**Name: - Vikas Virani (Student Id: s3715555)**

**RPubs Link: -** <https://rpubs.com/VikasVirani/613003>

**Code:-**

# installed.packages(magrittr)

# installed.packages(dplyr)

# installed.packages(janitor)

# installed.packages(ggrepel)

# installed.packages(directlabels)

library(ggplot2)

library(magrittr) # needs to be run every time you start R and want to use %>%

library(dplyr) # alternatively, this also loads %>%

library(janitor)

library(ggrepel)

library(directlabels)

library(scales)

# library(gganimate)

#loading data directly from web (github)

covid\_data <- read.csv(url("https://raw.githubusercontent.com/owid/covid-19-data/master/public/data/ecdc/COVID-2019%20-%20ECDC%20(2020).csv"))

#covid\_data

dataToContinentMapping <- read.csv(url("https://raw.githubusercontent.com/owid/covid-19-data/master/public/data/ecdc/locations.csv"))

#dataToContinentMapping

#Filter rows which are of not our use(i.e. not countries)

notCountry <- c('Africa', 'Asia', 'Asia excl. China', 'Europe', 'European Union', 'High income', 'Low income', 'Lower middle income', 'North America', 'Oceania', 'South America', 'Upper middle income', 'World', 'World excl. China', 'World excl. China and South Korea', 'World excl. China, South Korea, Japan and Singapore')

#join COVID data with country to continent, population mapping

Total\_covid\_data <- merge(covid\_data, dataToContinentMapping, by.x="Country", by.y="location", all.x = TRUE)

#Total\_covid\_data

#typeof(Total\_covid\_data)

#attributes(Total\_covid\_data)

#str(Total\_covid\_data)

# Total\_covid\_data <- Total\_covid\_data[Total\_covid\_data$"Days.since.the.total.confirmed.deaths.of.COVID.19.reached.5" >= 1,]

#filter & process data to make it ready for visualisation

Total\_covid\_data <- Total\_covid\_data[Total\_covid\_data$"Days.since.the.total.confirmed.deaths.of.COVID.19.per.million.people.reached.0.1" >= 0.1,]

Total\_covid\_data <- Total\_covid\_data[!is.na(Total\_covid\_data$"Country"),]

Total\_covid\_data <- Total\_covid\_data[!Total\_covid\_data$"Country" %in% notCountry, ]

#filtering needed columns/features

Total\_covid\_data\_filtered <- Total\_covid\_data[,c("Days.since.the.total.confirmed.deaths.of.COVID.19.per.million.people.reached.0.1","Total.confirmed.deaths.due.to.COVID.19.per.million.people","Country","continent")]

# Total\_covid\_data\_filtered <- Total\_covid\_data[,c("Days.since.the.total.confirmed.deaths.of.COVID.19.reached.5","Total.confirmed.deaths.due.to.COVID.19","Country","continent")]

#change feature names to make it less complex while working with them

names(Total\_covid\_data\_filtered) <- Total\_covid\_data\_filtered %>%

clean\_names() %>%

names()

names(Total\_covid\_data\_filtered) <- c("Days","Deaths","Country","Continent")

## Part used when processing/investigating data

# make.names(names(Total\_covid\_data\_filtered))

#names(Total\_covid\_data\_filtered)

#str(Total\_covid\_data\_filtered)

# Total\_covid\_data\_filtered <- Total\_covid\_data\_filtered %>% mutate(label = if\_else(Days == max(Days), as.character(Country), NA\_character\_))

# Total\_covid\_data\_filtered

## part ends

Total\_covid\_data\_filtered

#plot final filtered data

p1 <- ggplot(data=Total\_covid\_data\_filtered, aes(x=Days, y=Deaths, group=Country, color=Country)) +

geom\_line(size=1) +

geom\_point(size=2) +

guides(color = FALSE) + #remove legend

scale\_color\_brewer(palette="Dark2") +

facet\_wrap(~ Continent, ncol = 1,scales="free\_x") + #facet by continent

#scale\_x\_discrete(breaks = waiver(), labels = waiver(), expand=c(0, 3)) +

#scale\_x\_continuous() +

#scale\_y\_log10(breaks = trans\_breaks("log10", function(x) 10^x))

scale\_y\_continuous(trans = log10\_trans(),breaks = trans\_breaks("log10", function(x) 10^x),

labels = trans\_format("log10", math\_format(10^.x))) + #change y-scale for increase rate

#facet\_grid(rows = vars(Continent))

#ylim(0, 1000) +

labs(

title = "Total confirmed COVID-19 deaths: how rapidly are they increasing?",

x = "Days since total deaths reached 0.1 per million population of a country",

y = "Number of deaths (Per million population)"

) + # add labels

theme(axis.text.x = element\_text(colour = "black", size = 12, angle = 90, hjust = 0.5, vjust = 0.5),

axis.text.y = element\_text(colour = "red", size = 12), plot.title = element\_text(hjust = 0.5, size = 22, face = "bold.italic"),

text = element\_text(size = 16)) + #legend.position = "bottom"

geom\_dl(aes(label = Country), method = list(dl.combine("last.points"), cex = 0.8)) # add country name to each line

# theme\_classic()

# theme\_minimal()

# geom\_label\_repel(aes(label = label),

# nudge\_x = 1,

# na.rm = TRUE) +

# transition\_reveal(x) +

# anim\_save("COVID-19.gif")

p1