

WHO data

Nick Lauerman

4/8/2020

Contents

Libraries Used	1
Developed Functions	2
Read data in	3
Set variables	3
Get the data	3
Post process/clean the data	17
data check	18

Libraries Used

```
library(tidyverse)
```

```
## -- Attaching packages -----
## v ggplot2 3.3.0      v purrr  0.3.3
## v tibble  3.0.0      v dplyr  0.8.5
## v tidyr   1.0.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0

## -- Conflicts -----
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:dplyr':
##
##   intersect, setdiff, union

## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union
```

Developed Functions

```
readWHO <- function(path = "../data/WHO/",
                    file = "report 1.csv",
                    date = NULL,
                    reportNumber = 1){
  rawData <- NULL
  require(tidyverse)
  require(lubridate)
  toRead <- paste(path,
                  file,
                  sep = "")
  rawData <- read_csv(file = toRead)
  rawData$report <- reportNumber
  rawData$date <- date
  rawData$New[is.na(rawData$New)] <- FALSE
  return(rawData)
}

WHOClean <- function(dataSet = WHO){
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Federal Democratic Repu
    "Napal"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Republic of Singapore")
    "Singapore"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "the United Kingdom")] <-
    "United Kingdom"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "The United Kingdom")] <-
    "United Kingdom"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "the United States")] <-
    "United States of America"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Saint Barthlemy")] <-
    "Saint Barthelemy"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "occupied Palestinian t
    "Occupied Palestinian Territory"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "occupied Palestinian T
    "Occupied Palestinian Territory"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Cote d'Ivoire")] <-
    "Cote d Ivoire"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Kosovo[1]")] <-
    "Kosovo"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Lao People?s Democrati
    "Lao People's Democratic Republic"
  dataSet$`Country Territory area`[which(dataSet$`Country Territory area` == "Dominican")] <-
    "Dominican Republic"
  dataSet <- dataSet %>%
    filter(Region != "Other")
  return(dataSet)
}
```

Read data in

Set variables

```
endReport <- 78 # 77
WHO <- NULL
startDay <- as.Date("21-Jan-2020", format = "%d-%b-%Y")
```

Get the data

```
for (i in 1:endReport) {
  reportDate <- startDay + days(i - 1)
  fileName <- paste("report ",
                    i,
                    ".csv",
                    sep = "")
  temp <- readWHO(file = fileName,
                  date = reportDate,
                  reportNumber = i)
  WHO <- rbind(WHO,
               temp)
  rm(temp)
}

## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
```

```

## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),

```

```

## `total Confirmed Cases` = col_double(),
## `Confirmed new cases` = col_double(),
## `total deaths` = col_double(),
## `total new deaths` = col_double(),
## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```

```

## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(

```

```

##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),

```

```

## `total deaths` = col_double(),
## `total new deaths` = col_double(),
## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```



```

## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),

```

```

## `total Confirmed Cases` = col_double(),
## `Confirmed new cases` = col_double(),
## `total deaths` = col_double(),
## `total new deaths` = col_double(),
## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```

```

##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(

```

```

##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),

```

```

## `total deaths` = col_double(),
## `total new deaths` = col_double(),
## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```

```

## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),

```

```

## `total Confirmed Cases` = col_double(),
## `Confirmed new cases` = col_double(),
## `total deaths` = col_double(),
## `total new deaths` = col_double(),
## New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```

```

##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(

```



```

##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )
## Parsed with column specification:
## cols(
##   Region = col_character(),
##   `Country Territory area` = col_character(),
##   `total Confirmed Cases` = col_double(),
##   `Confirmed new cases` = col_double(),
##   `total deaths` = col_double(),
##   `total new deaths` = col_double(),
##   New = col_logical()
## )

```

Post process/clean the data

```
WHO <- WHOClean(dataSet = WHO)
```

data check

```
temp <- table(WHO$`Country Territory area`,
              useNA = "ifany")
temp[order(temp,
            decreasing = TRUE)]
```

```
##
##              China
##              78
##              Japan
##              78
##      Republic of Korea
##              78
##              Thailand
##              78
##      United States of America
##              77
##              Singapore
##              75
##              Viet Nam
##              75
##              Australia
##              74
##              Canada
##              73
##              France
##              73
##              Malaysia
##              73
##              Cambodia
##              71
##              Germany
##              71
##              Sri Lanka
##              71
##      United Arab Emirates
##              70
##              Finland
##              69
##              India
##              69
##      Philippines
##              69
##              Italy
##              67
##      Russian Federation
##              67
##              Spain
##              67
##              Sweden
##              67
##      United Kingdom
##              67
```

##	Belgium
##	63
##	Egypt
##	53
##	Lebanon
##	46
##	Kuwait
##	44
##	Bahrain
##	43
##	Iraq
##	43
##	Oman
##	43
##	Algeria
##	42
##	Austria
##	42
##	Brazil
##	42
##	Croatia
##	42
##	Switzerland
##	42
##	Denmark
##	41
##	Estonia
##	41
##	Georgia
##	41
##	Greece
##	41
##	North Macedonia
##	41
##	Norway
##	41
##	Pakistan
##	41
##	Romania
##	41
##	Belarus
##	40
##	Lithuania
##	40
##	Netherlands
##	40
##	New Zealand
##	40
##	Nigeria
##	40
##	San Marino
##	39
##	Azerbaijan
##	38

##	Ecuador
##	38
##	Ireland
##	38
##	Monaco
##	38
##	Napal
##	38
##	Qatar
##	38
##	Armenia
##	37
##	Czechia
##	37
##	Dominican Republic
##	37
##	Iceland
##	37
##	Indonesia
##	37
##	Luxembourg
##	37
##	Afghanistan
##	36
##	Andorra
##	36
##	Iran (Islamic Republic of)
##	36
##	Israel
##	36
##	Jordan
##	36
##	Latvia
##	36
##	Mexico
##	36
##	Morocco
##	36
##	Nepal
##	36
##	Portugal
##	36
##	Saudi Arabia
##	36
##	Senegal
##	36
##	Tunisia
##	36
##	Argentina
##	35
##	Chile
##	35
##	Poland
##	35

##	Ukraine
##	35
##	Bosnia and Herzegovina
##	34
##	Gibraltar
##	34
##	Hungary
##	34
##	Liechtenstein
##	34
##	Occupied Palestinian Territory
##	34
##	Saint Barthelemy
##	34
##	Saint Martin
##	34
##	Slovenia
##	34
##	Bhutan
##	33
##	Cameroon
##	33
##	Serbia
##	33
##	South Africa
##	33
##	Colombia
##	32
##	Holy See
##	32
##	Peru
##	32
##	Slovakia
##	32
##	Bulgaria
##	31
##	Costa Rica
##	31
##	Faroe Islands
##	31
##	French Guiana
##	31
##	Maldives
##	31
##	Malta
##	31
##	Martinique
##	31
##	Republic of Moldova
##	31
##	Togo
##	31
##	Albania
##	30

##	Bangladesh
##	30
##	Paraguay
##	30
##	Brunei Darussalam
##	29
##	Cyprus
##	29
##	Guernsey
##	29
##	Mongolia
##	29
##	Panama
##	29
##	Bolivia (Plurinational State of)
##	28
##	Burkina Faso
##	28
##	Democratic Republic of the Congo
##	28
##	Jamaica
##	28
##	Cote d Ivoire
##	27
##	French Polynesia
##	27
##	Honduras
##	27
##	Turkey
##	27
##	Cuba
##	26
##	Guyana
##	26
##	Jersey
##	26
##	Reunion
##	26
##	Saint Vincent and the Grenadines
##	26
##	Antigua and Barbuda
##	25
##	Cayman Islands
##	25
##	Ethiopia
##	25
##	Gabon
##	25
##	Ghana
##	25
##	Guadeloupe
##	25
##	Guinea
##	25

##	Kenya
##	25
##	Puerto Rico
##	25
##	Sudan
##	25
##	Trinidad and Tobago
##	25
##	Venezuela (Bolivarian Republic of)
##	25
##	Central African Republic
##	24
##	Congo
##	24
##	Curacao
##	24
##	Equatorial Guinea
##	24
##	Eswatini
##	24
##	Kazakhstan
##	24
##	Mauritania
##	24
##	Mayotte
##	24
##	Namibia
##	24
##	Guatemala
##	23
##	Rwanda
##	23
##	Saint Lucia
##	23
##	Seychelles
##	23
##	Suriname
##	23
##	Uruguay
##	23
##	Uzbekistan
##	23
##	Aruba
##	22
##	Bahamas
##	22
##	Benin
##	22
##	Guam
##	22
##	Liberia
##	22
##	Somalia
##	22

##	United Republic of Tanzania	
##		22
##	United States Virgin Islands	
##		22
##	Montenegro	
##		21
##	Barbados	
##		20
##	Djibouti	
##		20
##	Gambia	
##		20
##	Greenland	
##		20
##	Kyrgyzstan	
##		20
##	Mauritius	
##		20
##	Montserrat	
##		20
##	Zambia	
##		20
##	Bermuda	
##		19
##	Chad	
##		19
##	El Salvador	
##		19
##	Fiji	
##		19
##	New Caledonia	
##		19
##	Nicaragua	
##		19
##	Niger	
##		19
##	Sint Maarten	
##		19
##	Cabo Verde	
##		18
##	Haiti	
##		18
##	Isle of Man	
##		18
##	Papua New Guinea	
##		18
##	Timor-Leste	
##		18
##	Zimbabwe	
##		18
##	Angola	
##		17
##	Eritrea	
##		17

##	Madagascar	
##		17
##	Uganda	
##		17
##	Grenada	
##		16
##	Kosovo	
##		16
##	Mozambique	
##		16
##	Syrian Arab Republic	
##		16
##	Belize	
##		15
##	Dominica	
##		15
##	Myanmar	
##		15
##	Turks and Caicos Islands	
##		15
##	Lao People's Democratic Republic	
##		14
##	Libya	
##		14
##	Guinea-Bissau	
##		13
##	Mali	
##		13
##	Saint Kitts and Nevis	
##		13
##	Anguilla	
##		12
##	British Virgin Islands	
##		12
##	Iran	
##		10
##	Isreal	
##		10
##	Northern Mariana Islands (Commonwealth of the)	
##		10
##	Afganistan	
##		7
##	Botswana	
##		7
##	Burundi	
##		7
##	Sierra Leone	
##		7
##	Malawi	
##		5
##	Bonaire, Sint Eustatius and Saba	
##		4
##	Falkland Islands (Malvinas)	
##		3

```
##                                Mexio
##                                3
##                                Irab
##                                2
##                                South Sudan
##                                2
##                                Sao Tome and Principe
##                                1
```

```
temp[temp == 1]
```

```
## Sao Tome and Principe
##                                1
```

```
length(temp)
```

```
## [1] 216
```

```
table(WHO$date,
      useNA = "ifany")
```

```
##
## 2020-01-21 2020-01-22 2020-01-23 2020-01-24 2020-01-25 2020-01-26 2020-01-27
##          4          4          5          7          9          11          12
## 2020-01-28 2020-01-29 2020-01-30 2020-01-31 2020-02-01 2020-02-02 2020-02-03
##          15          16          19          19          24          24          24
## 2020-02-04 2020-02-05 2020-02-06 2020-02-07 2020-02-08 2020-02-09 2020-02-10
##          24          25          25          25          25          25          25
## 2020-02-11 2020-02-12 2020-02-13 2020-02-14 2020-02-15 2020-02-16 2020-02-17
##          25          25          25          25          26          26          26
## 2020-02-18 2020-02-19 2020-02-20 2020-02-21 2020-02-22 2020-02-23 2020-02-24
##          26          26          27          27          29          29          30
## 2020-02-25 2020-02-26 2020-02-27 2020-02-28 2020-02-29 2020-03-01 2020-03-02
##          34          38          47          55          54          59          65
## 2020-03-03 2020-03-04 2020-03-05 2020-03-06 2020-03-07 2020-03-08 2020-03-09
##          73          77          85          89          94          102         105
## 2020-03-10 2020-03-11 2020-03-12 2020-03-13 2020-03-14 2020-03-15 2020-03-16
##          110         114         118         123         135         144         150
## 2020-03-17 2020-03-18 2020-03-19 2020-03-20 2020-03-21 2020-03-22 2020-03-23
##          159         160         168         176         182         186         190
## 2020-03-24 2020-03-25 2020-03-26 2020-03-27 2020-03-28 2020-03-29 2020-03-30
##          194         196         199         201         201         202         202
## 2020-03-31 2020-04-01 2020-04-02 2020-04-03 2020-04-04 2020-04-05 2020-04-06
##          202         205         205         206         207         208         209
## 2020-04-07
##          210
```

```
length(table(WHO$date,
             useNA = "ifany"))
```

```
## [1] 78
```

```
table(WHO$report,
      useNA = "ifany")
```

```
##
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
##  4  4  5  7  9 11 12 15 16 19 19 24 24 24 24 25 25 25 25 25
```

```
## 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
## 25 25 25 25 25 26 26 26 26 26 27 27 29 29 30 34 38 47 55 54
## 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
## 59 65 73 77 85 89 94 102 105 110 114 118 123 135 144 150 159 160 168 176
## 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78
## 182 186 190 194 196 199 201 201 202 202 202 205 205 206 207 208 209 210
```

```
length(table(WHO$report,
             useNA = "ifany"))
```

```
## [1] 78
```

```
temp <- table(WHO$Region,
              useNA = "ifany")
temp[order(temp,
           decreasing = TRUE)]
```

```
##
##           European           Region of the Americas
##           2167             1021
##           African           Eastern Mediterranean
##           952             753
##           Western Pacific Region of the Americas Territories
##           469             371
##           Western Pacific Region           South-East Asia
##           345             307
##           European Territories           South-East Asia Region
##           176             155
##           Western Pacific Territories           African Territories
##           78             50
##           Eastern Mediterranean Territories
##           34
```

```
rm (temp)
```

```
tail(WHO)
```

```
## # A tibble: 6 x 9
##   Region `Country Territ~` `total Confirme~` `Confirmed new ~` `total deaths`
##   <chr> <chr>             <dbl>             <dbl>             <dbl>
## 1 Afric~ Malawi           4                 0                 0
## 2 Afric~ Sao Tome and Pr~  4                 4                 0
## 3 Afric~ Burundi          3                 0                 0
## 4 Afric~ South Sudan       1                 0                 0
## 5 Afric~ Reunion          349               5                 0
## 6 Afric~ Mayotte          164              17                 2
## # ... with 4 more variables: `total new deaths` <dbl>, New <lgl>, report <int>,
## #   date <date>
```