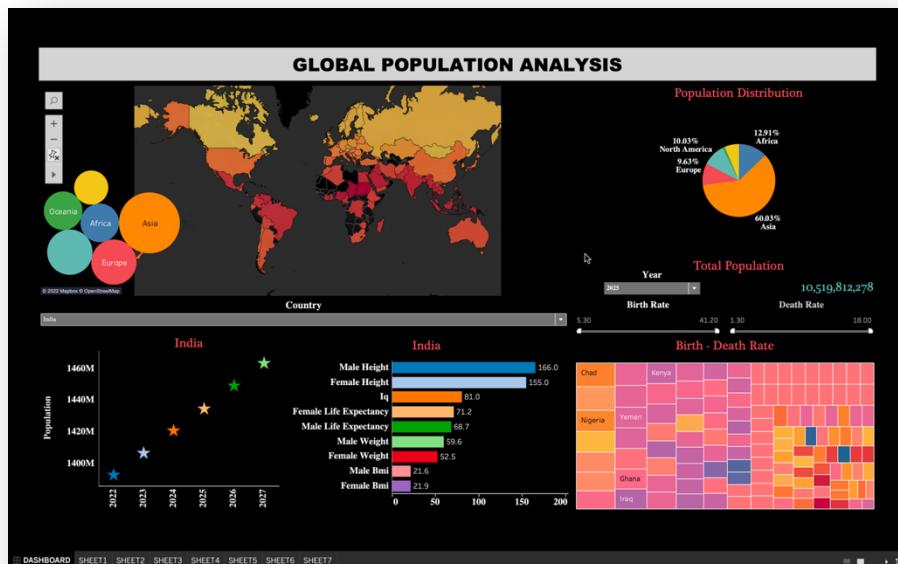


LAB EVALUATION - 2

TABLEAU DASHBOARD

GLOBAL POPULATION ANALYSIS



DASHBOARD LINK:

https://public.tableau.com/views/Dashboard_16689535501990/DASHBOARD?:language=en-GB&publish=yes&:display_count=n&:origin=viz_share_link

EFFORTS BY:

NAME: ROHAN GROVER

ROLL NO: 102003029

GROUP: 3CO2

Datasets:

Reading of Tables

```
1 library(dplyr)
2 library(writexl)
3
4 #reading data from datasets
5 continents <- read.csv('datasets/Countries-Continents.csv')
6 heights <- read.csv('datasets/height_weight_data.csv')
7 iq <- read.csv('datasets/iq.csv')
8 life_exp <- read.csv('datasets/life_expectancy.csv')
9 pop_dens <- read.csv('datasets/population_density.csv')
10 pop_change <- read.csv('datasets/pop_change.csv')
```

Table 1

	continent	country
1	Africa	Algeria
2	Africa	Angola
3	Africa	Benin
4	Africa	Botswana
5	Africa	Burkina
6	Africa	Burundi
7	Africa	Cameroon
8	Africa	Cape Verde
9	Africa	Central African Republic
10	Africa	Chad
11	Africa	Comoros
12	Africa	Congo
13	Africa	Congo (Dem. Republic)
14	Africa	Djibouti
15	Africa	Egypt
16	Africa	Equatorial Guinea
17	Africa	Eritrea
18	Africa	Ethiopia
19	Africa	Gabon
20	Africa	Gambia

Showing 1 to 20 of 194 entries, 2 total columns

Table 2

The screenshot shows a Jupyter Notebook interface with a cell containing a Pandas DataFrame. The DataFrame has columns: country, male_height, female_height, male_weight, female_weight, and male_bmi. The data consists of 20 rows, indexed from 1 to 20, representing various countries. The 'country' column lists countries like Netherlands, Montenegro, Estonia, Denmark, Bosnia and Herzegovina, Iceland, Czechia, Slovenia, Slovakia, Croatia, Serbia, Sweden, Norway, Lithuania, Poland, Ukraine, Finland, Latvia, Germany, and Dominica. The 'male_height' column contains values such as 184, 183, 182, etc. The 'female_height' column contains values like 170, 170, NA, 169, etc. The 'male_weight' column contains values like 87.9, 90.4, 89.9, 86.8, etc. The 'female_weight' column contains values like 73.2, 75.3, 73.7, 70.2, etc. The 'male_bmi' column contains values like 26.1, 27.0, 27.0, 26.3, etc. A footer message at the bottom of the cell says "Showing 1 to 20 of 126 entries, 6 total columns".

	country	male_height	female_height	male_weight	female_weight	male_bmi
1	Netherlands	184	170	87.9	73.2	26.1
2	Montenegro	183	170	90.4	75.3	27.0
3	Estonia	182	NA	89.9	73.7	27.0
4	Denmark	182	169	86.8	70.2	26.3
5	Bosnia and Herzegovina	182	167	87.1	70.6	26.4
6	Iceland	181	168	89.2	72.6	NA
7	Czechia	181	168	91.9	74.2	28.1
8	Slovenia	181	167	NA	74.2	26.6
9	Slovakia	NA	167	89.5	71.3	27.4
10	Croatia	181	167	91.3	74.7	28.0
11	Serbia	180	168	86.9	72.6	26.7
12	Sweden	180	167	87.1	NA	26.8
13	Norway	180	NA	89.1	72.6	27.4
14	Lithuania	180	167	88.1	73.1	27.1
15	Poland	180	165	89.0	71.5	27.4
16	Ukraine	180	166	87.6	72.9	NA
17	Finland	180	166	86.3	71.0	26.6
18	Latvia	NA	168	87.9	74.5	27.1
19	Germany	180	166	88.7	71.7	27.4
20	Dominica	180	167	80.7	80.8	25.0

Table 3

The screenshot shows a Jupyter Notebook interface with a cell containing a Pandas DataFrame. The DataFrame has columns: country, iq, education_expenditure_per_inhabitant, and daily_max_temp. The data consists of 20 rows, indexed from 1 to 20, representing various countries. The 'country' column lists countries like Singapore, Hong Kong, Taiwan, South Korea, Japan, China, Switzerland, Netherlands, North Korea, Macao, Iceland, Finland, Canada, Belgium, Germany, United Kingdom, Austria, New Zealand, Norway, and Sweden. The 'iq' column contains values like 108, 108, 106, 106, 105, 104, 102, 102, 102, 101, 101, 101, 101, 100, 100, 100, 100, 99, 99. The 'education_expenditure_per_inhabitant' column contains values like 908, 915, 537, 1,258, 27, 2,539, 1,701, 1,701, 628, 2,443, 1,990, 1,661, 1,676, 1,385, 1,455, 1,764, 1,283, 3,527, 2,359. The 'daily_max_temp' column contains values like 31.5, 26.2, 26.9, 18.2, 19.2, 19.0, 15.1, 14.4, 15.2, 26.0, 8.1, 8.2, 7.3, 14.8, 13.7, 12.9, 13.1, 17.3, 9.0, 9.9. A footer message at the bottom of the cell says "Showing 1 to 20 of 108 entries, 4 total columns".

	country	iq	education_expenditure_per_inhabitant	daily_max_temp
1	Singapore	108	908	31.5
2	Hong Kong	108	915	26.2
3	Taiwan	106		26.9
4	South Korea	106	537	18.2
5	Japan	105	1,258	19.2
6	China	104	27	19.0
7	Switzerland	102	2,539	15.1
8	Netherlands	102	1,701	14.4
9	North Korea	102		15.2
10	Macao	101	628	26.0
11	Iceland	101	2,443	8.1
12	Finland	101	1,990	8.2
13	Canada	101	1,661	7.3
14	Belgium	100	1,676	14.8
15	Germany	100	1,385	13.7
16	United Kingdom	100	1,455	12.9
17	Austria	100	1,764	13.1
18	New Zealand	100	1,283	17.3
19	Norway	99	3,527	9.0
20	Sweden	99	2,359	9.9

Table 4

A screenshot of a data viewer window titled "Queries.R". The window shows a table with 20 rows and 5 columns. The columns are: country, male_life_expectancy, female_life_expectancy, birth_rate, and death_rate. The rows are numbered 1 to 20. The data includes Hong Kong, Iceland, Japan, Norway, Singapore, Macao, Australia, Switzerland, Malta, Israel, Sweden, South Korea, Ireland, New Zealand, Italy, Netherlands, Canada, Spain, Denmark, and Luxembourg. The last row indicates "Showing 1 to 20 of 115 entries, 5 total columns".

	country	male_life_expectancy	female_life_expectancy	birth_rate	death_rate
1	Hong Kong	82.9	88.0	5.8	6.8
2	Iceland	81.7	84.5	12.3	7.8
3	Japan	81.6	87.7	6.8	8.8
4	Norway	81.6	84.9	9.8	9.8
5	Singapore	81.5	86.1	8.5	10.8
6	Macao	81.4	87.3	10.7	4.1
7	Australia	81.2	85.3	11.5	6.3
8	Switzerland	81.1	85.2	9.9	8.8
9	Malta	80.8	84.6	8.6	7.9
10	Israel	80.7	84.8	19.2	5.3
11	Sweden	80.7	84.2	10.9	9.5
12	South Korea	80.5	86.5	5.3	5.9
13	Ireland	80.4	84.1	11.2	6.4
14	New Zealand	80.3	83.9	11.3	6.4
15	Italy	80.1	84.7	6.8	12.6
16	Netherlands	79.8	83.1	9.7	9.7
17	Canada	79.7	83.9	9.4	8.1
18	Spain	79.7	85.1	7.1	10.4
19	Denmark	79.6	83.6	10.4	9.4
20	Luxembourg	79.4	84.2	10.2	7.3

Table 5

A screenshot of a data viewer window titled "Queries.R". The window shows a table with 20 rows and 4 columns. The columns are: country, area, population, and pop_per_km_sq. The rows are numbered 1 to 20. The data includes Macao, Monaco, Singapore, Hong Kong, Gibraltar, Holy See (Vatican City), Bahrain, Maldives, Malta, Saint Martin, Bermuda, Bangladesh, Guernsey, Jersey, Palestine, St. Martin, Mayotte, Barbados, Taiwan, and Lebanon. The last row indicates "Showing 1 to 20 of 91 entries, 4 total columns".

	country	area	population	pop_per_km_sq
1	Macao	30.0	0.66 M	21657.6
2	Monaco	2.0	0.04 M	19564.4
3	Singapore	719.0	5.45 M	7584.9
4	Hong Kong	1110.0	7.41 M	6678.5
5	Gibraltar	6.8	0.03 M	4954.6
6	Holy See (Vatican City)	0.4	0.00 M	2272.7
7	Bahrain	778.0	1.75 M	2247.2
8	Maldives	300.0	0.54 M	1812.1
9	Malta	320.0	0.52 M	1615.2
10	Saint Martin	34.0	0.04 M	1260.2
11	Bermuda	53.0	0.06 M	1200.5
12	Bangladesh	147630.0	166.30 M	1126.5
13	Guernsey	78.0	0.07 M	865.3
14	Jersey	120.0	0.10 M	854.1
15	Palestine	6020.0	4.92 M	817.7
16	St. Martin	53.0	0.04 M	737.6
17	Mayotte	374.0	0.26 M	685.9
18	Barbados	430.0	0.29 M	669.1
19	Taiwan	35980.0	23.58 M	655.4
20	Lebanon	10450.0	6.77 M	647.8

Table 6

The screenshot shows a data grid in RStudio's environment pane. The grid has two columns: 'country' and 'population_change'. The data is as follows:

	country	population_change
1	China	0.39%
2	India	0.99%
3	United States	0.59%
4	Indonesia	1.07%
5	Pakistan	2.00%
6	Brazil	0.72%
7	Nigeria	2.58%
8	Bangladesh	1.01%
9	Russia	0.04%
10	Mexico	1.06%
11	Japan	-0.30%
12	Ethiopia	2.57%
13	Philippines	1.35%
14	Egypt	1.94%
15	Vietnam	0.91%
16	DR Congo	3.19%
17	Turkey	1.09%
18	Iran	1.30%
19	Germany	0.32%
20	Thailand	0.25%

Showing 1 to 20 of 236 entries, 2 total columns

Joining Unprocessed Tables:

```
-- #joining the datasets on basis of country
12 df1 <- merge(x=heights,y=pop_dens,by="country",all.x=TRUE)
13 df2 <- merge(x=df1,y=iq,by="country",all.x=TRUE)
14 df3 <- merge(x=df2,y=life_exp,by="country",all.x=TRUE)
15 df4 <- merge(x=df3,y=continents,by="country",all.x=TRUE)
16 df <- merge(x=df4,y=pop_change,by="country",all.x=TRUE)
17
18
```

Unprocessed merged Table:

The screenshot shows a data frame titled 'df' in the RStudio environment. The table has 126 rows and 18 columns. The columns are: country, male_height, female_height, male_weight, female_weight, male_bmi, area, population, pop_per_km_sq, iq, education_expenditure_per_inhabitant, and daily. The data includes various countries like Afghanistan, Albania, Algeria, etc., with their respective demographic and economic metrics. The table is sorted by 'country'. A status bar at the bottom indicates 'Showing 1 to 19 of 126 entries, 18 total columns'.

country	male_height	female_height	male_weight	female_weight	male_bmi	area	population	pop_per_km_sq	iq	education_expenditure_per_inhabitant	daily
1 Afghanistan	168	155	63.6	NA	22.6	652860	39.84 M	61.0	80	13	
2 Albania	174	162	81.4	69.2	27.0	NA	NA	NA	84	55	
3 Algeria	174	162	74.6	69.4	24.7	NA	NA	NA	82	109	
4 American Samoa	177	167	103.2	NA	33.1	NA	NA	NA	NA	NA	
5 Andorra	178	165	87.3	71.7	NA	NA	NA	NA	NA	NA	
6 Antigua and Barbuda	178	165	81.6	75.9	25.7	NA	NA	NA	NA	NA	
7 Argentina	NA	161	84.7	71.4	27.9	2780400	45.81 M	16.5	90	278	
8 Australia	179	165	88.3	NA	27.7	7741220	25.74 M	3.3	99	1,390	
9 Austria	178	166	84.6	68.3	26.6	NA	NA	NA	100	1,764	
10 Bahrain	172	158	74.0	62.2	24.9	778	1.75 M	2247.2	83	401	
11 Bangladesh	165	152	57.7	50.5	21.3	147630	166.30 M	1126.5	77	9	
12 Belarus	178	166	84.1	74.4	26.5	NA	NA	NA	NA	NA	
13 Belgium	179	NA	85.9	68.8	26.8	30530	11.59 M	379.6	100	1,676	
14 Bermuda	179	166	88.4	80.4	27.6	NA	NA	NA	88	1,552	
15 Bolivia	168	155	71.2	66.8	25.3	NA	NA	NA	85	67	
16 Bosnia and Herzegovina	182	167	87.1	70.6	26.4	NA	NA	NA	NA	NA	
17 Brazil	175	162	80.7	70.3	26.3	8515770	213.99 M	25.1	83	223	
18 Brunei	166	155	74.7	65.0	27.1	5770	0.44 M	76.5	88	840	
19 Bulgaria	NA	164	81.8	69.6	27.2	NA	NA	NA	91	114	
20 Burma	166	154	NA	54.7	22.2	NA	NA	NA	83	3	

Data Pre-Processing:

```
19 #listing the countries whose continent data is not available
20 df$country[is.na(df$continent)]
21
22 #manually adding information of countries whose continent data is missing
23 df$continent[df$country=="Czechia"]="Europe"
24 df$continent[df$country=="Bermuda"]="North America"
25 df$continent[df$country=="Hong Kong"]="Asia"
26 df$continent[df$country=="Russia"]="Asia"
27 df$continent[df$country=="American Samoa"]="Oceania"
28 df$continent[df$country=="Cook Islands"]="Oceania"
29 df$continent[df$country=="French Polynesia"]="Oceania"
30 df$continent[df$country=="Taiwan"]="Asia"
31 df$continent[df$country=="Niue"]="Oceania"
32 df$continent[df$country=="Puerto Rico"]="North America"
33 df$continent[df$country=="Tokelau"]="Oceania"
34
35 #removing the unit information from population
36 #For example the population data of India before preprocessing is "1393.41 M"
37 #but after running the query it will convert to "1393.41"
38 df$population <- sapply(df$population, function(x) substring(x, 1, nchar(x) - 2))
39
40 #converting population data type from string to double format
41 df$population <- as.double(df$population)
42
```

```

48 #grouping countries by continent
49 df_cont <- group_by(df,continent)
50
51 #replacing the NA values of male height with mean value of male heights w.r.t continent
52 df_male_height <- summarise(df_cont,height = mean(male_height, na.rm = TRUE))
53 for(x in df_male_height$continent)
54 {
55   a <- is.na(df$male_height)
56   b <- df$continent==x
57   df$male_height[a&b] = trunc(df_male_height$height[df_male_height$continent==x])
58 }
59
60 #replacing the NA values of female height with mean value of female heights w.r.t continent
61 df_female_height<-summarise(df_cont,height=mean(female_height, na.rm = TRUE))
62 for(x in df_female_height$continent)
63 {
64   a<-is.na(df$female_height)
65   b<-df$continent==x
66   df$female_height[a&b]=trunc(df_female_height$height[df_female_height$continent==x])
67 }

```

```

69 #replacing the NA values of male weight with mean value of male weights w.r.t continent
70 df_male_weight<-summarise(df_cont,weight=mean(male_weight, na.rm = TRUE))
71 for(x in df_male_weight$continent)
72 {
73   a<-is.na(df$male_weight)
74   b<-df$continent==x
75   df$male_weight[a&b]=round(df_male_weight$weight[df_male_weight$continent==x],1)
76 }
77
78 #replacing the NA values of female weight with mean value of female weights w.r.t continent
79 df_female_weight<-summarise(df_cont,weight=mean(female_weight, na.rm = TRUE))
80 for(x in df_female_weight$continent)
81 {
82   a<-is.na(df$female_weight)
83   b<-df$continent==x
84   df$female_weight[a&b]=round(df_female_weight$weight[df_female_weight$continent==x],1)
85 }
86
87 #adding female BMI information by using the formula: weight/(height)^2
88 df <- mutate(df,female_bmi=round(female_weight*100*100/(female_height*female_height),1))
89
90 #replacing the NA values of male BMI with the the value calculated using the formula: weight/(height)^2
91 df$male_bmi <- round(mapply(function(x,y,z) if(is.na(x)){y*y*100*100/(z*z)} else {x},df$male_bmi,df$male_weight,df$male_height),1)
92
93 #replacing the NA values of IQ with its mean value across the world
94 df$iq[is.na(df$iq)] <- round(mean(df$iq, na.rm=TRUE), 1)
95

```

```

96 #converting education expenditure from string to numeric format
97 df$education_expenditure_per_inhabitant <- as.numeric(df$education_expenditure_per_inhabitant)
98
99 #replacing the NA values of education expenditure with its median value across the world
100 df$education_expenditure_per_inhabitant[is.na(df$education_expenditure_per_inhabitant)] <-
101   round(median(df$education_expenditure_per_inhabitant, na.rm=TRUE),0)
102
103 #replacing the NA values of max temperature with mean value of max temperature w.r.t continent
104 df_temp<-summarise(df_cont,temp=mean(daily_max_temp, na.rm = TRUE))
105 for(x in df_temp$continent)
106 {
107   a<-is.na(df$daily_max_temp)
108   b<-df$continent==x
109   df$daily_max_temp[a&b]=round(df_temp$temp[df_temp$continent==x],1)
110 }
111
112 #replacing the NA values of male life expectancy with its median value across the world
113 df$male_life_expectancy[is.na(df$male_life_expectancy)] <- round(median(df$male_life_expectancy, na.rm=TRUE),1)
114
115 #replacing the NA values of female life expectancy with its median value across the world
116 df$female_life_expectancy[is.na(df$female_life_expectancy)] <- round(median(df$female_life_expectancy, na.rm=TRUE),1)
117
118 #replacing the NA values of birth rate with its median value across the world
119 df$birth_rate[is.na(df$birth_rate)] <- round(median(df$birth_rate, na.rm=TRUE),1)
120
121 #replacing the NA values of death rate with its median value across the world
122 df$death_rate[is.na(df$death_rate)] <- round(median(df$death_rate, na.rm=TRUE),1)

```

```

124 #replacing the NA values of area of a country with its mean of total area covered by countries across the world
125 df$area[is.na(df$area)] <- round(mean(df$area, na.rm=TRUE),1)
126
127 #replacing the NA values of a country's population with mean of country population w.r.t the continent in which it is situated
128 df_pop<-summarise(df_cont,pop=mean(population, na.rm = TRUE))
129 for(x in df_pop$continent)
130 ~ {
131   a<-is.na(df$population)
132   b<-df$continent==x
133   df$population[a&b]=round(df_pop$pop[df_pop$continent==x],1)
134 ~ }
135
136 #replacing the missing values of population per sq km with the the value calculated using country population and its area
137 df$pop_per_km_sq <- round(mapply(function(x,y,z) if(is.na(x)){y*1000000/z} else {x},df$pop_per_km_sq,df$population,df$area),1)
138
139 #reordering the columns
140 df <- df[,c(1,17,7:9,18,12,2:6,19,13:16,10)]
141

```

Processed Merged Table:

Queries.R* df

	country	continent	area	population	pop_per_km_sq	population_change	daily_max_temp	male_height	female_height	male_weight	female_weight	male_bmi
1	Afghanistan	Asia	652860	39.84	61.0	2.33%	24.4	168	155	63.6	62.6	
2	Albania	Europe	1388170	25.90	18.7	-0.11%	21.2	174	162	81.4	69.2	
3	Algeria	Africa	1388170	69.40	50.0	1.85%	26.5	174	162	74.6	69.4	
4	American Samoa	Oceania	1388170	7.70	5.5	-0.22%	20.9	177	167	103.2	85.1	
5	Andorra	Europe	1388170	25.90	18.7	0.16%	15.5	178	165	87.3	71.7	
6	Antigua and Barbuda	North America	1388170	85.40	61.5	0.84%	21.8	178	165	81.6	75.9	
7	Argentina	South America	2780400	45.81	16.5	0.93%	23.9	171	161	84.7	71.4	
8	Australia	Oceania	7741220	25.74	3.3	1.18%	24.4	179	165	88.3	85.1	
9	Austria	Europe	1388170	25.90	18.7	0.57%	13.1	178	166	84.6	68.3	
10	Bahrain	Asia	778	1.75	2247.2	3.68%	32.6	172	158	74.0	62.2	
11	Bangladesh	Asia	147630	166.30	1126.5	1.01%	29.8	165	152	57.7	50.5	
12	Belarus	Europe	1388170	25.90	18.7	-0.03%	15.5	178	166	84.1	74.4	
13	Belgium	Europe	30530	11.59	379.6	0.44%	14.8	179	165	85.9	68.8	
14	Bermuda	North America	1388170	85.40	61.5	-0.36%	24.6	179	166	88.4	80.4	
15	Bolivia	South America	1388170	92.60	66.7	1.39%	27.0	168	155	71.2	66.8	
16	Bosnia and Herzegovina	Europe	1388170	25.90	18.7	-0.61%	15.5	182	167	87.1	70.6	
17	Brazil	South America	8515770	213.99	25.1	0.72%	30.7	175	162	80.7	70.3	
18	Brunei	Asia	5770	0.44	76.5	0.97%	31.7	166	155	74.7	65.0	
19	Bulgaria	Europe	1388170	25.90	18.7	-0.74%	18.1	179	164	81.8	69.6	
20	Burma	Asia	1388170	170.30	122.7	0.67%	32.2	166	154	71.5	54.7	

Showin 1 to 19 of 126 entries, 18 total columns

Tableau implementation:

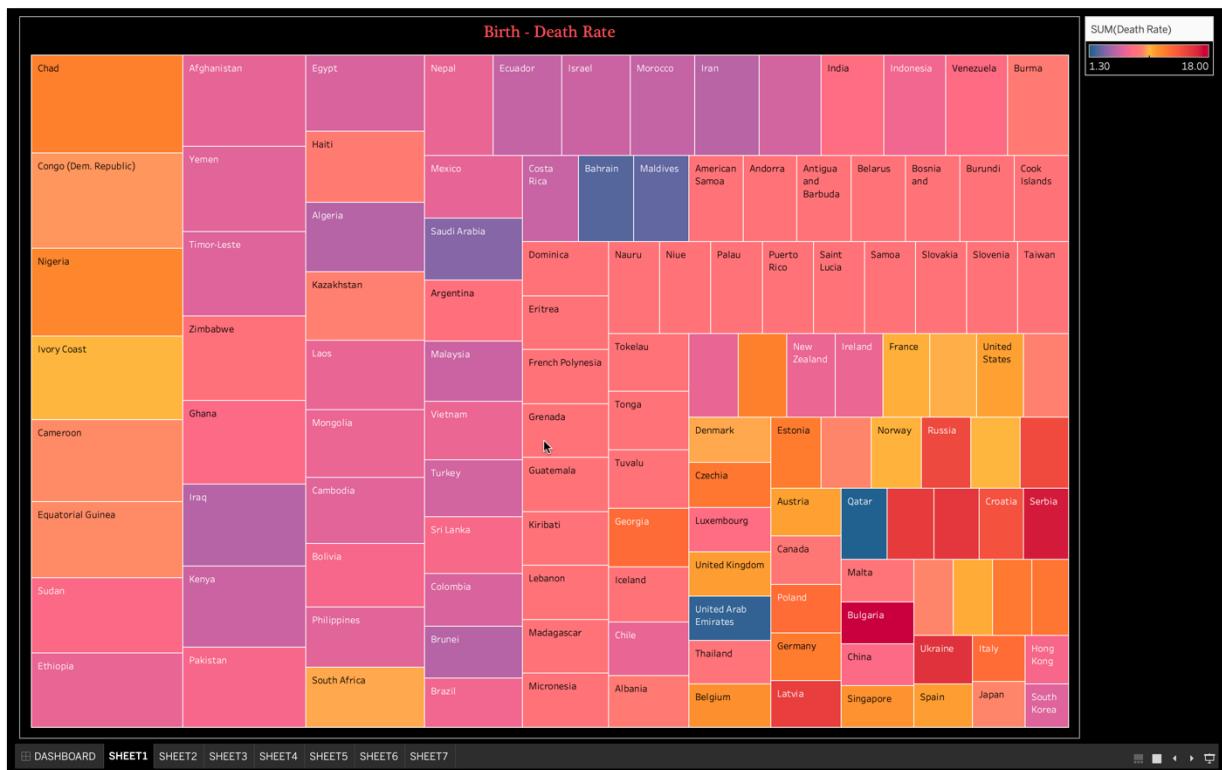
SHEET 1:

Visualizing Birth rate and death rate of each country using highlight table.

Generating Dataset for sheet 1:

```
142 sheet1 <- df[,c('continent', 'country', 'birth_rate', 'death_rate')]  
143 write_xlsx(sheet1,"sheet1.xlsx")  
144
```

	continent	country	birth_rate	death_rate
1	Asia	Afghanistan	31.2	6.2
2	Europe	Albania	11.5	8.3
3	Africa	Algeria	22.8	4.7
4	Oceania	American Samoa	13.0	7.8
5	Europe	Andorra	13.0	7.8
6	North America	Antigua and Barbuda	13.0	7.8
7	South America	Argentina	16.6	7.6
8	Oceania	Australia	11.5	6.3
9	Europe	Austria	9.4	10.3
10	Asia	Bahrain	13.3	2.4
11	Asia	Bangladesh	17.5	5.5
12	Europe	Belarus	13.0	7.8
13	Europe	Belgium	9.9	11.0
14	North America	Bermuda	8.4	8.8
15	South America	Bolivia	21.2	6.8
16	Europe	Bosnia and Herzegovina	13.0	7.8
17	South America	Brazil	13.5	6.6
18	Asia	Brunei	14.1	4.7
19	Europe	Bulgaria	8.5	18.0
20	Asia	Burma	17.2	8.3



Color shows death rate and size of each block shows birth rate i.e., darker color shows higher death rate whereas larger size of block depicts higher birth rate.

SHEET 2:

Visualizing population distribution using pie chart.

Generating Dataset for sheet 2:

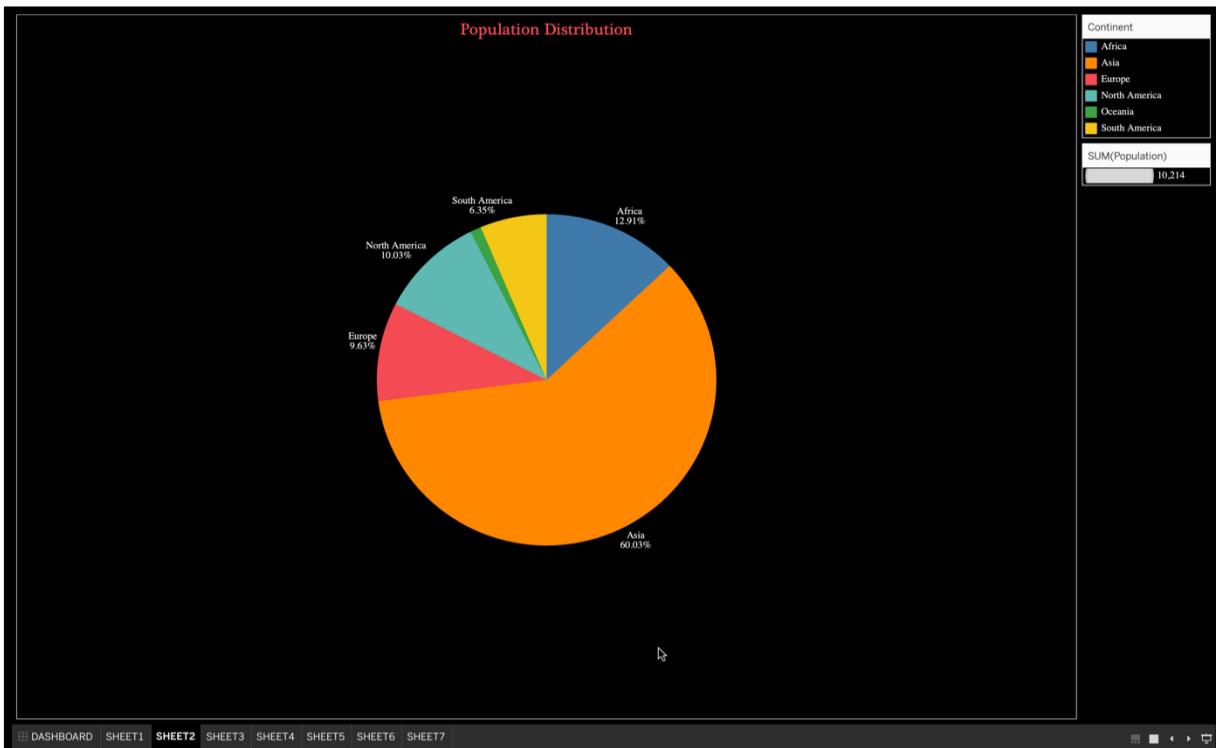
```
146
147 sheet2 <- df[,c('continent', 'country', 'population')]
148 write_xlsx(sheet2,"sheet2.xlsx")
149
```

Queries.R* x sheet1 x sheet2 x

Filter

	continent	country	population
1	Asia	Afghanistan	39.84
2	Europe	Albania	25.90
3	Africa	Algeria	69.40
4	Oceania	American Samoa	7.70
5	Europe	Andorra	25.90
6	North America	Antigua and Barbuda	85.40
7	South America	Argentina	45.81
8	Oceania	Australia	25.74
9	Europe	Austria	25.90
10	Asia	Bahrain	1.75
11	Asia	Bangladesh	166.30
12	Europe	Belarus	25.90
13	Europe	Belgium	11.59
14	North America	Bermuda	85.40
15	South America	Bolivia	92.60
16	Europe	Bosnia and Herzegovina	25.90
17	South America	Brazil	213.99
18	Asia	Brunei	0.44
19	Europe	Bulgaria	25.90
20	Asia	Burma	170.30

Showing 1 to 20 of 126 entries, 3 total columns



SHEET 3:

Visualizing continent area using bubble chart.

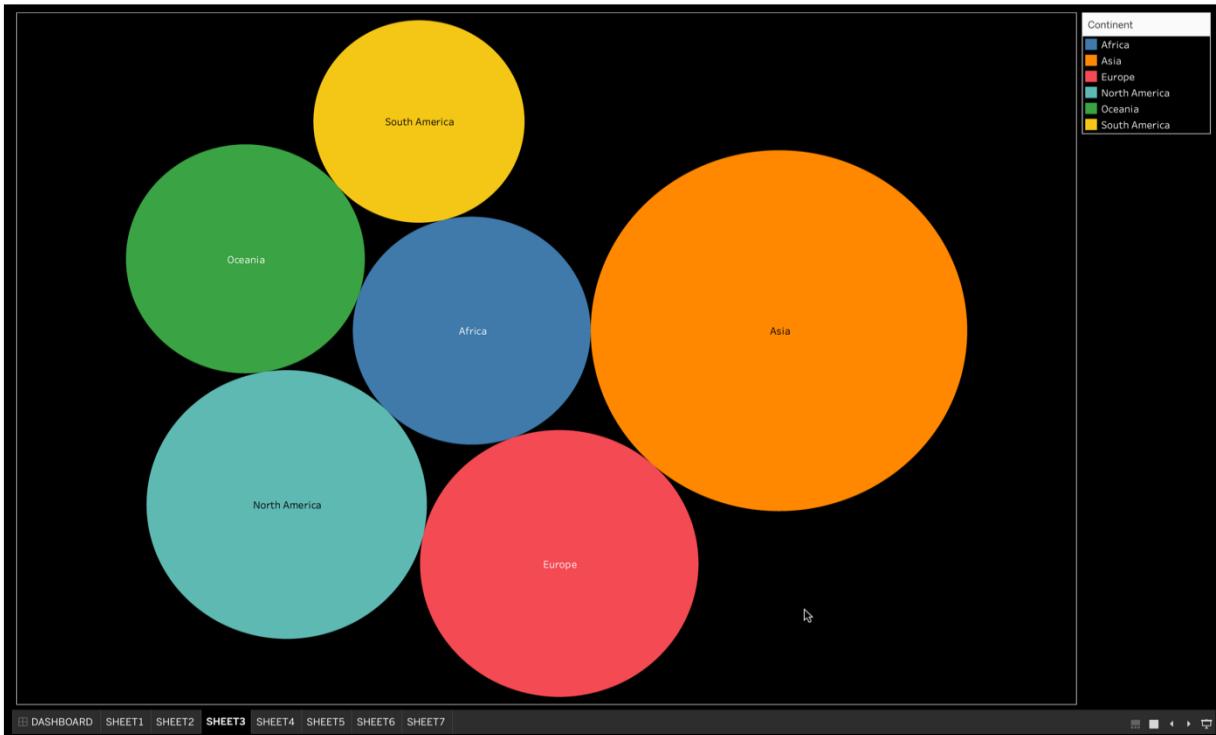
Generating Dataset for sheet 3:

```
151  
152 sheet3 <- df[,c('continent', 'country', 'area')]  
153 write_xlsx(sheet3,"sheet3.xlsx")  
154
```

The screenshot shows a Jupyter Notebook interface with multiple tabs at the top: 'Queries.R*', 'sheet3', 'sheet1', 'sheet2'. The 'sheet3' tab is active, displaying a table with three columns: 'continent', 'country', and 'area'. The data consists of 20 rows, each representing a country and its area, categorized by continent. The table is sorted by continent. A search bar and filter icon are visible at the top right of the table area.

	continent	country	area
1	Asia	Afghanistan	652860
2	Europe	Albania	1388170
3	Africa	Algeria	1388170
4	Oceania	American Samoa	1388170
5	Europe	Andorra	1388170
6	North America	Antigua and Barbuda	1388170
7	South America	Argentina	2780400
8	Oceania	Australia	7741220
9	Europe	Austria	1388170
10	Asia	Bahrain	778
11	Asia	Bangladesh	147630
12	Europe	Belarus	1388170
13	Europe	Belgium	30530
14	North America	Bermuda	1388170
15	South America	Bolivia	1388170
16	Europe	Bosnia and Herzegovina	1388170
17	South America	Brazil	8515770
18	Asia	Brunei	5770
19	Europe	Bulgaria	1388170
20	Asia	Burma	1388170

Showing 1 to 20 of 126 entries, 3 total columns



Greater size of bubble implies greater area.

SHEET 4:

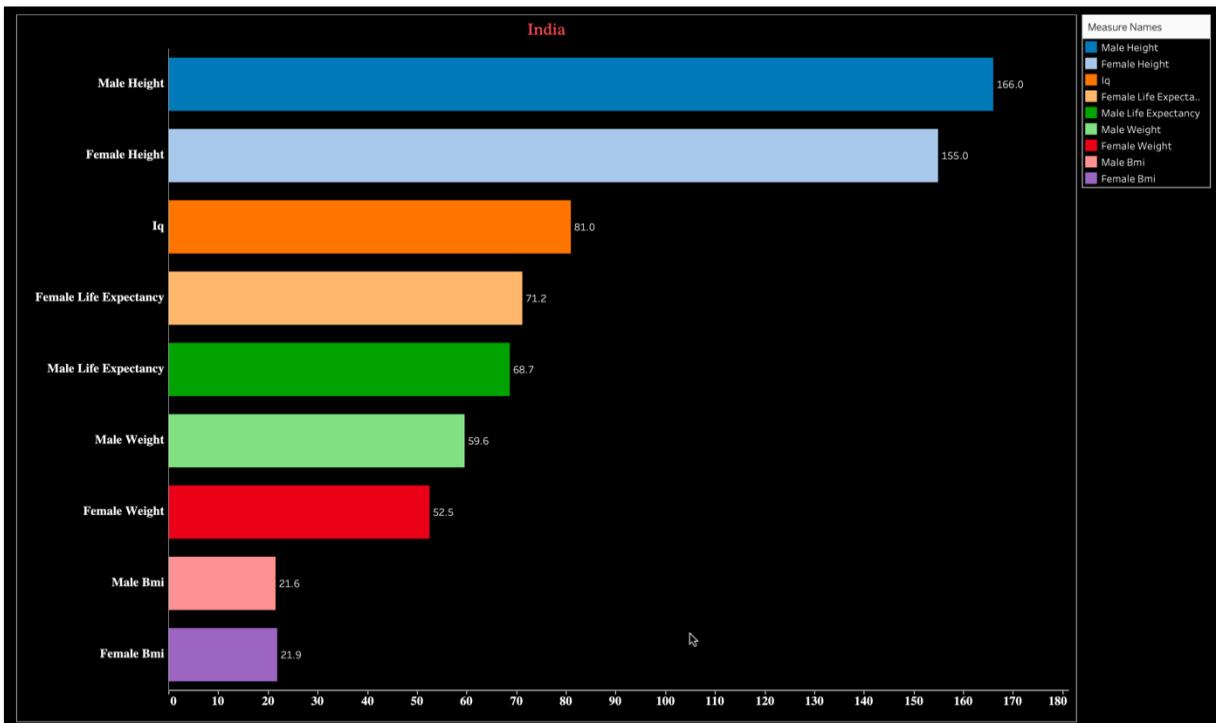
Visualizing country wise demographics using bar chart.

Generating Dataset for sheet 4:

```
157 sheet4 <- df[,c('continent', 'country', 'male_height', 'female_height', 'male_weight', 'female_weight',
158                           'male_bmi', 'female_bmi', 'male_life_expectancy', 'female_life_expectancy', 'iq')]
159 write_xlsx(sheet4,"sheet4.xlsx")
160
```

	continent	country	male_height	female_height	male_weight	female_weight	male_bmi	female_bmi	male_life_expectancy
1	Asia	Afghanistan	168	155	63.6	62.6	22.6	26.1	
2	Europe	Albania	174	162	81.4	69.2	27.0	26.4	
3	Africa	Algeria	174	162	74.6	69.4	24.7	26.4	
4	Oceania	American Samoa	177	167	103.2	85.1	33.1	30.5	
5	Europe	Andorra	178	165	87.3	71.7	27.6	26.3	
6	North America	Antigua and Barbuda	178	165	81.6	75.9	25.7	27.9	
7	South America	Argentina	171	161	84.7	71.4	27.9	27.5	
8	Oceania	Australia	179	165	88.3	85.1	27.7	31.3	
9	Europe	Austria	178	166	84.6	68.3	26.6	24.8	
10	Asia	Bahrain	172	158	74.0	62.2	24.9	24.9	
11	Asia	Bangladesh	165	152	57.7	50.5	21.3	21.9	
12	Europe	Belarus	178	166	84.1	74.4	26.5	27.0	
13	Europe	Belgium	179	165	85.9	68.8	26.8	25.3	
14	North America	Bermuda	179	166	88.4	80.4	27.6	29.2	
15	South America	Bolivia	168	155	71.2	66.8	25.3	27.8	
16	Europe	Bosnia and Herzegovina	182	167	87.1	70.6	26.4	25.3	
17	South America	Brazil	175	162	80.7	70.3	26.3	26.8	
18	Asia	Brunei	166	155	74.7	65.0	27.1	27.1	
19	Europe	Bulgaria	179	164	81.8	69.6	27.2	25.9	
20	Asia	Burma	166	154	71.5	54.7	22.2	23.1	

Showing 1 to 19 of 126 entries, 11 total columns



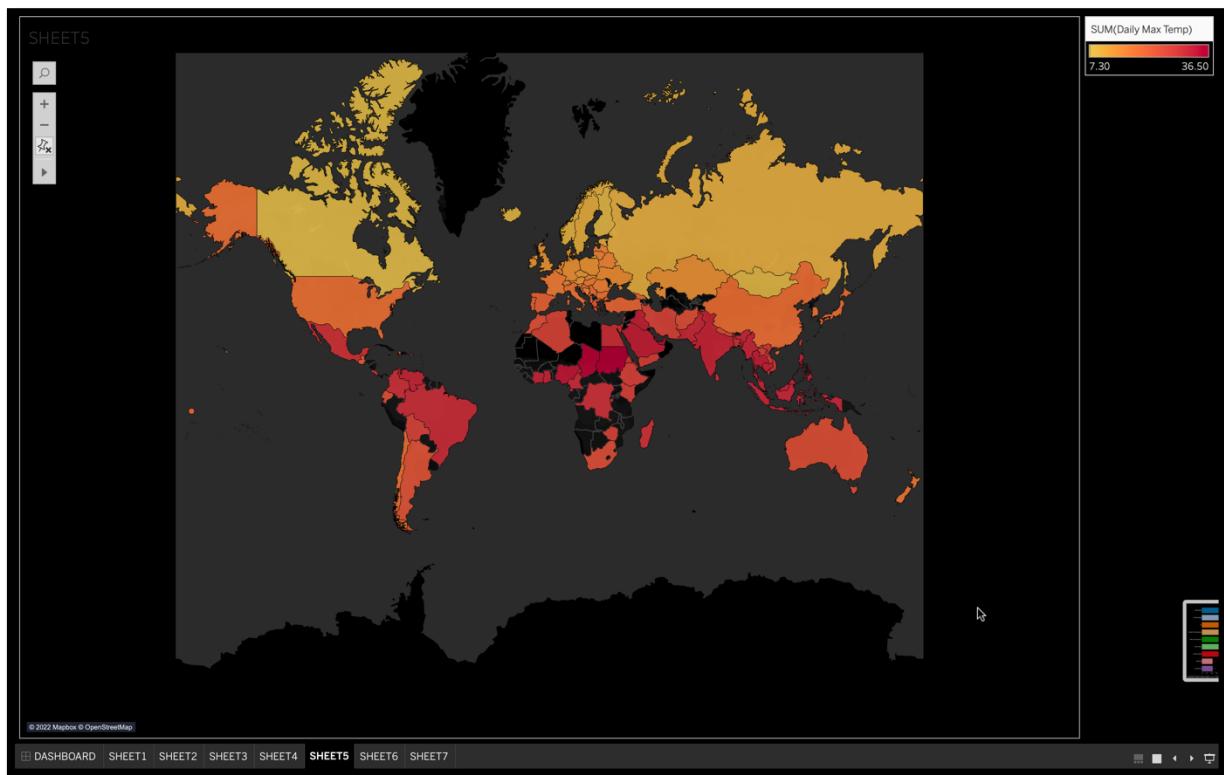
SHEET 5:

Visualizing max temperature of various countries using world map.

Generating Dataset for sheet 5:

```
162  
163 write_xlsx(sheet5,"sheet5.xlsx")  
164 sheet5 <- df[,c('continent', 'country', 'area', 'daily_max_temp')]  
165  
166
```

	continent	country	area	daily_max_temp
1	Asia	Afghanistan	652860	24.4
2	Europe	Albania	1388170	21.2
3	Africa	Algeria	1388170	26.5
4	Oceania	American Samoa	1388170	20.9
5	Europe	Andorra	1388170	15.5
6	North America	Antigua and Barbuda	1388170	21.8
7	South America	Argentina	2780400	23.9
8	Oceania	Australia	7741220	24.4
9	Europe	Austria	1388170	13.1
10	Asia	Bahrain	778	32.6
11	Asia	Bangladesh	147630	29.8
12	Europe	Belarus	1388170	15.5
13	Europe	Belgium	30530	14.8
14	North America	Bermuda	1388170	24.6
15	South America	Bolivia	1388170	27.0
16	Europe	Bosnia and Herzegovina	1388170	15.5
17	South America	Brazil	8515770	30.7
18	Asia	Brunei	5770	31.7
19	Europe	Bulgaria	1388170	18.1
20	Asia	Burma	1388170	32.2



Darker color depicts hotter region while lighter color depicts colder regions. We can also see its area by hovering over any country.

SHEET 6:

Visualizing population growth of each country using scatter plot.

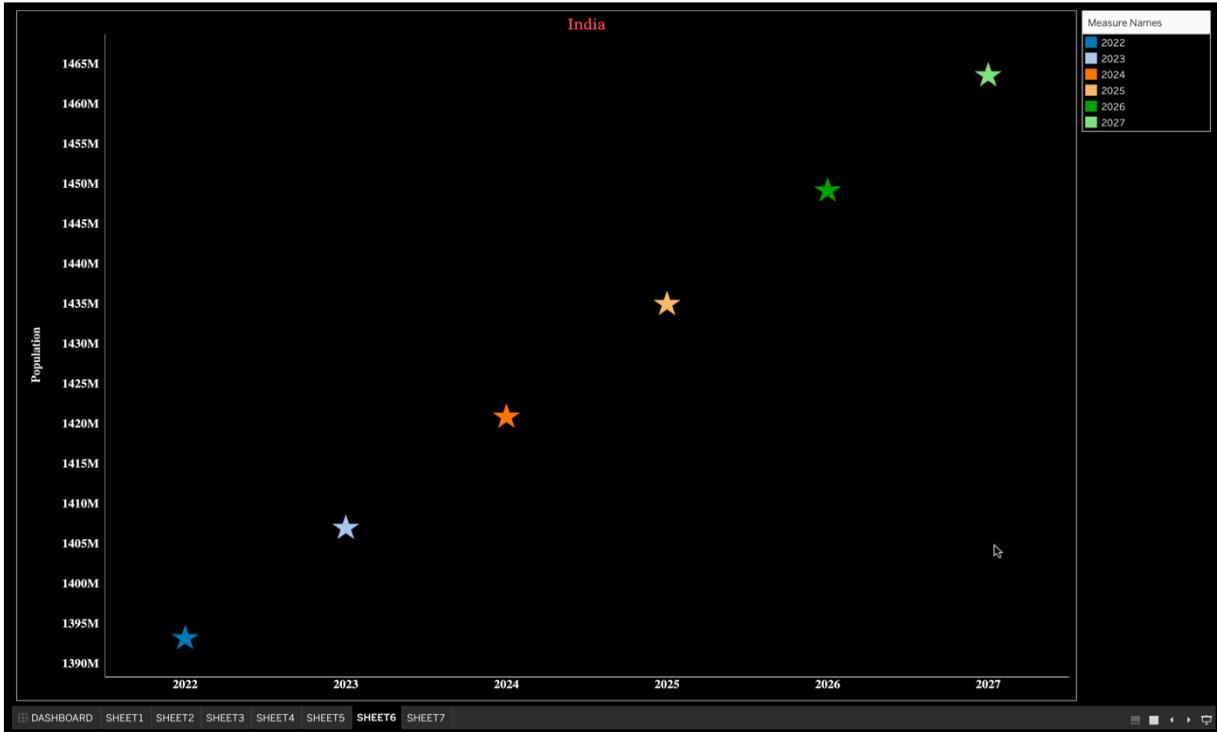
Generating Dataset for sheet 6:

```
167  
168 write_xlsx(sheet6,"sheet6.xlsx")  
169 sheet6 <- df[,c('continent', 'country', 'population', 'population_change')]  
170
```

A screenshot of a Jupyter Notebook interface. The current cell shows the code used to generate the dataset. Below the code, the resulting DataFrame is displayed as a table. The table has four columns: continent, country, population, and population_change. The data includes entries for various countries across continents like Asia, Europe, Africa, Oceania, and North/South America, with their respective populations and growth rates. The table is sorted by continent and country.

	continent	country	population	population_change
1	Asia	Afghanistan	39.84	2.33%
2	Europe	Albania	25.90	-0.11%
3	Africa	Algeria	69.40	1.85%
4	Oceania	American Samoa	7.70	-0.22%
5	Europe	Andorra	25.90	0.16%
6	North America	Antigua and Barbuda	85.40	0.84%
7	South America	Argentina	45.81	0.93%
8	Oceania	Australia	25.74	1.18%
9	Europe	Austria	25.90	0.57%
10	Asia	Bahrain	1.75	3.68%
11	Asia	Bangladesh	166.30	1.01%
12	Europe	Belarus	25.90	-0.03%
13	Europe	Belgium	11.59	0.44%
14	North America	Bermuda	85.40	-0.36%
15	South America	Bolivia	92.60	1.39%
16	Europe	Bosnia and Herzegovina	25.90	-0.61%
17	South America	Brazil	213.99	0.72%
18	Asia	Brunei	0.44	0.97%
19	Europe	Bulgaria	25.90	-0.74%
20	Asia	Burma	170.30	0.67%

Showing 1 to 20 of 126 entries, 4 total columns



SHEET 7:

Visualizing total world population filtered by year.

Generating Dataset for sheet 7:

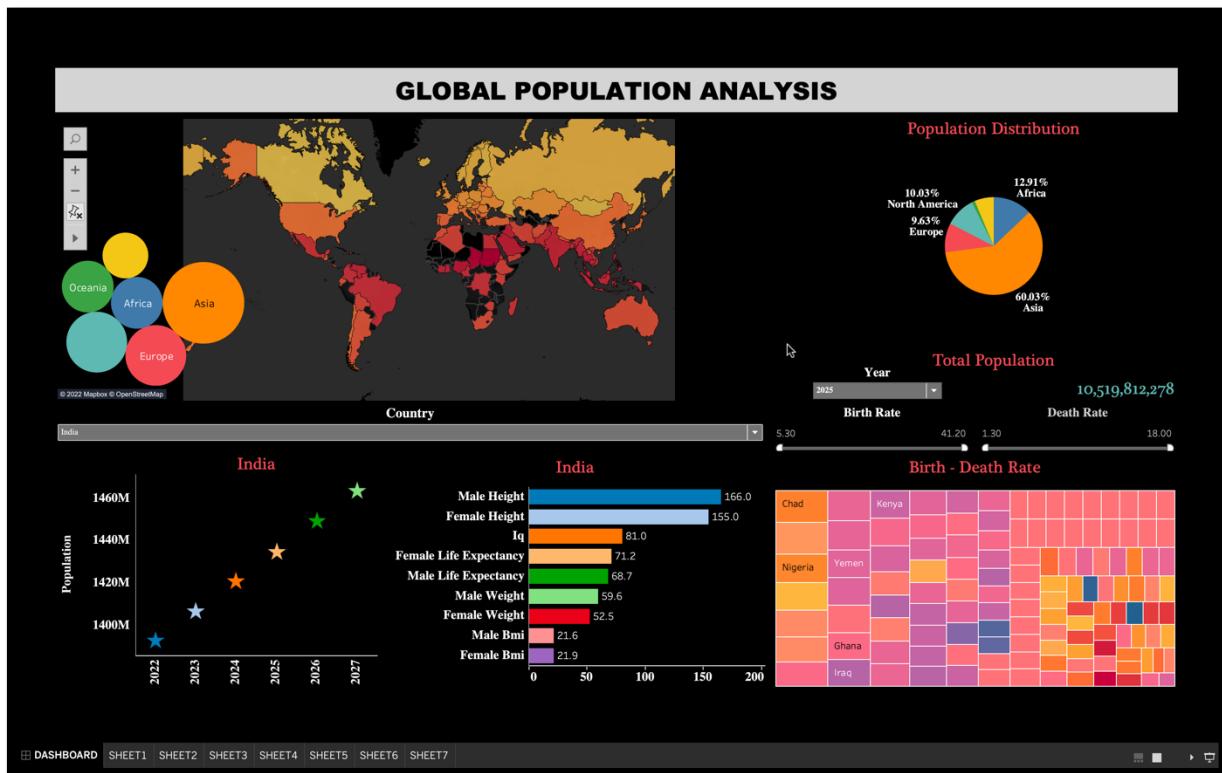
```
173 write_xlsx(sheet7,"sheet7.xlsx")
174 sheet7 <- df[,c('continent', 'country', 'population', 'population_change')]
175
```

	continent	country	population	population_change
1	Asia	Afghanistan	39.84	2.33%
2	Europe	Albania	25.90	-0.11%
3	Africa	Algeria	69.40	1.85%
4	Oceania	American Samoa	7.70	-0.22%
5	Europe	Andorra	25.90	0.16%
6	North America	Antigua and Barbuda	85.40	0.84%
7	South America	Argentina	45.81	0.93%
8	Oceania	Australia	25.74	1.18%
9	Europe	Austria	25.90	0.57%
10	Asia	Bahrain	1.75	3.68%
11	Asia	Bangladesh	166.30	1.01%
12	Europe	Belarus	25.90	-0.03%
13	Europe	Belgium	11.59	0.44%
14	North America	Bermuda	85.40	-0.36%
15	South America	Bolivia	92.60	1.39%
16	Europe	Bosnia and Herzegovina	25.90	-0.61%
17	South America	Brazil	213.99	0.72%
18	Asia	Brunei	0.44	0.97%
19	Europe	Bulgaria	25.90	-0.74%
20	Asia	Burma	170.30	0.67%

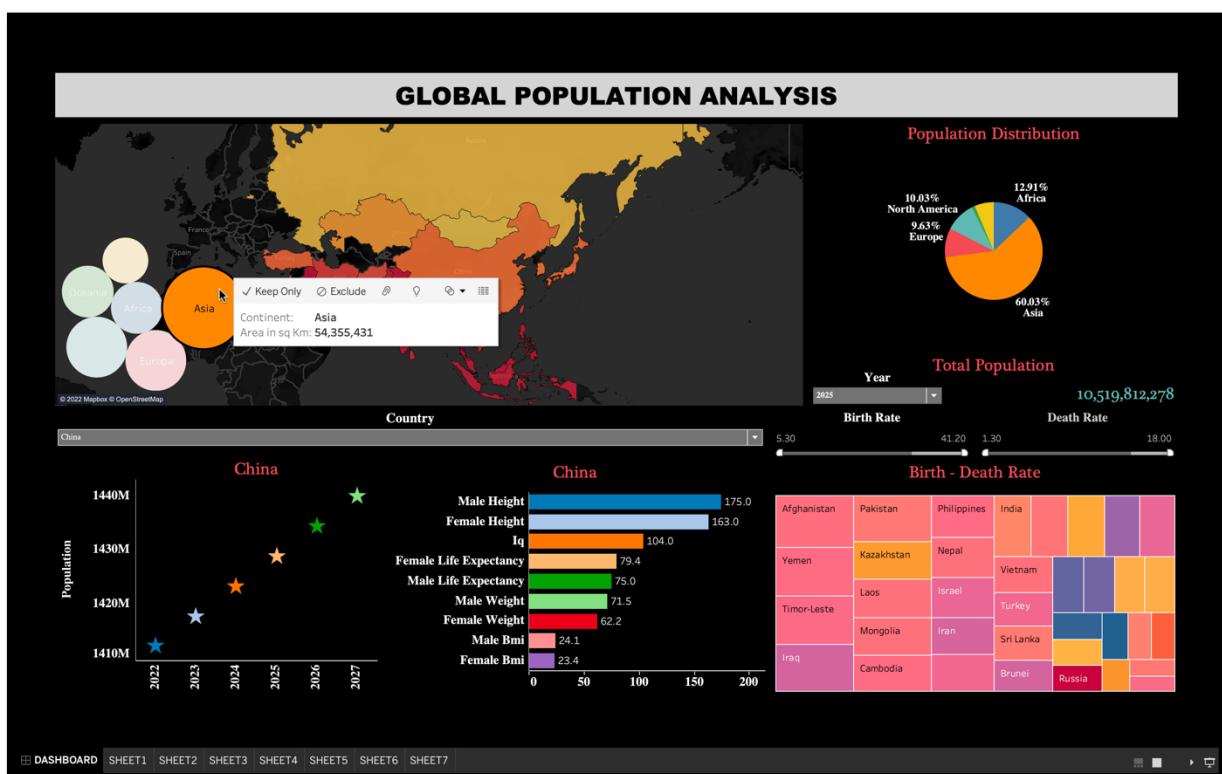
Showing 1 to 20 of 126 entries, 4 total columns



Final Dashboard:



Using Continent as filter:



Using Country as filter:

