_confirmed
```

- 6.) Reshape the data frames and also merge the confirmed, deaths, and recovered of each country into one data frame.

```
clean_frame <- function(df){
```

```
# transposing the dataframe and gathering dates and number of cases
```

```
suppressWarnings(df <- df %>% rownames_to_column() %>% gather(Date, Cases, -c()))
```

```
# remove extra column
```

```
df$rowname <- NULL
```

```
suppressWarnings(df$Cases <- as.numeric(df$Cases))
```

```
#remove extra rows
```

```
df <- df[-c(1, 2, 3, 4, 5), ]
```

COVID-19 Data Analysis

```
df$Date <- sapply(df$Date,function(x) {x <- gsub("X","",x)})

# convert date to actual date datatype
df$Date <- as.Date(df$Date, "%m.%d.%y")

return (df)
}

# combining confirmed, deaths and recoveries data frames
combine_frames = function(conf, death, rec) {
  combined_df = data.frame(matrix(ncol = 0, nrow = nrow(conf)))
  combined_df$Date = conf$Date
  combined_df$Confirmed = conf$Cases
  combined_df$Deaths = death$Cases
  combined_df$Recovered = rec$Cases

  return (combined_df)
}

us_df = combine_frames(clean_frame(us_confirmed), clean_frame(us_deaths),
  clean_frame(us_recovered))

germany_df = combine_frames(clean_frame(germany_confirmed),
  clean_frame(germany_deaths), clean_frame(germany_recovered))

italy_df = combine_frames(clean_frame(italy_confirmed), clean_frame(italy_deaths),
  clean_frame(italy_recovered))

sk_df = combine_frames(clean_frame(sk_confirmed), clean_frame(sk_deaths),
  clean_frame(sk_recovered))

head(us_df)
head(germany_df)
head(italy_df)
head(sk_df)
```

- 7.) Data visualization and descriptive analysis for each country. By plotting graph on canvas. Observing a graph and applying statistical formule, we obtain "Number of cases which had an outcome", "percentage of cases that had an outcome", "Deaths rate", "Recovery rate", "Currently Active cases" of chosen 4 countries (**"United States"**, **"Germany"**, **"Italy"**, **"South Korea"**).

```
fig <- function(width, heigth){
  options(repr.plot.width = width, repr.plot.height = heigth)
}
theme_set(theme_cowplot())
#Analysis on cases in United states
fig(14, 13)
```

#plotting graph

```
plot_grid(ggplot(data = us_df , aes(x = Date , y = Confirmed)) +
  scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y"))
+geom_line(colour = "yellow", size = 2) + scale_y_continuous(labels =
unit_format(unit = "M", scale = 1e-6)),
ggplot(data = us_df , aes(x = Date , y = Deaths)) + scale_x_date(date_breaks = "1
month", labels = date_format("%b %Y")) +geom_line(colour = "red", size = 2) +
scale_y_continuous(labels = unit_format(unit = "K", scale = 1)),
ggplot(data = us_df , aes(x = Date , y = Recovered)) + scale_x_date(date_breaks = "1
month", labels = date_format("%b %Y")) +geom_line(colour = "green", size = 2) +
scale_y_continuous(labels = unit_format(unit = "M", scale = 1e-6)),
labels = "", align = "v", ncol = 1)
```

```
# we obtain "Number of cases which had an outcome", "percentage of cases that had
an outcome", "Deaths rate", "Recovery rate", "Currently Active cases" of United States
us_cases_outcome = tail(us_df, 1)[3] + tail(us_df, 1)[4] # cases that had an outcome
us_cases_outcome = as.numeric(us_cases_outcome$Deaths)
us_outcome_percent = round((us_cases_outcome / tail(us_df, 1)[2]) * 100,
2)$Confirmed
us_deaths_percent = round((tail(us_df, 1)[3] / us_cases_outcome)$Deaths * 100, 2)
us_rec_percent = round((tail(us_df, 1)[4] / us_cases_outcome)$Recovered * 100, 2)
us_active = round(tail(us_df, 1)[2] - us_cases_outcome, 2)
```

```
glue("Number of cases which had an outcome: {us_cases_outcome}")
glue("percentage of cases that had an outcome: {us_outcome_percent}%")
glue("Deaths rate: {us_deaths_percent}%")
glue("Recovery rate: {us_rec_percent}%")
glue("Currently Active cases: {us_active}")
```

#Analysis on cases in Germany

fig(14, 13)

#plotting graph

```
plot_grid(ggplot(data = germany_df , aes(x = Date , y = Confirmed)) +
scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y")) +
geom_line(colour = "yellow", size = 2) + scale_y_continuous(labels = unit_format(unit =
"K", scale = 1)),
ggplot(data = germany_df , aes(x = Date , y = Deaths)) + scale_x_date(date_breaks =
"1 month", labels = date_format("%b %Y")) +geom_line(colour = "red", size = 2) +
scale_y_continuous(labels = unit_format(unit = "K", scale = 1)),
ggplot(data = germany_df , aes(x = Date , y = Recovered)) + scale_x_date(date_breaks
= "1 month", labels = date_format("%b %Y")) +
geom_line(colour = "green", size = 2) + scale_y_continuous(labels =unit_format(unit =
"K", scale = 1)),labels = "", align = "v", ncol = 1)
```

we obtain "Number of cases which had an outcome", "percentage of cases that had an outcome", "Death rate", "Recovery rate", "Currently Active cases" of Germany

```
germany_cases_outcome = tail(germany_df, 1)[3] + tail(germany_df, 1)[4] # cases that
had an outcome
germany_cases_outcome = as.numeric(germany_cases_outcome$Deaths)
germany_outcome_percent = round((germany_cases_outcome / tail(germany_df, 1)[2])
* 100, 2)$Confirmed
germany_deaths_percent = round((tail(germany_df, 1)[3] /
germany_cases_outcome)$Deaths * 100, 2)
germany_rec_percent = round((tail(germany_df, 1)[4] /
germany_cases_outcome)$Recovered * 100, 2)
germany_active = round(tail(germany_df, 1)[2] - germany_cases_outcome, 2)

glue("Number of cases which had an outcome: {germany_cases_outcome}")
glue("percentage of cases that had an outcome: {germany_outcome_percent}%")
glue("Death rate: {germany_deaths_percent}%")
glue("Recovery rate: {germany_rec_percent}%")
glue("Currently Active cases: {germany_active}")
```

#Analysis on cases in Italy

fig(14, 13)

#plotting graph

```
plot_grid(ggplot(data = italy_df , aes(x = Date , y = Confirmed)) +
scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y"))
+geom_line(colour = "yellow", size = 2) + scale_y_continuous(labels =
unit_format(unit = "K", scale = 1)),
ggplot(data = italy_df , aes(x = Date , y = Deaths)) + scale_x_date(date_breaks = "1
month", labels = date_format("%b %Y")) +
geom_line(colour = "red", size = 2) + scale_y_continuous(labels = unit_format(unit =
"K", scale = 1)),
ggplot(data = italy_df , aes(x = Date , y = Recovered)) + scale_x_date(date_breaks = "1
month", labels = date_format("%b %Y")) +
geom_line(colour = "green", size = 2) + scale_y_continuous(labels = unit_format(unit =
"K", scale = 1)), labels = "", align = "v", ncol = 1)
```

we obtain "Number of cases which had an outcome", "percentage of cases that had an outcome", "Death rate", "Recovery rate", "Currently Active cases" of Italy

```
italy_cases_outcome = tail(italy_df, 1)[3] + tail(italy_df, 1)[4] outcome
italy_cases_outcome = as.numeric(italy_cases_outcome$Deaths)
italy_outcome_percent = round((italy_cases_outcome / tail(italy_df, 1)[2]) * 100,
2)$Confirmed
italy_deaths_percent = round((tail(italy_df, 1)[3] / italy_cases_outcome)$Deaths * 100,
2)
italy_rec_percent = round((tail(italy_df, 1)[4] / italy_cases_outcome)$Recovered * 100,
2)
italy_active = round(tail(italy_df, 1)[2] - italy_cases_outcome, 2)
```

```
glue("Number of cases which had an outcome: {italy_cases_outcome}")
glue("percentage of cases that had an outcome: {italy_outcome_percent}%")
glue("Death rate: {italy_deaths_percent}%")
glue("Recovery rate: {italy_rec_percent}%")
glue("Currently Active cases: {italy_active}")
```

#Analysis on cases in South Korea

```
fig(14, 13)
```

#plotting graph

```
plot_grid(ggplot(data = sk_df , aes(x = Date , y = Confirmed)) +
scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y"))
+geom_line(colour = "yellow", size = 2), ggplot(data = sk_df , aes(x = Date , y =
Deaths)) + scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y")) +
geom_line(colour = "red", size = 2), ggplot(data = sk_df , aes(x = Date , y =
Recovered)) + scale_x_date(date_breaks = "1 month", labels = date_format("%b %Y"))
+geom_line(colour = "green", size = 2), labels = "", align = "v", ncol = 1)
```

COVID-19 Data Analysis

we obtain "Number of cases which had an outcome", "percentage of cases that had an outcome", "Death rate", "Recovery rate", "Currently Active cases" of South Korea

```
sk_cases_outcome = tail(sk_df, 1)[3] + tail(sk_df, 1)[4] # cases that had an outcome
sk_cases_outcome = as.numeric(sk_cases_outcome$Deaths)
sk_outcome_percent = round((sk_cases_outcome / tail(sk_df, 1)[2]) * 100,
2)$Confirmed
sk_deaths_percent = round((tail(sk_df, 1)[3] / sk_cases_outcome)$Deaths * 100, 2)
sk_rec_percent = round((tail(sk_df, 1)[4] / sk_cases_outcome)$Recovered * 100, 2)
sk_active = round(tail(sk_df, 1)[2] - sk_cases_outcome, 2)

glue("Number of cases which had an outcome: {sk_cases_outcome}")
glue("percentage of cases that had an outcome: {sk_outcome_percent}%")
glue("Death rate: {sk_deaths_percent}%")
glue("Recovery rate: {sk_rec_percent}%")
glue("Currently Active cases: {sk_active}")
```


RESULTS AND DISCUSSIONS

Result 1:

We converted all the empty columns to the value NA so it will be easier to work with later on. After reading dataset observe first few countries for confirmed case, deaths, and recoveries.

#Note: output shown is taken from Rstudio out of n rows 3 rows are displayed with n columns and x1.22.20 & so on be date as mentioned in csv file.

#Confirmed case from time_series_covid19_confirmed_global.csv

Province.State	Country	Region	Lat	Long	x1.22.20	x1.23.20	x1.24.20		
1	<NA>	Afghanistan	33.93911	67.70995	0	0	0		
2	<NA>	Albania	41.15330	20.16830	0	0	0		
3	<NA>	Algeria	28.03390	1.65960	0	0	0		
	x1.25.20	x1.26.20	x1.27.20	x1.28.20	x1.29.20	x1.30.20	x1.31.20	x2.1.20	x2.2.20
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
	x2.3.20	x2.4.20	x2.5.20	x2.6.20	x2.7.20	x2.8.20	x2.9.20	x2.10.20	x2.11.20
x2.12.20									
1	0	0	0	0	0	0	0	0	0
0									
2	0	0	0	0	0	0	0	0	0
0									
3	0	0	0	0	0	0	0	0	0
0									
	x2.13.20	x2.14.20	x2.15.20	x2.16.20	x2.17.20	x2.18.20	x2.19.20	x2.20.20	x2.21.20
1	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
	x2.22.20	x2.23.20	x2.24.20	x2.25.20	x2.26.20	x2.27.20	x2.28.20	x2.29.20	x3.1.20
1	0	0	1	1	1	1	1	1	1
2	0	0	0	0	0	0	0	0	0
3	0	0	0	1	1	1	1	1	1
	x3.2.20	x3.3.20	x3.4.20	x3.5.20	x3.6.20	x3.7.20	x3.8.20	x3.9.20	x3.10.20
1	1	2	4	4	4	4	5	7	8
2	0	0	0	0	0	0	0	2	10
3	3	5	12	12	17	17	19	20	20
	x3.12.20	x3.13.20	x3.14.20	x3.15.20	x3.16.20	x3.17.20	x3.18.20	x3.19.20	x3.20.20
1	12	13	15	16	18	20	24	25	29
2	23	33	38	42	51	55	59	64	70
3	24	26	37	48	54	60	74	87	90
	x3.21.20	x3.22.20	x3.23.20	x3.24.20	x3.25.20	x3.26.20	x3.27.20	x3.28.20	x3.29.20
1	30	34	41	43	76	80	91	107	118
2	76	89	104	123	146	174	186	197	212
3	139	201	230	264	302	367	409	454	511
	x3.30.20	x3.31.20	x4.1.20	x4.2.20	x4.3.20	x4.4.20	x4.5.20	x4.6.20	x4.7.20
1	146	175	197	240	275	300	338	368	424
2	223	243	259	277	304	333	361	377	383
3	584	716	847	986	1171	1251	1320	1423	1468
	x4.9.20	x4.10.20	x4.11.20	x4.12.20	x4.13.20	x4.14.20	x4.15.20	x4.16.20	x4.17.20
1	485	532	556	608	666	715	785	841	907
2	409	416	433	446	467	475	494	518	539

COVID-19 Data Analysis

3	1666	1761	1825	1914	1983	2070	2160	2268	2418
1	x4.18.20	x4.19.20	x4.20.20	x4.21.20	x4.22.20	x4.23.20	x4.24.20	x4.25.20	x4.26.20
2	934	997	1027	1093	1177	1236	1331	1464	1532
3	548	562	584	609	634	663	678	712	726
3	2534	2629	2718	2811	2910	3007	3127	3256	3382
1	x4.27.20	x4.28.20	x4.29.20	x4.30.20	x5.1.20	x5.2.20	x5.3.20	x5.4.20	x5.5.20
2	1704	1830	1940	2127	2291	2470	2705	2895	3225
3	736	750	766	773	782	789	795	803	820
3	3517	3649	3848	4006	4154	4295	4474	4648	4838
1	x5.6.20	x5.7.20	x5.8.20	x5.9.20	x5.10.20	x5.11.20	x5.12.20	x5.13.20	x5.14.20
2	3393	3564	3781	4042	4403	4687	4968	5227	5640
3	832	842	850	856	868	872	876	880	898
3	4997	5182	5369	5558	5723	5891	6067	6253	6442
1	x5.15.20	x5.16.20	x5.17.20	x5.18.20	x5.19.20	x5.20.20	x5.21.20	x5.22.20	x5.23.20
2	6054	6403	6665	7073	7654	8146	8677	9219	10001
3	916	933	946	948	949	964	969	981	989
3	6629	6821	7019	7201	7377	7542	7728	7918	8113
1	x5.24.20	x5.25.20	x5.26.20	x5.27.20	x5.28.20	x5.29.20	x5.30.20	x5.31.20	x6.1.20
2	10585	11176	11834	12459	13039	13662	14528	15208	15753
3	998	1004	1029	1050	1076	1099	1122	1137	1143
3	8306	8503	8697	8857	8997	9134	9267	9394	9513
1	x6.2.20	x6.3.20	x6.4.20	x6.5.20	x6.6.20	x6.7.20	x6.8.20	x6.9.20	x6.10.20
2	16512	17270	18057	18972	19554	20345	20920	21462	22146
3	1164	1184	1197	1212	1232	1246	1263	1299	1341
3	9626	9733	9831	9935	10050	10154	10265	10382	10484
1	x6.12.20	x6.13.20	x6.14.20	x6.15.20	x6.16.20	x6.17.20	x6.18.20	x6.19.20	x6.20.20
2	23550	24106	24770	25531	26314	26878	27536	27882	28428
3	1416	1464	1521	1590	1672	1722	1788	1838	1891
3	10698	10810	10919	11031	11147	11268	11385	11504	11631
1	x6.21.20	x6.22.20	x6.23.20	x6.24.20	x6.25.20	x6.26.20	x6.27.20	x6.28.20	x6.29.20
2	28837	29147	29471	29705	30165	30441	30606	30957	31228
3	1962	1995	2047	2114	2192	2269	2330	2402	2466
3	11771	11920	12076	12248	12445	12685	12968	13273	13571
1	x6.30.20	x7.1.20	x7.2.20	x7.3.20	x7.4.20	x7.5.20	x7.6.20	x7.7.20	x7.8.20
2	31507	31826	32012	32314	32662	32941	33180	33374	33584
3	2535	2580	2662	2752	2819	2893	2964	3038	3106
3	13907	14272	14657	15070	15500	15941	16404	16879	17348
1	x7.10.20	x7.11.20	x7.12.20	x7.13.20	x7.14.20	x7.15.20	x7.16.20	x7.17.20	x7.18.20
2	34184	34356	34441	34595	34730	34984	35060	35219	35279
3	3278	3371	3454	3571	3667	3752	3851	3906	4008
3	18242	18712	19195	19689	20216	20770	21355	21948	22549
1	x7.19.20	x7.20.20	x7.21.20	x7.22.20	x7.23.20	x7.24.20	x7.25.20	x7.26.20	x7.27.20
2	35453	35493	35605	35717	35918	35978	36026	36147	36253
3	4090	4171	4290	4358	4466	4570	4637	4763	4880
3	23084	23691	24278	24872	25484	26159	26764	27357	27973
1	x7.28.20	x7.29.20	x7.30.20	x7.31.20	x8.1.20	x8.2.20	x8.3.20	x8.4.20	x8.5.20
2	36358	36463	36532	36665	36700	36701	36737	36773	36820
3	4997	5105	5197	5276	5396	5519	5620	5750	5889
3	28615	29229	29831	30394	30950	31465	31972	32504	33055
1	x8.6.20	x8.7.20	x8.8.20	x8.9.20	x8.10.20	x8.11.20	x8.12.20	x8.13.20	x8.14.20
2	36928	37006	37046	37083	37153	37260	37336	37422	37497
3	6016	6151	6275	6411	6536	6676	6817	6971	7117
3	33626	34155	34693	35160	35712	36204	36699	37187	37664
1	x8.15.20	x8.16.20	x8.17.20	x8.18.20	x8.19.20	x8.20.20	x8.21.20	x8.22.20	x8.23.20
2	37542	37590	37667	37710	37750	37852	37885	37944	37990
3	7260	7380	7499	7654	7812	7967	8119	8275	8427
3	38133	38583	39025	39444	39847	40258	40667	41068	41460
1	x8.24.20	x8.25.20	x8.26.20	x8.27.20	x8.28.20	x8.29.20	x8.30.20	x8.31.20	x9.1.20
2	38045	38061	38103	38119	38130	38133	38155	38159	38193
3	8605	8759	8927	9083	9195	9279	9380	9513	9606
3	41858	42228	42619	43016	43403	43781	44146	44494	44833
1	x9.2.20	x9.3.20	x9.4.20	x9.5.20	x9.6.20	x9.7.20	x9.8.20	x9.9.20	x9.10.20
2	38243	38288	38304	38324	38398	38494	38520	38544	38572
3	9728	9844	9967	10102	10255	10406	10553	10704	10860
3	45158	45469	45773	46071	46364	46653	46938	47216	47488
									47752

COVID-19 Data Analysis

```

x9.12.20 x9.13.20 x9.14.20 x9.15.20 x9.16.20 x9.17.20 x9.18.20 x9.19.20 x9.20.20
1 38641 38716 38772 38815 38855 38872 38897 38919 39044
2 11185 11353 11520 11672 11816 11948 12073 12226 12385
3 48007 48254 48496 48734 48966 49194 49413 49623 49826
x9.21.20 x9.22.20 x9.23.20 x9.24.20 x9.25.20 x9.26.20 x9.27.20 x9.28.20 x9.29.20
1 39074 39096 39145 39170 39186 39192 39227 39239 39254
2 12535 12666 12787 12921 13045 13153 13259 13391 13518
3 50023 50214 50400 50579 50754 50914 51067 51213 51368
x9.30.20 x10.1.20 x10.2.20 x10.3.20 x10.4.20 x10.5.20 x10.6.20 x10.7.20 x10.8.20
1 39268 39285 39290 39297 39341 39422 39486 39548 39616
2 13649 13806 13965 14117 14266 14410 14568 14730 14899
3 51530 51690 51847 51995 52136 52270 52399 52520 52658
x10.9.20 x10.10.20 x10.11.20 x10.12.20 x10.13.20 x10.14.20 x10.15.20 x10.16.20
1 39693 39703 39799 39870 39928 39994 40026 40088
2 15066 15231 15399 15570 15752 15955 16212 16501
3 52804 52940 53072 53325 53399 53584 53777 53998
x10.17.20 x10.18.20 x10.19.20 x10.20.20 x10.21.20 x10.22.20 x10.23.20 x10.24.20
1 40141 40200 40287 40369 40510 40626 40687 40768
2 16774 17055 17350 17651 17948 18250 18556 18858
3 54203 54402 54616 54829 55081 55357 55630 55880
x10.25.20 x10.26.20 x10.27.20 x10.28.20 x10.29.20 x10.30.20 x10.31.20 x11.1.20
1 40833 40937 41032 41145 41268 41334 41425 41501
2 19157 19445 19729 20040 20315 20634 20875 21202
3 56143 56419 56706 57026 57332 57651 57942 58272
x11.2.20 x11.3.20 x11.4.20 x11.5.20 x11.6.20 x11.7.20 x11.8.20 x11.9.20 x11.10.20
1 41633 41728 41814 41935 41975 42033 42159 42297 42463
2 21523 21904 22300 22721 23210 23705 24206 24731 25294
3 58574 58979 59527 60169 60800 61381 62051 62693 63446
x11.11.20 x11.12.20 x11.13.20 x11.14.20 x11.15.20 x11.16.20 x11.17.20 x11.18.20
1 42609 42795 42969 43035 43240 43468 43681 43924
2 25801 26211 26701 27233 27830 28432 29126 29837
3 64257 65108 65975 66819 67679 68589 69591 70629
x11.19.20 x11.20.20 x11.21.20 x11.22.20 x11.23.20 x11.24.20 x11.25.20 x11.26.20
1 44177 44363 44503 44706 44988 45174 45384 45600
2 30623 31459 32196 32761 33556 34300 34944 35600
3 71652 72755 73774 74862 75867 77000 78025 79110
x11.27.20 x11.28.20 x11.29.20 x11.30.20 x12.1.20 x12.2.20 x12.3.20 x12.4.20
1 45723 45844 46116 46274 46516 46718 46837 46837
2 36245 36790 37625 38182 39014 39719 40501 41302
3 80168 81212 82221 83199 84152 85084 85927 86730
x12.5.20 x12.6.20 x12.7.20 x12.8.20 x12.9.20 x12.10.20 x12.11.20
1 47072 47306 47516 47716 47851 48053 48116
2 42148 42988 43683 44436 45188 46061 46863
3 87502 88252 88825 89416 90014 90579 91121
[ reached 'max' / getOption("max.print") -- omitted 3 rows ]

```

#Deaths case from time_series_covid19_deaths_global.csv

```

Province.State Country.Region Lat Long x1.22.20 x1.23.20 x1.24.20
1 <NA> Afghanistan 33.93911 67.70995 0 0 0
2 <NA> Albania 41.15330 20.16830 0 0 0
3 <NA> Algeria 28.03390 1.65960 0 0 0
x1.25.20 x1.26.20 x1.27.20 x1.28.20 x1.29.20 x1.30.20 x1.31.20 x2.1.20 x2.2.20
1 0 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0 0 0
x2.3.20 x2.4.20 x2.5.20 x2.6.20 x2.7.20 x2.8.20 x2.9.20 x2.10.20 x2.11.20
x2.12.20
1 0 0 0 0 0 0 0 0
0
2 0 0 0 0 0 0 0 0
0

```

COVID-19 Data Analysis

3	0	0	0	0	0	0	0	0	0	0
0	x2.13.20	x2.14.20	x2.15.20	x2.16.20	x2.17.20	x2.18.20	x2.19.20	x2.20.20	x2.21.20	
1	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	0	
1	x2.22.20	x2.23.20	x2.24.20	x2.25.20	x2.26.20	x2.27.20	x2.28.20	x2.29.20	x3.1.20	
2	0	0	0	0	0	0	0	0	0	
3	0	0	0	0	0	0	0	0	0	
1	x3.2.20	x3.3.20	x3.4.20	x3.5.20	x3.6.20	x3.7.20	x3.8.20	x3.9.20	x3.10.20	x3.11.20
2	0	0	0	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0	0
1	x3.12.20	x3.13.20	x3.14.20	x3.15.20	x3.16.20	x3.17.20	x3.18.20	x3.19.20	x3.20.20	
2	0	0	0	0	0	0	0	0	0	
3	1	1	1	1	1	1	2	2	2	
1	x3.21.20	x3.22.20	x3.23.20	x3.24.20	x3.25.20	x3.26.20	x3.27.20	x3.28.20	x3.29.20	
2	0	1	1	1	2	3	3	4	4	
3	15	17	17	19	21	25	26	29	31	
1	x3.30.20	x3.31.20	x4.1.20	x4.2.20	x4.3.20	x4.4.20	x4.5.20	x4.6.20	x4.7.20	x4.8.20
2	4	4	4	4	8	10	12	15	16	16
3	11	15	15	16	17	20	20	21	22	22
1	x4.9.20	x4.10.20	x4.11.20	x4.12.20	x4.13.20	x4.14.20	x4.15.20	x4.16.20	x4.17.20	
2	17	18	18	19	22	23	29	30	30	
3	235	256	275	293	313	326	336	348	364	
1	x4.18.20	x4.19.20	x4.20.20	x4.21.20	x4.22.20	x4.23.20	x4.24.20	x4.25.20	x4.26.20	
2	30	33	36	36	40	40	43	47	52	
3	26	26	26	26	27	27	27	27	28	
1	x4.27.20	x4.28.20	x4.29.20	x4.30.20	x5.1.20	x5.2.20	x5.3.20	x5.4.20	x5.5.20	
2	58	59	61	64	68	72	85	90	95	
3	28	30	30	31	31	31	31	31	31	
1	x5.6.20	x5.7.20	x5.8.20	x5.9.20	x5.10.20	x5.11.20	x5.12.20	x5.13.20	x5.14.20	
2	432	437	444	450	453	459	463	465	470	
3	104	106	109	115	120	122	127	132	136	
1	x5.15.20	x5.16.20	x5.17.20	x5.18.20	x5.19.20	x5.20.20	x5.21.20	x5.22.20	x5.23.20	
2	31	31	31	31	31	31	31	31	31	
3	476	483	488	494	502	507	515	522	529	
1	x5.24.20	x5.25.20	x5.26.20	x5.27.20	x5.28.20	x5.29.20	x5.30.20	x5.31.20	x6.1.20	
2	153	168	170	173	178	188	194	206	217	
3	31	31	31	31	31	31	31	31	31	
1	x6.2.20	x6.3.20	x6.4.20	x6.5.20	x6.6.20	x6.7.20	x6.8.20	x6.9.20	x6.10.20	x6.11.20
2	536	542	548	555	561	568	575	582	592	
3	219	220	221	228	236	247	250	258	266	
1	x6.12.20	x6.13.20	x6.14.20	x6.15.20	x6.16.20	x6.17.20	x6.18.20	x6.19.20	x6.20.20	
2	32	32	33	33	33	33	33	33	33	
3	600	609	617	623	630	638	646	653	661	
1	x6.21.20	x6.22.20	x6.23.20	x6.24.20	x6.25.20	x6.26.20	x6.27.20	x6.28.20	x6.29.20	
2	274	298	304	313	331	361	373	388	409	430
3	33	33	33	33	34	34	34	34	34	35
1	x6.30.20	x7.1.20	x7.2.20	x7.3.20	x7.4.20	x7.5.20	x7.6.20	x7.7.20	x7.8.20	x7.9.20
2	667	673	681	690	698	707	715	724	732	741
3	450	455	475	482	495	508	550	552	573	
1	x6.31.20	x7.1.20	x7.2.20	x7.3.20	x7.4.20	x7.5.20	x7.6.20	x7.7.20	x7.8.20	x7.9.20
2	36	36	36	36	37	38	39	42	43	
3	751	760	767	777	788	799	811	825	837	
1	x7.10.20	x7.11.20	x7.12.20	x7.13.20	x7.14.20	x7.15.20	x7.16.20	x7.17.20	x7.18.20	x7.19.20
2	585	597	617	638	674	682	716	728	740	
3	44	44	45	47	49	51	53	55	58	
1	x7.20.20	x7.21.20	x7.22.20	x7.23.20	x7.24.20	x7.25.20	x7.26.20	x7.27.20	x7.28.20	x7.29.20
2	845	852	861	869	878	885	892	897	905	
3	752	780	809	821	828	866	900	922	939	959
1	62	65	69	72	74	76	79	81	83	83

COVID-19 Data Analysis

3	912	920	928	937	946	952	959	968	978	988
1	x7.10.20	x7.11.20	x7.12.20	x7.13.20	x7.14.20	x7.15.20	x7.16.20	x7.17.20	x7.18.20	
2	973	996	1012	1040	1064	1096	1117	1149	1166	
3	85	89	93	95	97	101	104	107	111	
3	996	1004	1011	1018	1028	1040	1052	1057	1068	
1	x7.19.20	x7.20.20	x7.21.20	x7.22.20	x7.23.20	x7.24.20	x7.25.20	x7.26.20	x7.27.20	
2	1183	1185	1188	1192	1213	1227	1249	1261	1271	
3	112	113	117	120	123	128	134	138	144	
3	1078	1087	1100	1111	1124	1136	1146	1155	1163	
1	x7.28.20	x7.29.20	x7.30.20	x7.31.20	x8.1.20	x8.2.20	x8.3.20	x8.4.20	x8.5.20	
2	1272	1273	1283	1284	1285	1286	1290	1290	1296	
3	148	150	154	157	161	166	172	176	182	
3	1174	1186	1200	1210	1223	1231	1239	1248	1261	
1	x8.6.20	x8.7.20	x8.8.20	x8.9.20	x8.10.20	x8.11.20	x8.12.20	x8.13.20	x8.14.20	
2	1301	1310	1315	1321	1329	1346	1356	1365	1370	
3	188	189	193	199	200	205	208	213	219	
3	1273	1282	1293	1302	1312	1322	1333	1341	1351	
1	x8.15.20	x8.16.20	x8.17.20	x8.18.20	x8.19.20	x8.20.20	x8.21.20	x8.22.20	x8.23.20	
2	1371	1376	1378	1382	1384	1386	1386	1386	1388	
3	225	228	230	232	234	238	240	245	250	
3	1360	1370	1379	1391	1402	1411	1418	1424	1435	
1	x8.24.20	x8.25.20	x8.26.20	x8.27.20	x8.28.20	x8.29.20	x8.30.20	x8.31.20	x9.1.20	
2	1390	1398	1402	1402	1403	1403	1403	1403	1407	
3	254	259	263	266	271	275	280	284	290	
3	1446	1456	1465	1475	1483	1491	1501	1510	1518	
1	x9.2.20	x9.3.20	x9.4.20	x9.5.20	x9.6.20	x9.7.20	x9.8.20	x9.9.20	x9.10.20	x9.11.20
2	1410	1410	1410	1410	1413	1416	1419	1421	1421	1421
3	296	301	306	312	316	319	321	322	324	327
3	1523	1529	1539	1549	1556	1562	1571	1581	1591	1599
1	x9.12.20	x9.13.20	x9.14.20	x9.15.20	x9.16.20	x9.17.20	x9.18.20	x9.19.20	x9.20.20	
2	1421	1421	1427	1428	1438	1438	1439	1439	1443	
3	330	334	338	340	343	347	353	358	362	
3	1605	1612	1620	1632	1645	1654	1659	1665	1672	
1	x9.21.20	x9.22.20	x9.23.20	x9.24.20	x9.25.20	x9.26.20	x9.27.20	x9.28.20	x9.29.20	
2	1446	1447	1448	1453	1453	1455	1455	1458	1460	
3	364	367	370	370	373	375	377	380	384	
3	1679	1689	1698	1703	1707	1711	1714	1719	1726	
1	x9.30.20	x10.1.20	x10.2.20	x10.3.20	x10.4.20	x10.5.20	x10.6.20	x10.7.20	x10.8.20	
2	1460	1460	1460	1464	1464	1468	1469	1471	1472	
3	387	388	389	392	396	400	403	407	411	
3	1736	1741	1749	1756	1760	1768	1768	1771	1783	
1	x10.9.20	x10.10.20	x10.11.20	x10.12.20	x10.13.20	x10.14.20	x10.15.20	x10.16.20		
2	1474	1475	1479	1481	1482	1483	1483	1487		
3	413	416	420	424	429	434	439	443		
3	1789	1795	1801	1809	1818	1827	1827	1841		
1	x10.17.20	x10.18.20	x10.19.20	x10.20.20	x10.21.20	x10.22.20	x10.23.20	x10.24.20		
2	1490	1494	1499	1501	1503	1507	1509	1513		
3	448	451	454	458	462	465	469	473		
3	1846	1856	1865	1873	1880	1888	1897	1907		
1	x10.25.20	x10.26.20	x10.27.20	x10.28.20	x10.29.20	x10.30.20	x10.31.20	x11.1.20		
2	1516	1520	1525	1531	1534	1535	1538	1538		
3	477	480	487	493	499	502	509	518		
3	1914	1922	1931	1941	1949	1956	1964	1973		
1	x11.2.20	x11.3.20	x11.4.20	x11.5.20	x11.6.20	x11.7.20	x11.8.20	x11.9.20	x11.10.20	
2	1544	1547	1551	1557	1557	1559	1565	1577	1580	
3	527	532	536	543	549	557	559	571	579	
3	1980	1980	1999	2011	2024	2036	2048	2062	2077	
1	x11.11.20	x11.12.20	x11.13.20	x11.14.20	x11.15.20	x11.16.20	x11.17.20	x11.18.20		
2	1584	1594	1598	1608	1620	1635	1641	1648		
3	590	598	605	612	623	631	637	646		
3	2093	2111	2124	2139	2154	2168	2186	2206		
1	x11.19.20	x11.20.20	x11.21.20	x11.22.20	x11.23.20	x11.24.20	x11.25.20	x11.26.20		
2	1653	1666	1678	1690	1702	1715	1728	1737		
3	657	672	685	699	716	735	743	753		
3	2224	2236	2255	2272	2294	2309	2329	2352		

COVID-19 Data Analysis

```

x11.27.20 x11.28.20 x11.29.20 x11.30.20 x12.1.20 x12.2.20 x12.3.20 x12.4.20
1      1740      1752      1774      1795      1822      1841      1846      1846
2       771       787       798       810       822       839       852       870
3      2372      2393      2410      2431      2447      2464      2480      2492
x12.5.20 x12.6.20 x12.7.20 x12.8.20 x12.9.20 x12.10.20 x12.11.20
1      1864      1874      1900      1906      1919      1935      1945
2       889       905       922       936       951       965       977
3      2501      2516      2527      2539      2554      2564      2575
[ reached 'max' / getOption("max.print") -- omitted 3 rows ]

```

#Recovered case from time_series_covid19_recovered_global.csv

```

Province.State Country.Region Lat Long x1.22.20 x1.23.20 x1.24.20
1 <NA> Afghanistan 33.93911 67.70995 0 0 0
2 <NA> Albania 41.15330 20.16830 0 0 0
3 <NA> Algeria 28.03390 1.65960 0 0 0
x1.25.20 x1.26.20 x1.27.20 x1.28.20 x1.29.20 x1.30.20 x1.31.20 x2.1.20 x2.2.20
1 0 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0 0 0
x2.3.20 x2.4.20 x2.5.20 x2.6.20 x2.7.20 x2.8.20 x2.9.20 x2.10.20 x2.11.20
x2.12.20
1 0 0 0 0 0 0 0 0
0
2 0 0 0 0 0 0 0 0
0
3 0 0 0 0 0 0 0 0
0
x2.13.20 x2.14.20 x2.15.20 x2.16.20 x2.17.20 x2.18.20 x2.19.20 x2.20.20 x2.21.20
1 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0 0
x2.22.20 x2.23.20 x2.24.20 x2.25.20 x2.26.20 x2.27.20 x2.28.20 x2.29.20 x3.1.20
1 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0 0
x3.2.20 x3.3.20 x3.4.20 x3.5.20 x3.6.20 x3.7.20 x3.8.20 x3.9.20 x3.10.20 x3.11.20
1 0 0 0 0 0 0 0 0 0
2 0 0 0 0 0 0 0 0 0
3 0 0 0 0 0 0 0 0 0
x3.12.20 x3.13.20 x3.14.20 x3.15.20 x3.16.20 x3.17.20 x3.18.20 x3.19.20 x3.20.20
1 0 0 0 0 1 1 1 1
2 0 0 0 0 0 0 0 0 0
3 8 8 12 12 12 12 12 32 32
x3.21.20 x3.22.20 x3.23.20 x3.24.20 x3.25.20 x3.26.20 x3.27.20 x3.28.20 x3.29.20
1 1 1 1 1 2 2 2 2
2 2 2 2 10 17 17 31 31
3 32 65 65 24 65 29 29 31
x3.30.20 x3.31.20 x4.1.20 x4.2.20 x4.3.20 x4.4.20 x4.5.20 x4.6.20 x4.7.20 x4.8.20
1 2 5 5 10 10 15 18 18 29
2 44 52 67 76 89 99 104 116 131
3 37 46 61 61 62 90 90 113 237
x4.9.20 x4.10.20 x4.11.20 x4.12.20 x4.13.20 x4.14.20 x4.15.20 x4.16.20 x4.17.20
1 32 32 32 32 32 40 43 54 99
2 165 182 197 217 232 248 251 277 283
3 347 405 460 591 601 691 708 783 846
x4.18.20 x4.19.20 x4.20.20 x4.21.20 x4.22.20 x4.23.20 x4.24.20 x4.25.20 x4.26.20
1 112 131 135 150 166 179 188 188 207
2 302 314 327 345 356 385 394 403 410
3 894 1047 1099 1152 1204 1355 1408 1479 1508
x4.27.20 x4.28.20 x4.29.20 x4.30.20 x5.1.20 x5.2.20 x5.3.20 x5.4.20 x5.5.20
1 220 228 252 260 310 331 345 397 421

```

COVID-19 Data Analysis

2	422	431	455	470	488	519	531	543	570
3	1558	1651	1702	1779	1821	1872	1936	1998	2067
	x5.6.20	x5.7.20	x5.8.20	x5.9.20	x5.10.20	x5.11.20	x5.12.20	x5.13.20	x5.14.20
1	458	468	472	502	558	558	610	648	691
2	595	605	620	627	650	654	682	688	694
3	2197	2323	2467	2546	2678	2841	2998	3058	3158
	x5.15.20	x5.16.20	x5.17.20	x5.18.20	x5.19.20	x5.20.20	x5.21.20	x5.22.20	x5.23.20
1	745	745	778	801	850	930	938	996	1040
2	705	714	715	727	742	758	771	777	783
3	3271	3409	3507	3625	3746	3968	4062	4256	4426
	x5.24.20	x5.25.20	x5.26.20	x5.27.20	x5.28.20	x5.29.20	x5.30.20	x5.31.20	x6.1.20
1	1075	1097	1128	1138	1209	1259	1303	1328	1428
2	789	795	803	812	823	851	857	872	877
3	4784	4747	4918	5129	5277	5422	5549	5748	5894
	x6.2.20	x6.3.20	x6.4.20	x6.5.20	x6.6.20	x6.7.20	x6.8.20	x6.9.20	x6.10.20
1	1450	1522	1585	1762	1830	1875	2171	2651	3013
2	891	898	898	910	925	938	945	960	980
3	6067	6218	6297	6453	6631	6717	6799	6951	7074
	x6.12.20	x6.13.20	x6.14.20	x6.15.20	x6.16.20	x6.17.20	x6.18.20	x6.19.20	x6.20.20
1	3928	4201	4725	5164	5508	6158	7660	7962	8292
2	1034	1039	1044	1055	1064	1077	1086	1114	1126
3	7322	7420	7606	7735	7842	7943	8078	8196	8324
	x6.21.20	x6.22.20	x6.23.20	x6.24.20	x6.25.20	x6.26.20	x6.27.20	x6.28.20	x6.29.20
1	8764	8841	9260	9869	10174	10306	10674	12604	13934
2	1134	1159	1195	1217	1250	1298	1346	1384	1438
3	8422	8559	8674	8792	8920	9066	9202	9371	9674
	x6.30.20	x7.1.20	x7.2.20	x7.3.20	x7.4.20	x7.5.20	x7.6.20	x7.7.20	x7.8.20
1	14131	15651	16041	17331	19164	19366	20103	20179	20700
2	1459	1516	1559	1592	1637	1657	1702	1744	1791
3	9897	10040	10342	10832	11181	11492	11884	12094	12329
	x7.10.20	x7.11.20	x7.12.20	x7.13.20	x7.14.20	x7.15.20	x7.16.20	x7.17.20	x7.18.20
1	20882	21135	21216	21254	21454	22456	22824	23151	23273
2	1875	1881	1946	2014	2062	2091	2137	2214	2264
3	13124	13124	13743	14019	14295	14792	15107	15430	15744
	x7.19.20	x7.20.20	x7.21.20	x7.22.20	x7.23.20	x7.24.20	x7.25.20	x7.26.20	x7.27.20
1	23634	23741	23741	23924	24550	24602	24793	25180	25198
2	2311	2352	2397	2463	2523	2608	2637	2682	2745
3	16051	16400	16646	16983	17369	17369	18076	18088	18837
	x7.28.20	x7.29.20	x7.30.20	x7.31.20	x8.1.20	x8.2.20	x8.3.20	x8.4.20	x8.5.20
1	25358	25389	25471	25509	25509	25510	25669	25669	25742
2	2789	2830	2883	2952	2961	3018	3031	3031	3123
3	19233	19592	20082	20537	20988	21419	21901	22375	22802
	x8.6.20	x8.7.20	x8.8.20	x8.9.20	x8.10.20	x8.11.20	x8.12.20	x8.13.20	x8.14.20
1	25840	25903	25960	25960	26228	26415	26694	26714	26714
2	3155	3227	3268	3342	3379	3480	3552	3616	3695
3	23238	23667	24083	24506	24920	25263	25627	26004	26308
	x8.15.20	x8.16.20	x8.17.20	x8.18.20	x8.19.20	x8.20.20	x8.21.20	x8.22.20	x8.23.20
1	27166	27166	27166	27166	27166	27681	28016	28016	28180
2	3746	3794	3816	3871	3928	3986	4096	4184	4332
3	26644	27017	27347	27653	27971	28281	28587	28874	29142
	x8.24.20	x8.25.20	x8.26.20	x8.27.20	x8.28.20	x8.29.20	x8.30.20	x8.31.20	x9.1.20
1	28360	28440	29042	29046	29059	29063	29089	29089	29231
2	4413	4530	4633	4791	4923	5020	5139	5214	5441
3	29369	29587	29886	30157	30436	30717	30978	31244	31493
	x9.2.20	x9.3.20	x9.4.20	x9.5.20	x9.6.20	x9.7.20	x9.8.20	x9.9.20	x9.10.20
1	29315	29390	29713	30082	30537	30557	30715	31048	31129
2	5582	5732	5882	5976	6106	6186	6239	6284	6346
3	31746	32006	32259	32481	32745	32985	33183	33379	33562
	x9.12.20	x9.13.20	x9.14.20	x9.15.20	x9.16.20	x9.17.20	x9.18.20	x9.19.20	x9.20.20
1	31234	31638	32073	32098	32503	32505	32576	32576	32576
2	6494	6569	6615	6668	6733	6788	6831	6888	6940
3	33875	34037	34204	34385	34517	34675	34818	34923	35047
	x9.21.20	x9.22.20	x9.23.20	x9.24.20	x9.25.20	x9.26.20	x9.27.20	x9.28.20	x9.29.20
1	32576	32576	32610	32619	32619	32635	32642	32642	32746
2	6995	7042	7139	7239	7309	7397	7397	7629	7732

COVID-19 Data Analysis

```

3      35180      35307      35428      35544      35654      35756      35860      35962      36063
x9.30.20 x10.1.20 x10.2.20 x10.3.20 x10.4.20 x10.5.20 x10.6.20 x10.7.20 x10.8.20
1      32789      32842      32842      32842      32852      32879      32977      33045      33058
2      7847       8077       8342       8536       8675       8825       8965       9115       9215
3      36174      36282      36385      36482      36578      36672      36763      36857      36958
x10.9.20 x10.10.20 x10.11.20 x10.12.20 x10.13.20 x10.14.20 x10.15.20 x10.16.20
1      33058      33064      33114      33118      33308      33354      33447      33516
2      9304       9406       9500       9585       9675       9762       9864       9957
3      37067      37170      37170      37382      37492      37603      37603      37856
x10.17.20 x10.18.20 x10.19.20 x10.20.20 x10.21.20 x10.22.20 x10.23.20 x10.24.20
1      33561      33614      33760      33790      33824      33831      34010      34023
2      10001      10071      10167      10225      10341      10395      10466      10548
3      37971      38088      38215      38346      38482      38618      38788      38932
x10.25.20 x10.26.20 x10.27.20 x10.28.20 x10.29.20 x10.30.20 x10.31.20 x11.1.20
1      34129      34150      34217      34237      34239      34258      34321      34326
2      10654      10705      10808      10893      11007      11097      11189      11246
3      39095      39273      39444      39635      39635      40014      40201      40395
x11.2.20 x11.3.20 x11.4.20 x11.5.20 x11.6.20 x11.7.20 x11.8.20 x11.9.20 x11.10.20
1      34342      34355      34362      34440      34440      34446      34458      34721      34954
2      11367      11473      11578      11696      11861      12002      12092      12203      12353
3      40577      40577      41001      41244      41510      41783      42037      42325      42626
x11.11.20 x11.12.20 x11.13.20 x11.14.20 x11.15.20 x11.16.20 x11.17.20 x11.18.20
1      34967      35024      35036      35067      35092      35137      35160      35295
2      12493      12574      12667      12767      12889      13453      13804      14216
3      42980      42980      43779      44199      44633      45148      45148      46326
x11.19.20 x11.20.20 x11.21.20 x11.22.20 x11.23.20 x11.24.20 x11.25.20 x11.26.20
1      35350      35370      35422      35934      35976      36122      36145      36232
2      14565      15055      15469      15842      16230      16666      17031      17352
3      46962      47581      48183      48794      49421      50070      50712      51334
x11.27.20 x11.28.20 x11.29.20 x11.30.20 x12.1.20 x12.2.20 x12.3.20 x12.4.20
1      36295      36709      36716      36831      36946      37218      37260      37260
2      17755      18152      18481      18849      19384      19912      20484      20974
3      51946      52568      53204      53809      54405      54990      55538      56079
x12.5.20 x12.6.20 x12.7.20 x12.8.20 x12.9.20 x12.10.20 x12.11.20
1      37393      37685      37879      37920      38032      38099      38141
2      21286      21617      22180      22527      23072      23609      24136
3      56617      57146      57648      58146      58146      59135      59590
[ reached 'max' / getOption("max.print") -- omitted 3 rows ]

```

#Rawdata case from owid-covid-data.csv

iso_code	continent	location	date	total_cases	new_cases	new_cases_smoothed
1	AFG	Asia Afghanistan	2019-12-31			0
NA						
2	AFG	Asia Afghanistan	2020-01-01			0
NA						
3	AFG	Asia Afghanistan	2020-01-02			0
NA						
4	AFG	Asia Afghanistan	2020-01-03			0
NA						
5	AFG	Asia Afghanistan	2020-01-04			0
NA						
6	AFG	Asia Afghanistan	2020-01-05			0
NA						
	total_deaths	new_deaths	new_deaths_smoothed	total_cases_per_million		
1	0	0		NA	0	
2	0	0		NA	0	
3	0	0		NA	0	
4	0	0		NA	0	
5	0	0		NA	0	

COVID-19 Data Analysis

6	0	0	NA	0	
	new_cases_per_million	new_cases_smoothed_per_million	total_deaths_per_million		
1		0	NA	0	
2		0	NA	0	
3		0	NA	0	
4		0	NA	0	
5		0	NA	0	
6		0	NA	0	
	new_deaths_per_million	new_deaths_smoothed_per_million	new_tests	total_tests	
1		0	NA	NA	
2		0	NA	NA	
3		0	NA	NA	
4		0	NA	NA	
5		0	NA	NA	
6		0	NA	NA	
	total_tests_per_thousand	new_tests_per_thousand	new_tests_smoothed		
1		NA	NA	NA	
2		NA	NA	NA	
3		NA	NA	NA	
4		NA	NA	NA	
5		NA	NA	NA	
6		NA	NA	NA	
	new_tests_smoothed_per_thousand	tests_per_case	positive_rate	tests_units	
1		NA	NA	<NA>	
2		NA	NA	<NA>	
3		NA	NA	<NA>	
4		NA	NA	<NA>	
5		NA	NA	<NA>	
6		NA	NA	<NA>	
	stringency_index	population	population_density	median_age	aged_65_older
1	NA	38928341	54.422	18.6	2.581
2	0	38928341	54.422	18.6	2.581
3	0	38928341	54.422	18.6	2.581
4	0	38928341	54.422	18.6	2.581
5	0	38928341	54.422	18.6	2.581
6	0	38928341	54.422	18.6	2.581
	aged_70_older	gdp_per_capita	extreme_poverty	cardiovasc_death_rate	
1	1.337	1803.987	NA	597.029	
2	1.337	1803.987	NA	597.029	
3	1.337	1803.987	NA	597.029	
4	1.337	1803.987	NA	597.029	
5	1.337	1803.987	NA	597.029	
6	1.337	1803.987	NA	597.029	
	diabetes_prevalence	female_smokers	male_smokers	handwashing_facilities	
1	9.59	NA	NA	37.746	
2	9.59	NA	NA	37.746	
3	9.59	NA	NA	37.746	
4	9.59	NA	NA	37.746	
5	9.59	NA	NA	37.746	
6	9.59	NA	NA	37.746	
	hospital_beds_per_thousand	life_expectancy			
1		0.5	64.83		
2		0.5	64.83		
3		0.5	64.83		
4		0.5	64.83		
5		0.5	64.83		
6		0.5	64.83		

COVID-19 Data Analysis

Result 2:

From the output Below Canada as a region appears multiple times, and that's due to the inclusion of states, there are other countries where the states are included as well. To get a list of all the countries with different states we can check to see if the State column is null or not.

#Note: output shown is taken from Rstudio out of n rows 3 rows are displayed with n columns and x1.22.20 & so on be date as mentioned in csv file.

Province.State	Country	Region	Lat	Long	x1.22.20	x1.23.20	x1.24.20			
40	Alberta	Canada	53.9333	-116.5765		0	0	0		
41	British Columbia	Canada	53.7267	-127.6476		0	0	0		
42	Diamond Princess	Canada	0.0000	0.0000		0	0	0		
	x1.25.20	x1.26.20	x1.27.20	x1.28.20	x1.29.20	x1.30.20	x1.31.20	x2.1.20	x2.2.20	
40	0	0	0	0	0	0	0	0	0	
41	0	0	0	1	1	1	1	1	1	
42	0	0	0	0	0	0	0	0	0	
	x2.3.20	x2.4.20	x2.5.20	x2.6.20	x2.7.20	x2.8.20	x2.9.20	x2.10.20	x2.11.20	
40	0	0	0	0	0	0	0	0	0	
41	1	1	2	2	4	4	4	4	4	
42	0	0	0	0	0	0	0	0	0	
	x2.12.20	x2.13.20	x2.14.20	x2.15.20	x2.16.20	x2.17.20	x2.18.20	x2.19.20	x2.20.20	
40	0	0	0	0	0	0	0	0	0	
41	4	4	4	4	4	5	5	5	5	
42	0	0	0	0	0	0	0	0	0	
	x2.21.20	x2.22.20	x2.23.20	x2.24.20	x2.25.20	x2.26.20	x2.27.20	x2.28.20	x2.29.20	
40	0	0	0	0	0	0	0	0	0	
41	6	6	6	6	7	7	7	7	8	
42	0	0	0	0	0	0	0	0	0	
	x3.1.20	x3.2.20	x3.3.20	x3.4.20	x3.5.20	x3.6.20	x3.7.20	x3.8.20	x3.9.20	x3.10.20
40	0	0	0	0	0	1	2	4	7	7
41	8	8	9	12	13	21	21	27	32	32
42	0	0	0	0	0	0	0	0	0	0
	x3.11.20	x3.12.20	x3.13.20	x3.14.20	x3.15.20	x3.16.20	x3.17.20	x3.18.20	x3.19.20	
40	19	19	29	29	39	56	74	97	119	
41	39	46	64	64	73	103	103	186	231	
42	0	0	0	0	0	0	0	0	0	
	x3.20.20	x3.21.20	x3.22.20	x3.23.20	x3.24.20	x3.25.20	x3.26.20	x3.27.20	x3.28.20	
40	146	195	259	301	359	358	486	542	542	
41	271	424	424	472	617	617	725	725	884	
42	0	0	0	0	0	0	0	0	0	
	x3.29.20	x3.30.20	x3.31.20	x4.1.20	x4.2.20	x4.3.20	x4.4.20	x4.5.20	x4.6.20	
40	621	661	690	754	969	969	1075	1181	1250	
41	884	970	1013	1013	1121	1174	1203	1203	1266	
42	0	0	0	0	0	0	0	0	0	
	x4.7.20	x4.8.20	x4.9.20	x4.10.20	x4.11.20	x4.12.20	x4.13.20	x4.14.20	x4.15.20	
40	1373	1373	1423	1451	1567	1567	1732	1870	1870	
41	1266	1291	1336	1370	1445	1445	1490	1490	1517	
42	0	0	0	0	0	0	0	0	0	
	x4.16.20	x4.17.20	x4.18.20	x4.19.20	x4.20.20	x4.21.20	x4.22.20	x4.23.20	x4.24.20	
40	1996	2397	2562	2803	2908	3095	3401	3720	4017	
41	1561	1575	1618	1647	1647	1724	1795	1824	1853	
42	0	0	0	0	0	0	0	0	0	
	x4.25.20	x4.26.20	x4.27.20	x4.28.20	x4.29.20	x4.30.20	x5.1.20	x5.2.20	x5.3.20	
40	4233	4480	4696	4850	5165	5355	5573	5670	5766	
41	1948	1948	1998	2053	2087	2112	2145	2171	2171	
42	0	0	0	0	0	0	1	1	1	
	x5.4.20	x5.5.20	x5.6.20	x5.7.20	x5.8.20	x5.9.20	x5.10.20	x5.11.20	x5.12.20	
40	5836	5893	5963	6017	6098	6157	6253	6300	6345	

COVID-19 Data Analysis

41	2224	2232	2255	2288	2315	2330	2330	2353	2360
42	1	1	1	1	1	1	1	1	1
40	x5.13.20	x5.14.20	x5.15.20	x5.16.20	x5.17.20	x5.18.20	x5.19.20	x5.20.20	x5.21.20
41	6407	6457	6515	6587	6644	6683	6716	6735	6768
42	2376	2392	2407	2428	2428	2444	2446	2467	2479
40	1	1	1	1	1	1	1	1	1
41	x5.22.20	x5.23.20	x5.24.20	x5.25.20	x5.26.20	x5.27.20	x5.28.20	x5.29.20	x5.30.20
42	6800	6818	6860	6879	6901	6926	6955	6979	6992
40	2507	2517	2517	2530	2541	2550	2558	2562	2573
41	1	1	1	1	1	1	1	1	1
42	x5.31.20	x6.1.20	x6.2.20	x6.3.20	x6.4.20	x6.5.20	x6.6.20	x6.7.20	x6.8.20
40	7010	7044	7057	7076	7091	7098	7138	7138	7202
41	2573	2597	2601	2623	2632	2632	2632	2632	2659
42	1	1	0	0	0	0	0	0	0
40	x6.10.20	x6.11.20	x6.12.20	x6.13.20	x6.14.20	x6.15.20	x6.16.20	x6.17.20	x6.18.20
41	7276	7316	7346	7383	7433	7453	7482	7530	7579
42	2680	2694	2709	2709	2709	2745	2756	2775	2783
40	0	0	0	0	0	0	0	0	0
41	x6.19.20	x6.20.20	x6.21.20	x6.22.20	x6.23.20	x6.24.20	x6.25.20	x6.26.20	x6.27.20
42	7625	7673	7704	7736	7781	7825	7851	7888	7957
40	2790	2790	2790	2822	2835	2849	2869	2878	2878
41	0	0	0	0	0	0	0	0	0
42	x6.28.20	x6.29.20	x6.30.20	x7.1.20	x7.2.20	x7.3.20	x7.4.20	x7.5.20	x7.6.20
40	7996	8067	8108	8108	8202	8259	8259	8259	8389
41	2878	2904	2916	2916	2940	2947	2947	2947	2978
42	0	0	0	0	0	0	0	0	0
40	x7.7.20	x7.8.20	x7.9.20	x7.10.20	x7.11.20	x7.12.20	x7.13.20	x7.14.20	x7.15.20
41	8436	8482	8519	8596	8596	8596	8826	8912	8994
42	2990	3008	3028	3053	3053	3053	3053	3128	3149
40	0	0	0	0	0	0	0	0	0
41	x7.16.20	x7.17.20	x7.18.20	x7.19.20	x7.20.20	x7.21.20	x7.22.20	x7.23.20	x7.24.20
42	9114	9219	9219	9219	9587	9728	9728	9975	10086
40	3170	3198	3198	3198	3300	3328	3328	3392	3392
41	0	0	0	0	0	0	0	0	0
42	x7.25.20	x7.26.20	x7.27.20	x7.28.20	x7.29.20	x7.30.20	x7.31.20	x8.1.20	x8.2.20
40	10086	10086	10390	10470	10603	10716	10843	10843	10843
41	3419	3419	3500	3523	3562	3591	3641	3641	3641
42	0	0	0	0	0	0	0	0	0
40	x8.3.20	x8.4.20	x8.5.20	x8.6.20	x8.7.20	x8.8.20	x8.9.20	x8.10.20	x8.11.20
41	10843	11146	11240	11296	11430	11430	11430	11687	11772
42	3641	3787	3834	3881	3934	3934	3934	4065	4111
40	0	0	0	0	0	0	0	0	0
41	x8.12.20	x8.13.20	x8.14.20	x8.15.20	x8.16.20	x8.17.20	x8.18.20	x8.19.20	x8.20.20
42	11893	11969	12053	12053	12053	12053	12419	12501	12501
40	4111	4274	4358	4358	4358	4358	4677	4745	4745
41	0	0	0	0	0	0	0	0	0
42	x8.21.20	x8.22.20	x8.23.20	x8.24.20	x8.25.20	x8.26.20	x8.27.20	x8.28.20	x8.29.20
40	12748	12748	12748	13006	13083	13210	13318	13476	13476
41	4915	4915	4915	5184	5242	5304	5372	5496	5496
42	0	0	0	0	0	0	0	0	0
40	x8.30.20	x8.31.20	x9.1.20	x9.2.20	x9.3.20	x9.4.20	x9.5.20	x9.6.20	x9.7.20
41	13476	13902	14066	14180	14310	14474	14474	14474	14474
42	15093	5496	5790	5848	5952	6041	6162	6162	6162
40	6591	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	0	0	0
42	x9.9.20	x9.10.20	x9.11.20	x9.12.20	x9.13.20	x9.14.20	x9.15.20	x9.16.20	x9.17.20
40	15093	15304	15415	15415	15415	15833	15957	16128	16274
41	6591	6830	6962	6962	6962	7279	7376	7498	7663
42	0	0	0	0	0	0	0	0	0
40	x9.18.20	x9.19.20	x9.20.20	x9.21.20	x9.22.20	x9.23.20	x9.24.20	x9.25.20	x9.26.20
41	16381	16381	16381	16739	16889	17032	17190	17343	17343
42	7842	7842	7842	8208	8304	8395	8395	8641	8641

COVID-19 Data Analysis

```

42      0      0      0      0      0      0      0      0      0
40 x9.27.20 x9.28.20 x9.29.20 x9.30.20 x10.1.20 x10.2.20 x10.3.20 x10.4.20 x10.5.20
41      17343      17749      17909      18062      18235      18357      18357      18357      18935
42      8641      8908      9013      9138      9220      9381      9381      9381      9739
42      0      0      0      0      0      0      0      0      0
40 x10.6.20 x10.7.20 x10.8.20 x10.9.20 x10.10.20 x10.11.20 x10.12.20 x10.13.20
41      19211      19354      19718      19995      19995      19995      19995      20956
42      9841      9956      10066      10185      10185      10185      10185      10734
42      0      0      0      0      0      0      0      0      0
40 x10.14.20 x10.15.20 x10.16.20 x10.17.20 x10.18.20 x10.19.20 x10.20.20 x10.21.20
41      21199      21443      21775      21775      21775      22673      22996      23402
42      10892      11034      11189      11189      11189      11687      11854      12057
42      0      0      0      0      0      0      0      0      0
40 x10.22.20 x10.23.20 x10.24.20 x10.25.20 x10.26.20 x10.27.20 x10.28.20 x10.29.20
41      23829      24261      24261      24261      25733      26155      26565      27042
42      12057      12554      12554      12554      13371      13588      13875      14109
42      0      0      0      0      0      0      0      0      0
40 x10.30.20 x10.31.20 x11.1.20 x11.2.20 x11.3.20 x11.4.20 x11.5.20 x11.6.20
41      27664      27664      27664      27664      29932      29932      30447      31858
42      14381      14381      14381      15501      15800      16135      16560      17149
42      0      0      0      0      0      0      0      0      0
40 x11.7.20 x11.8.20 x11.9.20 x11.10.20 x11.11.20 x11.12.20 x11.13.20 x11.14.20
41      32777      33504      34160      34873      35545      36405      37312      38338
42      17716      17716      18714      19239      19239      20369      20986      20986
42      0      0      0      0      0      0      0      0      0
40 x11.15.20 x11.16.20 x11.17.20 x11.18.20 x11.19.20 x11.20.20 x11.21.20 x11.22.20
41      39329      40189      40962      41692      42797      43952      45288      46872
42      20986      22945      23662      24422      24960      25474      25474      25474
42      0      0      0      0      0      0      0      0      0
40 x11.23.20 x11.24.20 x11.25.20 x11.26.20 x11.27.20 x11.28.20 x11.29.20 x11.30.20
41      48421      49536      50801      51878      53105      54836      56444      58177
42      27407      28348      29086      29973      30884      30884      30884      33238
42      0      0      0      0      0      0      0      0      0
40 x12.1.20 x12.2.20 x12.3.20 x12.4.20 x12.5.20 x12.6.20 x12.7.20 x12.8.20 x12.9.20
41      59484      61169      63023      64851      66730      68566      70301      72028      73488
42      33894      34728      35422      36132      36132      36132      38152      38718      39337
42      0      0      0      0      0      0      0      0      0
40 x12.10.20 x12.11.20
41      75054      76792
42      40060      40797
42      0      0
[ reached 'max' / getOption("max.print") -- omitted 13 rows ]

```

Result 3:

Using unique() we will be displaying only Country regions belonging to time_series_covid19_confirmed_global.csv.

```

[1] "Australia"      "Canada"          "China"           "Denmark"
[5] "France"         "Netherlands"     "United kingdom"

```

COVID-19 Data Analysis

Result 4:

We have created data frames for each type of cases and for each country, and to make it easier to work with its better to merge those data frames into a more encompassing one.

Province.State	Country	Region	Lat	Long	x1.22.20	x1.23.20	x1.24.20	x1.25.20	
248	<NA>	US	40	-100	1	1	2	2	
	x1.26.20	x1.27.20	x1.28.20	x1.29.20	x1.30.20	x1.31.20	x2.1.20	x2.2.20	x2.3.20
248	5	5	5	6	6	8	8	8	11
	x2.4.20	x2.5.20	x2.6.20	x2.7.20	x2.8.20	x2.9.20	x2.10.20	x2.11.20	x2.12.20
248	11	11	12	12	12	12	12	13	13
	x2.13.20	x2.14.20	x2.15.20	x2.16.20	x2.17.20	x2.18.20	x2.19.20	x2.20.20	
x2.21.20									
248	14	14	14	14	14	14	14	14	14
16									
	x2.22.20	x2.23.20	x2.24.20	x2.25.20	x2.26.20	x2.27.20	x2.28.20	x2.29.20	x3.1.20
248	16	16	16	16	16	17	17	25	32
	x3.2.20	x3.3.20	x3.4.20	x3.5.20	x3.6.20	x3.7.20	x3.8.20	x3.9.20	x3.10.20
248	55	74	107	184	237	403	519	594	782
	x3.11.20	x3.12.20	x3.13.20	x3.14.20	x3.15.20	x3.16.20	x3.17.20	x3.18.20	
x3.19.20									
248	1147	1586	2219	2978	3212	4679	6512	9169	
13663									
	x3.20.20	x3.21.20	x3.22.20	x3.23.20	x3.24.20	x3.25.20	x3.26.20	x3.27.20	
x3.28.20									
248	20030	26025	34855	46086	56698	68773	86613	105293	
124900									
	x3.29.20	x3.30.20	x3.31.20	x4.1.20	x4.2.20	x4.3.20	x4.4.20	x4.5.20	x4.6.20
248	143779	165861	192177	218060	248447	280417	313432	341629	371802
	x4.7.20	x4.8.20	x4.9.20	x4.10.20	x4.11.20	x4.12.20	x4.13.20	x4.14.20	x4.15.20
248	403212	435407	469989	503474	532782	559709	585518	614082	644247
	x4.16.20	x4.17.20	x4.18.20	x4.19.20	x4.20.20	x4.21.20	x4.22.20	x4.23.20	
x4.24.20									
248	675648	708317	736244	761933	790353	816413	845727	878911	
912662									
	x4.25.20	x4.26.20	x4.27.20	x4.28.20	x4.29.20	x4.30.20	x5.1.20	x5.2.20	x5.3.20
248	944234	971078	994265	1018926	1046737	1076224	1110464	1138228	1162685
	x5.4.20	x5.5.20	x5.6.20	x5.7.20	x5.8.20	x5.9.20	x5.10.20	x5.11.20	x5.12.20
248	1186067	1210577	1235666	1263402	1290151	1315099	1333970	1353397	1376122
	x5.13.20	x5.14.20	x5.15.20	x5.16.20	x5.17.20	x5.18.20	x5.19.20	x5.20.20	
x5.21.20									
248	1397085	1424243	1449498	1473514	1491829	1513816	1534871	1557933	
1583798									
	x5.22.20	x5.23.20	x5.24.20	x5.25.20	x5.26.20	x5.27.20	x5.28.20	x5.29.20	
x5.30.20									
248	1607109	1628212	1648158	1666505	1685956	1704489	1727357	1751612	
1775428									
	x5.31.20	x6.1.20	x6.2.20	x6.3.20	x6.4.20	x6.5.20	x6.6.20	x6.7.20	x6.8.20
x6.9.20									
248	1794465	1811393	1832782	1852788	1874156	1899505	1920910	1938572	1956064
1974425									
	x6.10.20	x6.11.20	x6.12.20	x6.13.20	x6.14.20	x6.15.20	x6.16.20	x6.17.20	
x6.18.20									
248	1995376	2018418	2043332	2068600	2087619	2107075	2130768	2157267	
2185134									
	x6.19.20	x6.20.20	x6.21.20	x6.22.20	x6.23.20	x6.24.20	x6.25.20	x6.26.20	
x6.27.20									

COVID-19 Data Analysis

248 2216115 2248032 2274000 2304676 2340884 2375310 2415666 2461019
2502378
x6.28.20 x6.29.20 x6.30.20 x7.1.20 x7.2.20 x7.3.20 x7.4.20 x7.5.20 x7.6.20
248 2542787 2582844 2628824 2680312 2735900 2787741 2833290 2882870 2927032
x7.7.20 x7.8.20 x7.9.20 x7.10.20 x7.11.20 x7.12.20 x7.13.20 x7.14.20 x7.15.20
248 2987787 3047527 3110094 3177953 3237991 3297076 3355869 3423296 3490931
x7.16.20 x7.17.20 x7.18.20 x7.19.20 x7.20.20 x7.21.20 x7.22.20 x7.23.20
x7.24.20
248 3568033 3639705 3702175 3762891 3824368 3888607 3960522 4029033
4102241
x7.25.20 x7.26.20 x7.27.20 x7.28.20 x7.29.20 x7.30.20 x7.31.20 x8.1.20 x8.2.20
248 4167616 4222480 4278427 4344212 4416044 4483643 4551554 4609393 4655662
x8.3.20 x8.4.20 x8.5.20 x8.6.20 x8.7.20 x8.8.20 x8.9.20 x8.10.20 x8.11.20
248 4700229 4757512 4811945 4871426 4929748 4984197 5030566 5079299 5126209
x8.12.20 x8.13.20 x8.14.20 x8.15.20 x8.16.20 x8.17.20 x8.18.20 x8.19.20
x8.20.20
248 5183020 5234762 5299526 5346316 5387497 5423981 5469015 5516408
5560495
x8.21.20 x8.22.20 x8.23.20 x8.24.20 x8.25.20 x8.26.20 x8.27.20 x8.28.20
x8.29.20
248 5608748 5652401 5686911 5723343 5763322 5808842 5853962 5900891
5946305
x8.30.20 x8.31.20 x9.1.20 x9.2.20 x9.3.20 x9.4.20 x9.5.20 x9.6.20 x9.7.20
248 5981092 6016230 6057907 6098858 6142639 6192768 6235960 6267471 6291016
x9.8.20 x9.9.20 x9.10.20 x9.11.20 x9.12.20 x9.13.20 x9.14.20 x9.15.20 x9.16.20
248 6317861 6351756 6387822 6435261 6476264 6510569 6544411 6583796 6622256
x9.17.20 x9.18.20 x9.19.20 x9.20.20 x9.21.20 x9.22.20 x9.23.20 x9.24.20
x9.25.20
248 6667178 6715928 6758440 6796884 6848379 6887861 6926442 6972152
7020206
x9.26.20 x9.27.20 x9.28.20 x9.29.20 x9.30.20 x10.1.20 x10.2.20 x10.3.20
x10.4.20
248 7064600 7101880 7134164 7176856 7218052 7263409 7317878 7366505
7402597
x10.5.20 x10.6.20 x10.7.20 x10.8.20 x10.9.20 x10.10.20 x10.11.20 x10.12.20
248 7441432 7485381 7536022 7594351 7650345 7704419 7750367 7790890
x10.13.20 x10.14.20 x10.15.20 x10.16.20 x10.17.20 x10.18.20 x10.19.20 x10.20.20
248 7843769 7903242 7967600 8036049 8093003 8142357 8209226 8270314
x10.21.20 x10.22.20 x10.23.20 x10.24.20 x10.25.20 x10.26.20 x10.27.20 x10.28.20
248 8333340 8408675 8489662 8572483 8634285 8700789 8776790 8855474
x10.29.20 x10.30.20 x10.31.20 x11.1.20 x11.2.20 x11.3.20 x11.4.20 x11.5.20
248 8945642 9044161 9133098 9208669 9292156 9416709 9519582 9645772
x11.6.20 x11.7.20 x11.8.20 x11.9.20 x11.10.20 x11.11.20 x11.12.20 x11.13.20
248 9771480 9899516 10009542 10130752 10268446 10412095 10573242 10750810
x11.14.20 x11.15.20 x11.16.20 x11.17.20 x11.18.20 x11.19.20 x11.20.20 x11.21.20
248 10917560 11053501 11211321 11371895 11542408 11730441 11925959 12104056
x11.22.20 x11.23.20 x11.24.20 x11.25.20 x11.26.20 x11.27.20 x11.28.20 x11.29.20
248 12246863 12418378 12591366 12772653 12883307 13088821 13244701 13383334
x11.30.20 x12.1.20 x12.2.20 x12.3.20 x12.4.20 x12.5.20 x12.6.20 x12.7.20
248 13541185 13721822 13921929 14139609 14367437 14581318 14756982 14949417
x12.8.20 x12.9.20 x12.10.20 x12.11.20
248 15165295 15386562 15611014 15842789

Result 5:

Data frames we can see that it's represented as a single row with the date values as columns, this format is not ideal and cannot be graphed or used in model creation as shown below only for chosen country as mentioned in question.

```
> head(us_df)                                     #for United States
  Date Confirmed Deaths Recovered
1 2020-01-22      1      0         0
2 2020-01-23      1      0         0
3 2020-01-24      2      0         0
4 2020-01-25      2      0         0
5 2020-01-26      5      0         0
6 2020-01-27      5      0         0
> head(germany_df)                                #for Germany
  Date Confirmed Deaths Recovered
1 2020-01-22      0      0         0
2 2020-01-23      0      0         0
3 2020-01-24      0      0         0
4 2020-01-25      0      0         0
5 2020-01-26      0      0         0
6 2020-01-27      1      0         0
> head(italy_df)                                   #for Italy
  Date Confirmed Deaths Recovered
1 2020-01-22      0      0         0
2 2020-01-23      0      0         0
3 2020-01-24      0      0         0
4 2020-01-25      0      0         0
5 2020-01-26      0      0         0
6 2020-01-27      0      0         0
> head(sk_df)                                       #for South Korea
  Date Confirmed Deaths Recovered
1 2020-01-22      1      0         0
2 2020-01-23      1      0         0
3 2020-01-24      2      0         0
4 2020-01-25      2      0         0
5 2020-01-26      3      0         0
6 2020-01-27      4      0         0
```

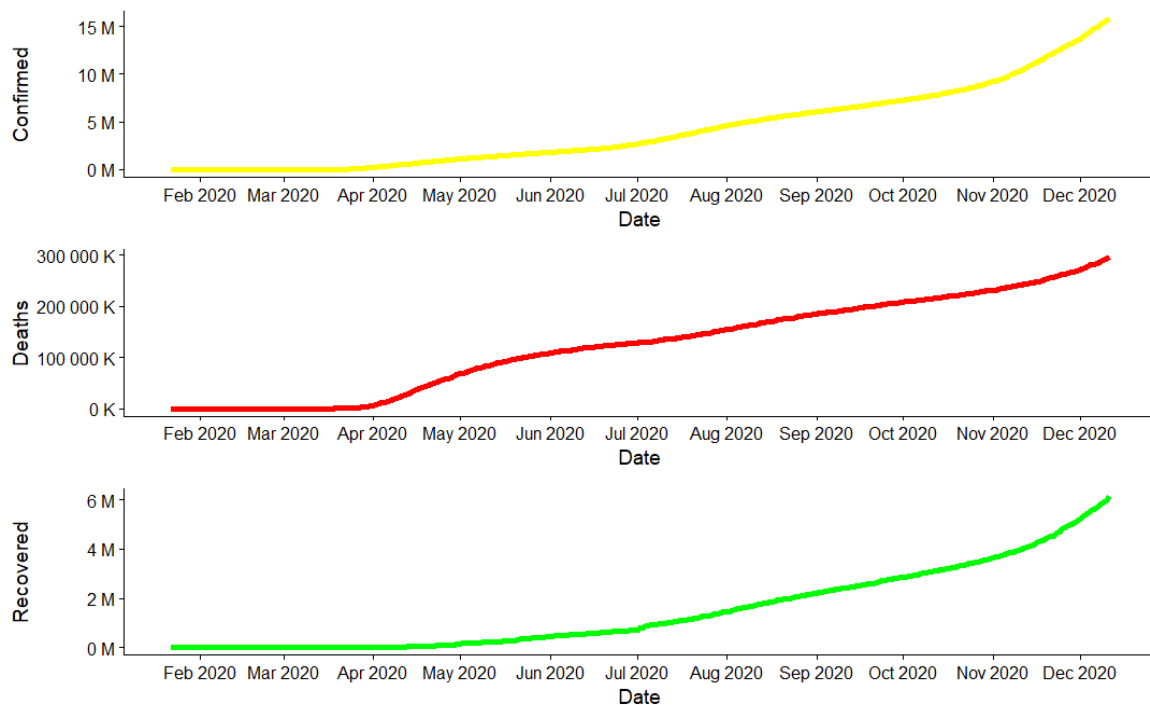
Result 6:

In USA the number of confirmed cases started up slow until around April, it started to go up at a much faster rate and it kept that pace even during quarantine, in July the rate at which the cases are increasing got higher and the cases started increasing faster, this can be attributed to the recent protests and people's ignorance to the CDC guidelines, as of writing this on July 22nd the total number of cases is 4.4M.

Deaths are the only cases that have had a continuously increasing rate, all the way from April the number of deaths is increasing at an increasing rate, till Dec 2020 it was 136K.

when it comes to the recoveries of the 4.4M cases only 1.3M have recovered so far, the recovery starts at the same time as the confirmed cases with a very unstable increase rate, the highest increase rate is also from around July which is surprising considering the rate of confirmed cases also went up around that time.

Hence this Range is described in graph shown below:



Number of cases which had an outcome: 6430764
percentage of cases that had an outcome: 40.59%
Deaths rate: 4.59%
Recovery rate: 95.41%
Currently Active cases: 9412025

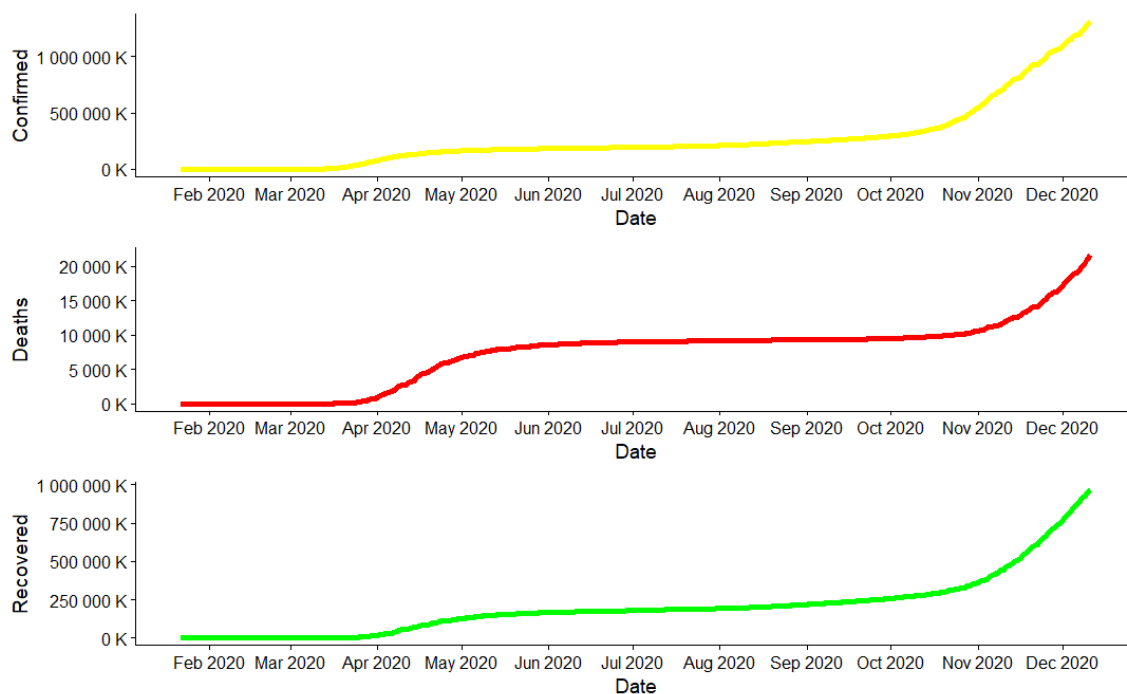
Result 7:

When it comes to Germany, the virus started creeping up early compared to the USA, the first confirmed case was around March 20th after that the cases increased at a high rate until May where the increase rate stabilized and the situation is much better, as of July the number of confirmed cases is around 207k.

Deaths wise, Germany is also doing very well, in the beginning, the numbers were going up quickly, but as of July the death toll seems to be stable and rarely increasing daily, till Dec 2020 it was 9.1k.

when it comes to recovery, the recovery rate started increasing as of April and it's still going up steadily if the current progress is maintained Germany will be corona free by the end of this year, currently, the number of recoveries is 190k.

Hence this Range is described in graph shown below:



Number of cases which had an outcome: 987805
percentage of cases that had an outcome: 75.16%
Death rate: 2.18%
Recovery rate: 97.82%
Currently Active cases: 326504

COVID-19 Data Analysis

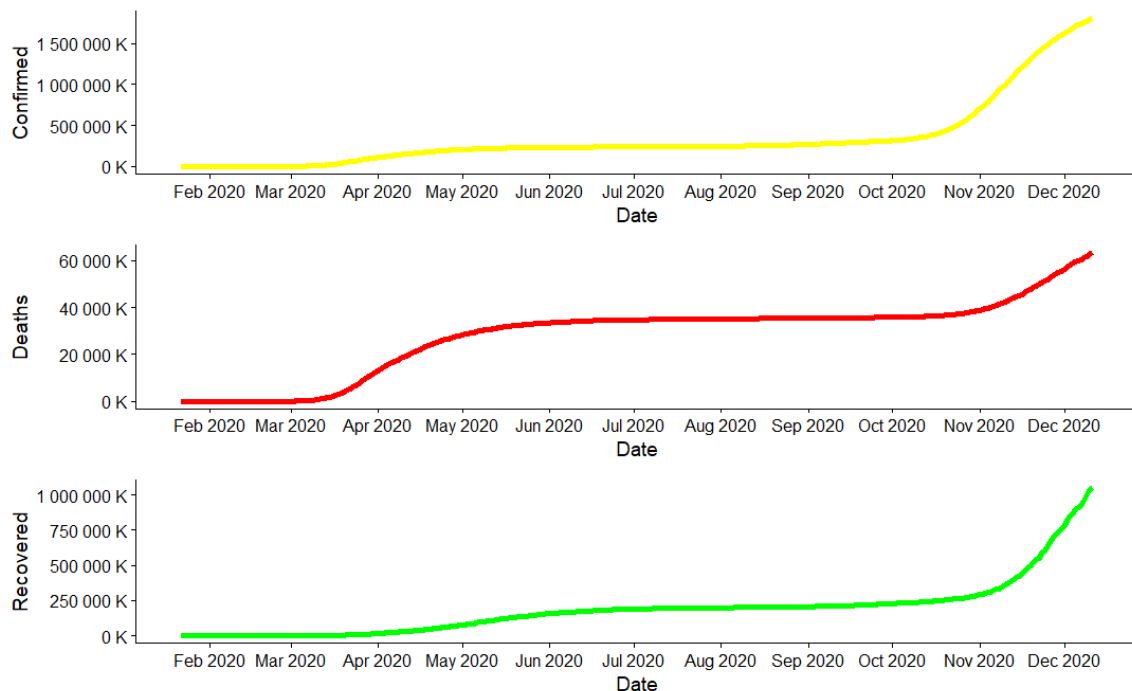
Result 8:

When the virus first appeared in Italy in the middle of March and the numbers started to suddenly increase very fast, especially the number of deaths, when it comes to the number of confirmed cases, it has stabilized in June and for the past 2 months, Italy only had a few daily cases.

the death rate is the first thing that jumps up when you look at data about Italy, the number started increasing exponentially all the way until June, from June and until now it's been much less daily deaths compared to earlier months.

When it comes to recoveries, the numbers are similar to confirmed cases minus the deaths, meaning the recovery was quite fast if you take into account the high numbers in the first few weeks.

Hence this Range is described in graph shown below:



Number of cases which had an outcome: 1115550
percentage of cases that had an outcome: 61.77%
Death rate: 5.68%
Recovery rate: 94.32%
Currently Active cases: 690323

COVID-19 Data Analysis

Result 9:

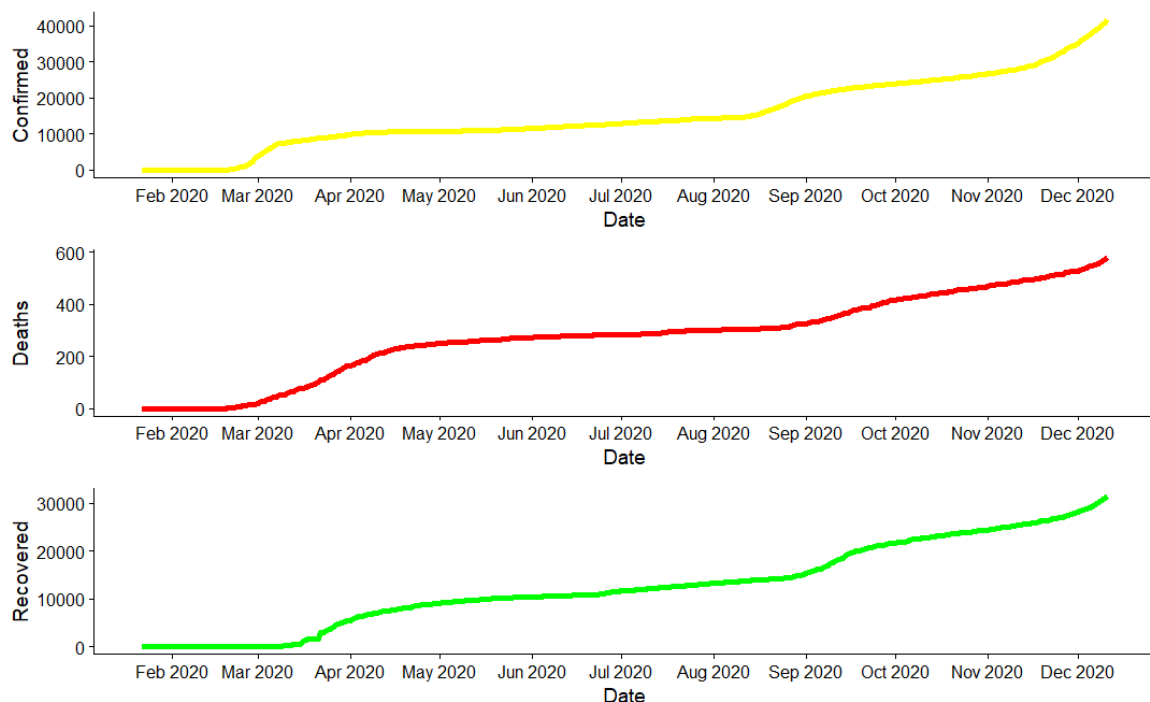
South Korea has the most diverse graph of the 4 countries we are analyzing, each graph has its own shape and points.

COVID-19 appeared in South Korea very early at the end of February due to its proximity to China, the numbers went up at a very high rate at the beginning of March and started stabilizing and going down around April until they started increasing again in June.

The number of deaths seems under control in South Korea, at the start it was slowly creeping up all the way to mid-April where it increasing slower until around may where the numbers were almost stable.

When it comes to recoveries, the graphs fluctuate very much until mid-march where it starts increasing exponentially all the way to the end of June where the numbers suddenly jump and stay on that pace.

Hence this Range is described in graph shown below:



Number of cases which had an outcome: 32071
percentage of cases that had an outcome: 76.84%
Death rate: 1.8%
Recovery rate: 98.2%
Currently Active cases: 9665

CONCLUSION AND FUTURE SCOPE

Analyzing COVID-19 data based on recovery, deaths, and confirmed cases using R can provide insights into the impact of the virus on different aspects. Here's a structure for drawing conclusions and outlining the scope of your analysis:

Conclusion:

1. Mortality and Recovery Rates:

- **Case Fatality Rate (CFR):** Calculate the percentage of deaths among confirmed cases.
- **Recovery Rate:** Assess the proportion of individuals who have recovered from the virus.

2. Temporal Trends:

- **Rate of Change:** Evaluate how recovery, death, and confirmed case numbers have changed over time.
- **Fluctuations:** Identify periods of significant changes in recovery and mortality rates.

3. Demographic Variations:

- **Age Groups:** Examine how recovery and mortality rates vary across different age groups.
- **Comorbidities:** Investigate the impact of underlying health conditions on outcomes.

4. Healthcare System Capacity:

- **Hospitalization Rates:** Analyze the burden on healthcare facilities concerning recovery and mortality.
- **Critical Care:** Assess the demand for critical care resources based on outcomes.

5. Geographical Patterns:

- **Regional Disparities:** Identify variations in recovery and mortality rates among different geographic regions.
- **Hotspots:** Explore areas with higher mortality or slower recovery rates.

6. Public Health Measures:

- **Effectiveness:** Assess the impact of public health interventions on recovery and mortality rates.
- **Vaccination Impact:** Explore the association between vaccination rates and outcomes.

Scope:

1. Further Research:

- **Long-Term Effects:** Investigate potential long-term health effects on individuals who have recovered.
- **Immunity Duration:** Study the duration of immunity post-recovery.
- **Factors Influencing Recovery:** Explore factors contributing to a faster or slower recovery.

2. Policy Implications:

- **Resource Allocation:** Provide insights for optimizing resource allocation in healthcare systems.
- **Targeted Interventions:** Support the development of interventions based on recovery and mortality patterns.

3. Healthcare System Strengthening:

- **Capacity Building:** Suggest measures to enhance healthcare system capacity for better outcomes.
- **Training and Preparedness:** Identify areas for training and preparedness in dealing with severe cases.

4. Community Education:

- **Risk Communication:** Develop strategies for communicating risks associated with severe outcomes.
- **Preventive Measures:** Emphasize the importance of preventive measures for vulnerable populations.

5. Global Comparisons:

- **Cross-Country Analysis:** Compare recovery and mortality rates across different countries.
- **Learning from Success/Failure:** Identify successful strategies in reducing mortality and promoting recovery.

6. Continuous Monitoring:

- **Surveillance Systems:** Enhance surveillance systems for early detection of outbreaks and monitoring recovery patterns.
- **Adaptive Strategies:** Be prepared to adapt strategies based on evolving recovery and mortality data.

Future Scope:

- Develop an interactive dashboard using tools, allowing users to interact with the data, customize visualization, and gain deeper insights.
- Create user-friendly reports or visualizations that can be easily interpreted by policymakers, economists, or other stakeholders, facilitating data-driven decision-making.

REFERENCES

- [1] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8404394/>
- [2] [R Tutorial: Analyzing COVID-19 Data | by Mykyta Solonko | Towards Data Science](#)
- [3] <https://www.kaggle.com/code/abubakaryagob/covid-19-data-analysis-using-r/input>
- [4] [The COVID19 package, an interface to the COVID-19 Data Hub · R Views \(rstudio.com\)](#)