**Covid Analysis**

**Code**

library(ggplot2)

library(scales)

data = read.csv('transformed\_data.csv', fileEncoding="UTF-8-BOM")

data2 = read.csv('raw\_data.csv', fileEncoding="UTF-8-BOM")

head(data)

head(data2)

data

table(data['COUNTRY'])

find\_mode <- function(x) {

u <- unique(x)

tab <- tabulate(match(x, u))

u[tab == max(tab)]

}

mode = find\_mode(table(data['COUNTRY']))

code = unique(data['CODE'])

country = unique(data['COUNTRY'])

hdi = data.frame()

tc = data.frame()

td = data.frame()

sti = data.frame()

population = unique(data['POP'])

gdp = data.frame()

for (j in 1:nrow(country)){

i = country[j, 1]

print(paste0(j," ",i))

hdi[j, 1] = sum(data[which(data$COUNTRY == i), 'HDI'], na.rm = T)/mode

tc[j, 1] = sum(data2[which(data2$location == i), 'total\_cases'], na.rm = T)

td[j, 1] = sum(data2[which(data2$location == i), 'total\_deaths'], na.rm = T)

sti[j, 1] = sum(data[which(data$COUNTRY == i), 'STI'], na.rm = T)/mode

population[j, 1] = sum(data2[which(data2$location == i), 'population'], na.rm = T)/ mode

}

code <- unname(code)

country <- unname(country)

hdi <- unname(hdi)

tc <- unname(tc)

td <- unname(td)

sti <-unname(sti)

population <- unname(population)

aggregated\_data = cbind(code, country, hdi,tc,td,sti,population, row.names = NULL)

aggregated\_data

colnames(aggregated\_data) <- c("Country Code", "Country", "HDI",

"Total Cases", "Total Deaths",

"Stringency Index", "Population")

head(aggregated\_data)

data = aggregated\_data[order(aggregated\_data$`Total Cases`, decreasing = T),]

data = head(data, 10)

data

data$`GDP Before Covid` = c(65279.53, 8897.49, 2100.75,

11497.65, 7027.61, 9946.03,

29564.74, 6001.40, 6424.98, 42354.41)

data$`GDP During Covid` = c(63543.58, 6796.84, 1900.71,

10126.72, 6126.87, 8346.70,

27057.16, 5090.72, 5332.77, 40284.64)

data

plotData <- data

plotData$Country <- factor(plotData$Country, # Factor levels in increasing order

levels = plotData$Country[order(plotData$`Total Cases`, decreasing = T)])

ggplot(plotData, aes(x = `Country`, y = `Total Cases`)) + geom\_bar( stat="identity", fill='sky blue') +

ggtitle("Countries with Highest Covid Cases") + theme(plot.title = element\_text(hjust = 0.5))

ggplot(plotData, aes(x = `Country`, y = `Total Deaths`)) + geom\_bar( stat="identity", fill='sky blue') +

ggtitle("Countries with Highest Deaths") + theme(plot.title = element\_text(hjust = 0.5))

dfm <- pivot\_longer(plotData, c('Total Cases', 'Total Deaths'), names\_to="variable", values\_to="value")

ggplot(dfm, aes(x = `Country`, y = value)) + geom\_bar(aes(fill = variable), stat="identity", position = "dodge") +

ggtitle("Countries with Highest Deaths") + theme(plot.title = element\_text(hjust = 0.5))

cases = sum(data$`Total Cases`)

deceased = sum(data$`Total Deaths`)

labels = c("Total Cases", "Total Deaths")

values = c(cases, deceased)

dfm = as.data.frame(cbind(labels, values))

dfm$values <- as.numeric(dfm$values)

sum\_of\_obsrv <- sum(dfm$values)

ggplot(dfm, aes(x="", y=values, fill=labels)) +

geom\_col() +

geom\_label(aes(label = percent(values/sum\_of\_obsrv)),

position = position\_stack(vjust = 0.5),

show.legend = FALSE) +

coord\_polar(theta = "y") +

theme\_void() +

ggtitle("Percentage of Total Cases and Deaths") + theme(plot.title = element\_text(hjust = 0.5))

death\_rate = (sum(data$`Total Deaths`) / sum(data$`Total Cases`)) \* 100

paste0("Death Rate =", death\_rate)

ggplot(plotData, aes(x = `Country`, y = `Total Cases`, fill=`Stringency Index`)) + geom\_bar( stat="identity") +

ggtitle("Stringency Index during Covid-19") + theme(plot.title = element\_text(hjust = 0.5))

ggplot(plotData, aes(x = `Country`, y = `Total Cases`, fill=`GDP Before Covid`)) + geom\_bar( stat="identity") +

ggtitle("GDP Per Capita Before Covid-19") + theme(plot.title = element\_text(hjust = 0.5))

ggplot(plotData, aes(x = `Country`, y = `Total Cases`, fill=`GDP During Covid`)) + geom\_bar( stat="identity") +

ggtitle("GDP Per Capita During Covid-19") + theme(plot.title = element\_text(hjust = 0.5))

dfm <- pivot\_longer(plotData, c('GDP Before Covid', 'GDP During Covid'), names\_to="variable", values\_to="value")

ggplot(dfm, aes(x = `Country`, y = value)) + geom\_bar(aes(fill = variable), stat="identity", position = "dodge") +

ggtitle("GDP Comparision") + theme(plot.title = element\_text(hjust = 0.5))

**Screenshots**

Table

Description automatically generated

Graphical user interface, text

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

Text

Description automatically generated

Text

Description automatically generated with medium confidence

Text, letter

Description automatically generated

Chart

Description automatically generated