**Global Population Dashboard**

Name: Anurag Dinesh Karmarkar

Dashboard Link: <https://lordreaper29.shinyapps.io/Assignment3/>

Code:

User Interface:

library(shiny)

library(shinydashboard)

shinyUI(

dashboardPage(

dashboardHeader(title = "Global Population"),

dashboardSidebar(

sidebarMenu(

menuItem("Overview", tabName = "tab1", icon = icon("user-friends")),

menuItem("Top 10 Populated Countries", tabName = "tab2", icon = icon("list-ol")),

menuItem("Factors", tabName = "tab3", icon = icon("layer-group"))

)

),

dashboardBody(

tabItems(

#First Tab

tabItem(

tabName = "tab1",

fluidRow(box("This tab shows the overview of the Global Population, its growth and Gender Ratio.")),

fluidRow(

box(plotOutput("plot1")),

box(plotOutput("plot9"))

),

fluidRow(

box(plotOutput("plot2")),

box(plotOutput("plot3"))

)

),

# Second Tab

tabItem(

tabName = "tab2",

fluidRow(box("This tab shows the 10 most populated countries in the World and factors influencing these population numbers.")),

fluidRow(

plotOutput("plot4")

),

fluidRow(

box(plotOutput("plot8")),

box(plotOutput("plot6"))

)

),

# Third Tab

tabItem(

tabName = "tab3",

fluidRow(box("This tab shows the Birth Rate, Death Rate and Fertility Rate of the 10 most populated countries.")),

fluidRow(

plotOutput("plot7", click = "plot\_click"),

verbatimTextOutput("info")

),

fluidRow(

box(plotOutput("plot5")),

verbatimTextOutput("references")

)

)

)

)

)

)

Server:

library(shiny)

library(shinydashboard)

# Importing the necessary libraries

library(ggplot2)

library(readr)

library(tidyr)

library(dplyr)

require(scales)

library(viridis)

library(hrbrthemes)

library(RColorBrewer)

library(readxl)

library(ggpol)

library(plotly)

# Import dataset

population\_data <- read\_csv("population.csv",

col\_types = cols(Code = col\_skip()))

#head(population\_data)

# Source

# https://www.gapminder.org/data/documentation/gd003/

# Renaming Columns

names(population\_data)[names(population\_data) == "Entity"] <- "Region"

names(population\_data)[names(population\_data) == "Year"] <- "year"

names(population\_data)[names(population\_data) == "Total population (Gapminder, HYDE & UN)"] <- "total\_population"

#head(population\_data)

# Storing the data of the world in a seperate data frame

world\_population <- population\_data[population\_data$Region == "World" & population\_data$year >= 1900,]

#head(world\_population)

p1 <- ggplot(world\_population, aes(x =year, y=total\_population)) +

geom\_line() +

xlab("Year") +

ylab("Population") +

ggtitle("Global Population Growth") +

theme(plot.title = element\_text(hjust = 0.5)) +

scale\_y\_continuous(label = unit\_format(unit = "Billion", scale = 1e-9))

#p1

# Storing region vise data in a different plot

region\_population <- population\_data[(population\_data$Region == "Asia" |

population\_data$Region == "Africa" |

population\_data$Region == "Europe" |

population\_data$Region == "Oceania" |

population\_data$Region == "North America" |

population\_data$Region == "Latin America" ) &

population\_data$year >= 1900,]

#head(region\_population)

p2 <- ggplot(region\_population, aes(x=year, y=total\_population, fill=Region)) +

geom\_area() +

geom\_area(alpha=0.6 , size=.5, colour="white") +

scale\_fill\_viridis(discrete = T) +

#theme\_ipsum() +

xlab("Year") +

ylab("Population") +

ggtitle("Population Growth by Region") +

theme(plot.title = element\_text(hjust = 0.5)) +

scale\_y\_continuous(label = unit\_format(unit = "Billion", scale = 1e-9))

#p2 <- ggplotly(p2, tooltip="text")

#p2

#p2

# Importing data

population\_age\_group <- read\_csv("population-by-broad-age-group.csv",

col\_types = cols(Code = col\_skip()))

#head(population\_age\_group)

# Source

# https://population.un.org/wpp2019/Download/Standard/Population/

# Renaming Columns

names(population\_age\_group)[names(population\_age\_group) == "Entity"] <- "entity"

names(population\_age\_group)[names(population\_age\_group) == "Year"] <- "year"

names(population\_age\_group)[names(population\_age\_group) == "Estimates, 1950 - 2020: Total population by broad age group, both sexes combined (thousands) - Population under age 5"] <- "Below 5"

names(population\_age\_group)[names(population\_age\_group) == "Estimates, 1950 - 2020: Total population by broad age group, both sexes combined (thousands) - Population aged 5-14"] <- "5 to 14"

names(population\_age\_group)[names(population\_age\_group) == "Estimates, 1950 - 2020: Total population by broad age group, both sexes combined (thousands) - Population aged 15-24"] <- "15 to 24"

names(population\_age\_group)[names(population\_age\_group) == "Estimates, 1950 - 2020: Total population by broad age group, both sexes combined (thousands) - Population aged 25-64"] <- "25 to 64"

names(population\_age\_group)[names(population\_age\_group) == "Estimates, 1950 - 2020: Total population by broad age group, both sexes combined (thousands) - Population aged 65 or over"] <- "Above 64"

#head(population\_age\_group)

# Storing the data of the world

population\_age\_group <- population\_age\_group[population\_age\_group$entity == "World",]

#head(population\_age\_group)

population\_age\_group\_con <- population\_age\_group %>% gather(key = "Age Group", value = "population", c('Below 5':'Above 64'))

#population\_age\_group\_con

p3 <- ggplot(population\_age\_group\_con, aes(x=year, y=population, fill=`Age Group`)) +

geom\_area() +

geom\_area(alpha=0.6 , size=.5, colour="white") +

scale\_fill\_viridis(discrete = T) +

#theme\_ipsum() +

xlab("Year") +

ylab("Population") +

ggtitle("Population Growth by Age Group") +

theme(plot.title = element\_text(hjust = 0.5)) +

scale\_y\_continuous(label = unit\_format(unit = "Billion", scale = 1e-9))

#p3

# Selecting and storing top 10 countries according to population in a seperate dataframe

population\_data\_top10 <- population\_data[population\_data$year == 2019 &

(population\_data$Region != "World" &

population\_data$Region != "Asia" &

population\_data$Region != "Africa" &

population\_data$Region != "Europe" &

population\_data$Region != "North America" &

population\_data$Region != "Latin America" &

population\_data$Region != "Oceania"),]

population\_data\_top10 <- population\_data\_top10 %>% arrange(desc(population\_data\_top10$total\_population))

population\_data\_top10 <- head(population\_data\_top10,10)

names(population\_data\_top10)[names(population\_data\_top10) == "Region"] <- "Country"

#population\_data\_top10

p4<-ggplot(data=population\_data\_top10, aes(x=Country, y=total\_population, fill=Country)) +

geom\_bar(stat="identity", position = "dodge") +

scale\_fill\_brewer(palette="Paired") +

xlab("Country") +

ylab("Population") +

ggtitle("Top 10 Countries by Population") +

theme(plot.title = element\_text(hjust = 0.5)) +

scale\_y\_continuous(label = unit\_format(unit = "Billion", scale = 1e-9)) +

theme(axis.title.x=element\_blank(),

axis.text.x=element\_blank(),

axis.ticks.x=element\_blank())

#p4

# Importing data

children\_per\_woman <- read\_csv("children-per-woman-UN.csv",

col\_types = cols(Code = col\_skip()))

#head(children\_per\_woman)

# Source

# https://population.un.org/wpp2019/Download/Standard/Interpolated/

names(children\_per\_woman)[names(children\_per\_woman) == "Entity"] <- "Country"

names(children\_per\_woman)[names(children\_per\_woman) == "Year"] <- "year"

names(children\_per\_woman)[names(children\_per\_woman) == "Estimates, 1950 - 2020: Annually interpolated demographic indicators - Total fertility (live births per woman)"] <- "total\_fertility"

#head(children\_per\_woman)

children\_per\_woman\_top10 <- children\_per\_woman[(children\_per\_woman$Country == "China" |

children\_per\_woman$Country == "India" |

children\_per\_woman$Country == "Russia" |

children\_per\_woman$Country == "Indonesia" |

children\_per\_woman$Country == "Brazil" |

children\_per\_woman$Country == "Bangladesh" |

children\_per\_woman$Country == "Pakistan" |

children\_per\_woman$Country == "United States" |

children\_per\_woman$Country == "Mexico" |

children\_per\_woman$Country == "Nigeria"),]

#children\_per\_woman\_top10

cw\_china <- children\_per\_woman\_top10 %>% filter(Country == "China")

cw\_india <- children\_per\_woman\_top10 %>% filter(Country == "India")

cw\_russia <- children\_per\_woman\_top10 %>% filter(Country == "Russia")

cw\_bangladesh <- children\_per\_woman\_top10 %>% filter(Country == "Bangladesh")

cw\_indonesia <- children\_per\_woman\_top10 %>% filter(Country == "Indonesia")

cw\_brazil <- children\_per\_woman\_top10 %>% filter(Country == "Brazil")

cw\_pakistan <- children\_per\_woman\_top10 %>% filter(Country == "Pakistan")

cw\_us <- children\_per\_woman\_top10 %>% filter(Country == "United States")

cw\_mexico <- children\_per\_woman\_top10 %>% filter(Country == "Mexico")

cw\_nigeria <- children\_per\_woman\_top10 %>% filter(Country == "Nigeria")

p5 <- ggplot(cw\_india, aes(x = year), legend = TRUE) +

geom\_line(aes(y = total\_fertility, color = "red"), size = 1) +

geom\_line(aes(y = cw\_china$total\_fertility, color = "blue"), size = 1) +

geom\_line(aes(y = cw\_bangladesh$total\_fertility, color = "green"), size = 1) +

geom\_line(aes(y = cw\_brazil$total\_fertility, color = "yellow"), size = 1) +

geom\_line(aes(y = cw\_indonesia$total\_fertility, color = "khaki4"), size = 1) +

geom\_line(aes(y = cw\_mexico$total\_fertility, color = "hotpink"), size = 1) +

geom\_line(aes(y = cw\_nigeria$total\_fertility, color = "darkgreen"), size = 1) +

geom\_line(aes(y = cw\_pakistan$total\_fertility, color = "coral"), size = 1) +

geom\_line(aes(y = cw\_russia$total\_fertility, color = "cyan"), size = 1) +

geom\_line(aes(y = cw\_us$total\_fertility, color = "darkorange"), size = 1) +

ggtitle("Fertility Rate of the top 10 Countries") +

xlab('Year') +

ylab('Fertility Rate') +

theme\_minimal() +

scale\_x\_continuous(limits=c(1950, 2020)) +

scale\_y\_continuous(label = comma)+

scale\_colour\_identity(guide ="legend", name = "Countries", labels = c("China","Pakistan","Russia","Nigeria","United States", "Bangladesh", "Mexico", "Indonesia", "India", "Brazil")) +

theme(plot.title = element\_text(hjust = 0.5))

#p5

# Importing Child Mortality Data

child\_mortality <- read\_excel("child\_mortality\_unicef.xlsx",

col\_types = c("skip", "text", "skip",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric"))

#head(child\_mortality)

# Source

# https://data.unicef.org/resources/dataset/child-mortality/

child\_mortality\_con <- child\_mortality %>% gather(key = "Year", value = "child\_mortality", c('1990':'2019'))

#child\_mortality\_con

p6 <- ggplot(child\_mortality\_con, aes(x=Country.Name, y=child\_mortality, fill=Country.Name)) +

geom\_boxplot(

color="blue",

fill="steelblue",

alpha=0.2,

# Notch?

notch=TRUE,

notchwidth = 0.8,

# custom outliers

outlier.colour="red",

outlier.fill="red",

outlier.size=3

) +

coord\_flip() +

xlab("Country") +

ylab("Child Mortality per 1000 live births") +

ggtitle("Child Mortality Rate of the top 10 Countries (1990-2019)") +

theme(plot.title = element\_text(hjust = 0.5))

#p6

# Importing the Birth Rate and Death Rate Data

birth\_rate <- read\_excel("birth\_rate.xlsx",

col\_types = c("text", "skip", "skip",

"skip", "numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric"))

#head(birth\_rate)

death\_rate <- read\_excel("death\_rate.xlsx",

col\_types = c("text", "skip", "skip",

"skip", "numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric"))

#head(death\_rate)

birth\_death <- read\_excel("birth\_death.xlsx")

#head(birth\_death)

# Source

# https://data.worldbank.org/indicator/SP.DYN.CBRT.IN

# https://data.worldbank.org/indicator/SP.DYN.CDRT.IN

# Merging the birth and death rates with population

birth\_death <- merge(birth\_death,population\_data\_top10)

#birth\_death

p7 <- birth\_death %>%

arrange(desc(total\_population)) %>%

mutate(Country = factor(Country, Country)) %>%

ggplot(aes(x=birth\_death$birthrate, y=birth\_death$deathrate, size = birth\_death$total\_population, color=Country)) +

geom\_point(alpha=0.5) +

scale\_size(range = c(.1, 36), name="Population") +

scale\_fill\_viridis(discrete=TRUE, guide=FALSE, option="A") +

ylab("Death Rate") +

xlab("Birth Rate") +

ggtitle("Death Rate versus Birth Rate of the top 10 Countries") +

theme(plot.title = element\_text(hjust = 0.5))

#p7

# Importing Life Expectancy data

life\_expectancy <- read\_excel("life\_expectancy.xlsx",

col\_types = c("text", "skip", "skip",

"skip", "numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric",

"numeric", "numeric", "numeric"))

#head(life\_expectancy)

# Source

# https://data.worldbank.org/indicator/SP.DYN.LE00.IN

life\_expectancy\_con <- life\_expectancy %>% gather(key = "Year", value = "life\_expectancy", c('1960':'2019'))

life\_expectancy\_con$Year <- as.numeric(life\_expectancy\_con$Year)

#life\_expectancy\_con

le\_china <- life\_expectancy\_con %>% filter(Country == "China")

le\_india <- life\_expectancy\_con %>% filter(Country == "India")

le\_russia <- life\_expectancy\_con %>% filter(Country == "Russia")

le\_bangladesh <- life\_expectancy\_con %>% filter(Country == "Bangladesh")

le\_indonesia <- life\_expectancy\_con %>% filter(Country == "Indonesia")

le\_brazil <- life\_expectancy\_con %>% filter(Country == "Brazil")

le\_pakistan <- life\_expectancy\_con %>% filter(Country == "Pakistan")

le\_us <- life\_expectancy\_con %>% filter(Country == "United States")

le\_mexico <- life\_expectancy\_con %>% filter(Country == "Mexico")

le\_nigeria <- life\_expectancy\_con %>% filter(Country == "Nigeria")

p8 <- ggplot(le\_india, aes(x = Year), legend = TRUE) +

geom\_line(aes(y = life\_expectancy, color = "red"), size = 1) +

geom\_line(aes(y = le\_china$life\_expectancy, color = "blue"), size = 1) +

geom\_line(aes(y = le\_bangladesh$life\_expectancy, color = "green"), size = 1) +

geom\_line(aes(y = le\_brazil$life\_expectancy, color = "yellow"), size = 1) +

geom\_line(aes(y = le\_indonesia$life\_expectancy, color = "khaki4"), size = 1) +

geom\_line(aes(y = le\_mexico$life\_expectancy, color = "hotpink"), size = 1) +

geom\_line(aes(y = le\_nigeria$life\_expectancy, color = "darkgreen"), size = 1) +

geom\_line(aes(y = le\_pakistan$life\_expectancy, color = "coral"), size = 1) +

geom\_line(aes(y = le\_russia$life\_expectancy, color = "cyan"), size = 1) +

geom\_line(aes(y = le\_us$life\_expectancy, color = "darkorange"), size = 1) +

ggtitle("Life Expectancy of the top 10 Countries") +

xlab('Year') +

ylab('Life Expectancy') +

theme\_minimal() +

scale\_x\_continuous(limits=c(1960, 2019)) +

#scale\_y\_continuous(label = comma)+

scale\_colour\_identity(guide ="legend", name = "Countries", labels = c("China","Pakistan","Russia","Nigeria","United States", "Bangladesh", "Mexico", "Indonesia", "India", "Brazil")) +

theme(plot.title = element\_text(hjust = 0.5))

#p8

# Importing gender ratio data

gender\_ratio <- read\_excel("gender\_ratio.xlsx",

col\_types = c("skip", "skip", "text",

"numeric", "numeric"))

#head(gender\_ratio)

# Source

# https://population.un.org/

gender\_ratio\_con <- gender\_ratio %>% gather(key = "Gender", value = "Population", c('Female':'Male'))

gender\_ratio\_con$Time <- as.numeric(gender\_ratio\_con$Time)

#gender\_ratio\_con

p9 <- ggplot(gender\_ratio\_con, aes(fill=Gender, y=Population, x=Time)) +

geom\_bar(position="dodge", stat="identity") +

scale\_y\_continuous(label = unit\_format(unit = "Billion", scale = 1e-6)) +

ggtitle("Global Gender Ratio") +

theme(plot.title = element\_text(hjust = 0.5))

#p9

shinyServer(function(input,output){

output$plot1 <- renderPlot(p1)

output$plot2 <- renderPlot(p2)

output$plot3 <- renderPlot(p3)

output$plot4 <- renderPlot(p4)

output$plot5 <- renderPlot(p5)

output$plot6 <- renderPlot(p6)

output$plot7 <- renderPlot(p7)

output$info <- renderText({

paste0("Birth Rate = ", input$plot\_click$x, "\nDeath Rate = ", input$plot\_click$y)

})

output$plot8 <- renderPlot(p8)

output$plot9 <- renderPlot(p9)

output$references <- renderText({

paste("References:-",

"1. Gapminder.org. 2020. GD003 | Gapminder. [online] Available at: <https://www.gapminder.org/data/documentation/gd003/> [Accessed 6 June 2021].",

"2. Population.un.org. 2021. World Population Prospects - Population Division - United Nations. [online] Available at: <https://population.un.org/wpp2019/Download/Standard/Interpolated/> [Accessed 6 June 2021].",

"3. UNICEF Data. 2020. Child mortality data - UNICEF DATA. [online] Available at: <https://data.unicef.org/resources/dataset/child-mortality/> [Accessed 6 June 2021].",

"4. Data.worldbank.org. 2021. Birth rate, crude (per 1,000 people) . [online] Available at: <https://data.worldbank.org/indicator/SP.DYN.CBRT.IN> [Accessed 6 June 2021].",

"5. Data.worldbank.org. 2021. Death rate, crude (per 1,000 people). [online] Available at: <https://data.worldbank.org/indicator/SP.DYN.CDRT.IN> [Accessed 6 June 2021].",

"6. Data.worldbank.org. 2021. Life expectancy at birth, total (years). [online] Available at: <https://data.worldbank.org/indicator/SP.DYN.LE00.IN> [Accessed 6 June 2021].",

sep="\n")

})

})