Sample Report

Created by Diego Velandia
Publication date: 11/2/2022

Introduction

Sample datasets used as test subject for the present report describes evolution of confirmed positive cases, deceases and recovered patients along the current Covid-19 epidemics. Datasets are publicly available thanks to the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.

Datasets come in separate .csv documents and they can be downloaded through CSSE public repository on github.

- Positive cases
- Confirmed deceases
- Recovered patients

R Markdown file is written in a way that users do not have to scroll into dense monolithic code to search for a specific code snippet to modify. Instead, individual file management allow users to quickly navigate, understand, modify or even change the coding structure to make a easy to maintain reporting tool.

```
source("RFunctions/DataSource.R")

df_positives_raw = get_positive_cases_df()
df_deceases_raw  = get_deceases_cases_df()
df_recovered_raw = get_recovered_cases_df()
```

R's source() function allows remote calling of functions, parameters and from .R and .Rmd

Data Exploration

Each dataset comes from a separate source. Therefore we begin by exploring the structure of each dataset.

Shape and variable type

Positive Cases

```
## row count: 282
## column count: 757
##  -  character :  2
##  -  integer :  753
##  -  numeric :  2
```

Deceased Cases

```
## row count: 282
## column count: 755
##  -  character :  2
##  -  integer :  751
##  -  numeric :  2
```

Recovered Cases

```
## row count: 267
## column count: 755
##  -  character :  2
##  -  integer :  751
##  -  numeric :  2
```

Character and numeric type columns

Since all Datasets share a common property of having 2 character and 2 numeric columns, it is possible that all Datasets have the same 4 columns (fields). We can verify these two columns on each case to verify the hypothesis and, then, start exploring the integer type columns.

```
## [1] "From positive cases dataset"

## columns of type: character
## Province.State
## Country.Region

## [1] "----------------------------"

## [1] "From deceases cases dataset"

## columns of type: character
## Province.State
## Country.Region

## [1] "----------------------------"

## [1] "From recovered cases dataset"

## columns of type: character
## Province.State
## Country.Region
```

Similar scenario can be expected regarding the numeric type variables.

```
## [1] "From positive cases dataset"

## columns of type: numeric
## Lat
## Long

## [1] "----------------------------"

## [1] "From deceases cases dataset"

## columns of type: numeric
## Lat
## Long

## [1] "----------------------------"
```

```
## [1] "From recovered cases dataset"
## columns of type: numeric
## Lat
## Long
```