

Health Data Analysis

Umang Parti

2023-12-10

This Project is an Assignment for Data Analysis with R
Submitted to Edureka By Umang Parti

```
getwd()
```

```
## [1] "C:/Users/umang/Desktop/New folder"
```

Let's Read the data

```
health_data<- read.csv("health_data.csv")  
head(health_data)
```

```
##           Country           Region Population Under15 Over60  
## 1      Afghanistan Eastern Mediterranean    29825   47.42   3.82  
## 2          Albania             Europe      3162   21.33  14.93  
## 3          Algeria             Africa   38482   27.42   7.17  
## 4          Andorra             Europe      78   15.20  22.86  
## 5          Angola             Africa   20821   47.58   3.84  
## 6 Antigua and Barbuda           Americas      89   25.96  12.35  
##      FertilityRate LifeExpectancy ChildMortality CellularSubscribers  
LiteracyRate  
## 1           5.40           60           98.5           54.26  
NA  
## 2           1.75           74           16.7           96.39  
NA  
## 3           2.83           73           20.0           98.99  
NA  
## 4           NA           82           3.2           75.49  
NA  
## 5           6.10           51          163.5           48.38  
70.1  
## 6           2.12           75           9.9           196.41  
99.0  
##      GNI PrimarySchoolEnrollmentMale PrimarySchoolEnrollmentFemale  
## 1 1140           NA           NA  
## 2 8820           NA           NA  
## 3 8310           98.2           96.4  
## 4  NA           78.4           79.4  
## 5 5230           93.1           78.2  
## 6 17900          91.1           84.5
```

Let's Clean the data

```
#Task 1: Load the dataset into R by replacing empty
#Ensure that the dataset is consisting NA values by replacing empty strings
with NA.
```

```
health_data<- read.csv("health_data.csv", na.strings = c("", "NA"))
```

```
#Task 2: Locating Missing Data
```

```
missing_data <- sum(is.na(health_data))
```

```
missing_data
```

```
## [1] 330
```

```
#Task 3: Filtering data using WHICH()
```

```
# We want to filter out rows having LifeExpectancy value=71
```

```
filtered_data <- health_data[-which(health_data$LifeExpectancy == 71),]
```

```
filtered_data
```

```
##              Country              Region
Population
## 1      Afghanistan Eastern Mediterranean
29825
## 2           Albania              Europe
3162
## 3           Algeria              Africa
38482
## 4           Andorra              Europe
78
## 5           Angola              Africa
20821
## 6  Antigua and Barbuda              Americas
89
## 7           Argentina              Americas
41087
## 9           Australia      Western Pacific
23050
## 10          Austria              Europe
8464
## 12          Bahamas              Americas
372
## 13          Bahrain Eastern Mediterranean
1318
## 14        Bangladesh      South-East Asia
155000
## 15          Barbados              Americas
283
## 17          Belgium              Europe
11060
## 18          Belize              Americas
324
## 19          Benin              Africa
10051
```

## 20	Bhutan	South-East Asia
742		
## 21	Bolivia (Plurinational State of)	Americas
10496		
## 22	Bosnia and Herzegovina	Europe
3834		
## 23	Botswana	Africa
2004		
## 24	Brazil	Americas
199000		
## 25	Brunei Darussalam	Western Pacific
412		
## 26	Bulgaria	Europe
7278		
## 27	Burkina Faso	Africa
16460		
## 28	Burundi	Africa
9850		
## 29	Cambodia	Western Pacific
14865		
## 30	Cameroon	Africa
21700		
## 31	Canada	Americas
34838		
## 32	Cape Verde	Africa
494		
## 33	Central African Republic	Africa
4525		
## 34	Chad	Africa
12448		
## 35	Chile	Americas
17465		
## 36	China	Western Pacific
1390000		
## 37	Colombia	Americas
47704		
## 38	Comoros	Africa
718		
## 39	Congo	Africa
4337		
## 40	Cook Islands	Western Pacific
21		
## 41	Costa Rica	Americas
4805		
## 42	Ivory Coast	Africa
19840		
## 43	Croatia	Europe
4307		
## 44	Cuba	Americas
11271		

## 45	Cyprus	Europe
1129		
## 46	Czech Republic	Europe
10660		
## 47	Democratic People's Republic of Korea	South-East Asia
24763		
## 48	Democratic Republic of the Congo	Africa
65705		
## 49	Denmark	Europe
5598		
## 50	Djibouti	Eastern Mediterranean
860		
## 51	Dominica	Americas
72		
## 52	Dominican Republic	Americas
10277		
## 53	Ecuador	Americas
15492		
## 54	Egypt	Eastern Mediterranean
80722		
## 55	El Salvador	Americas
6297		
## 56	Equatorial Guinea	Africa
736		
## 57	Eritrea	Africa
6131		
## 58	Estonia	Europe
1291		
## 59	Ethiopia	Africa
91729		
## 60	Fiji	Western Pacific
875		
## 61	Finland	Europe
5408		
## 62	France	Europe
63937		
## 63	Gabon	Africa
1633		
## 64	Gambia	Africa
1791		
## 65	Georgia	Europe
4358		
## 66	Germany	Europe
82800		
## 67	Ghana	Africa
25366		
## 68	Greece	Europe
11125		
## 69	Grenada	Americas
105		

## 70 15083	Guatemala	Americas
## 71 11451	Guinea	Africa
## 72 1664	Guinea-Bissau	Africa
## 73 795	Guyana	Americas
## 74 10174	Haiti	Americas
## 75 7936	Honduras	Americas
## 76 9976	Hungary	Europe
## 77 326	Iceland	Europe
## 78 1240000	India	South-East Asia
## 79 247000	Indonesia	South-East Asia
## 80 76424	Iran (Islamic Republic of)	Eastern Mediterranean
## 81 32778	Iraq	Eastern Mediterranean
## 82 4576	Ireland	Europe
## 83 7644	Israel	Europe
## 84 60885	Italy	Europe
## 85 2769	Jamaica	Americas
## 86 127000	Japan	Western Pacific
## 87 7009	Jordan	Eastern Mediterranean
## 88 16271	Kazakhstan	Europe
## 89 43178	Kenya	Africa
## 90 101	Kiribati	Western Pacific
## 91 3250	Kuwait	Eastern Mediterranean
## 92 5474	Kyrgyzstan	Europe
## 93 6646	Lao People's Democratic Republic	Western Pacific
## 94 2060	Latvia	Europe

## 95	Lebanon	Eastern Mediterranean
4647		
## 96	Lesotho	Africa
2052		
## 97	Liberia	Africa
4190		
## 98	Libya	Eastern Mediterranean
6155		
## 99	Lithuania	Europe
3028		
## 100	Luxembourg	Europe
524		
## 101	Madagascar	Africa
22294		
## 102	Malawi	Africa
15906		
## 103	Malaysia	Western Pacific
29240		
## 104	Maldives	South-East Asia
338		
## 105	Mali	Africa
14854		
## 106	Malta	Europe
428		
## 107	Marshall Islands	Western Pacific
53		
## 108	Mauritania	Africa
3796		
## 109	Mauritius	Africa
1240		
## 110	Mexico	Americas
121000		
## 111	Micronesia (Federated States of)	Western Pacific
103		
## 112	Monaco	Europe
38		
## 113	Mongolia	Western Pacific
2796		
## 114	Montenegro	Europe
621		
## 115	Morocco	Eastern Mediterranean
32521		
## 116	Mozambique	Africa
25203		
## 117	Myanmar	South-East Asia
52797		
## 118	Namibia	Africa
2259		
## 120	Nepal	South-East Asia
27474		

## 121 16714	Netherlands	Europe
## 122 4460	New Zealand	Western Pacific
## 123 5992	Nicaragua	Americas
## 124 17157	Niger	Africa
## 125 169000	Nigeria	Africa
## 126 1	Niue	Western Pacific
## 127 4994	Norway	Europe
## 128 3314	Oman	Eastern Mediterranean
## 129 179000	Pakistan	Eastern Mediterranean
## 130 21	Palau	Western Pacific
## 131 3802	Panama	Americas
## 132 7167	Papua New Guinea	Western Pacific
## 133 6687	Paraguay	Americas
## 134 29988	Peru	Americas
## 135 96707	Philippines	Western Pacific
## 136 38211	Poland	Europe
## 137 10604	Portugal	Europe
## 138 2051	Qatar	Eastern Mediterranean
## 139 49003	Republic of Korea	Western Pacific
## 141 21755	Romania	Europe
## 142 143000	Russian Federation	Europe
## 143 11458	Rwanda	Africa
## 144 54	Saint Kitts and Nevis	Americas
## 145 181	Saint Lucia	Americas
## 146 109	Saint Vincent and the Grenadines	Americas

## 147 189	Samoa	Western Pacific
## 148 31	San Marino	Europe
## 149 188	Sao Tome and Principe	Africa
## 150 28288	Saudi Arabia	Eastern Mediterranean
## 151 13726	Senegal	Africa
## 152 9553	Serbia	Europe
## 153 92	Seychelles	Africa
## 154 5979	Sierra Leone	Africa
## 155 5303	Singapore	Western Pacific
## 156 5446	Slovakia	Europe
## 157 2068	Slovenia	Europe
## 158 550	Solomon Islands	Western Pacific
## 159 10195	Somalia	Eastern Mediterranean
## 160 52386	South Africa	Africa
## 161 10838	South Sudan	Eastern Mediterranean
## 162 46755	Spain	Europe
## 163 21098	Sri Lanka	South-East Asia
## 164 37195	Sudan	Eastern Mediterranean
## 165 535	Suriname	Americas
## 166 1231	Swaziland	Africa
## 167 9511	Sweden	Europe
## 168 7997	Switzerland	Europe
## 169 21890	Syrian Arab Republic	Eastern Mediterranean
## 170 8009	Tajikistan	Europe
## 171 66785	Thailand	South-East Asia

## 172	The former Yugoslav Republic of Macedonia	Europe
2106		
## 173	Timor-Leste	South-East Asia
1114		
## 174	Togo	Africa
6643		
## 175	Tonga	Western Pacific
105		
## 177	Tunisia	Eastern Mediterranean
10875		
## 178	Turkey	Europe
73997		
## 179	Turkmenistan	Europe
5173		
## 180	Tuvalu	Western Pacific
10		
## 181	Uganda	Africa
36346		
## 183	United Arab Emirates	Eastern Mediterranean
9206		
## 184	United Kingdom	Europe
62783		
## 185	United Republic of Tanzania	Africa
47783		
## 186	United States of America	Americas
318000		
## 187	Uruguay	Americas
3395		
## 188	Uzbekistan	Europe
28541		
## 189	Vanuatu	Western Pacific
247		
## 190	Venezuela (Bolivarian Republic of)	Americas
29955		
## 191	Viet Nam	Western Pacific
90796		
## 192	Yemen	Eastern Mediterranean
23852		
## 193	Zambia	Africa
14075		
## 194	Zimbabwe	Africa
13724		

##	Under15	Over60	FertilityRate	LifeExpectancy	ChildMortality
## 1	47.42	3.82	5.40	60	98.5
## 2	21.33	14.93	1.75	74	16.7
## 3	27.42	7.17	2.83	73	20.0
## 4	15.20	22.86	NA	82	3.2
## 5	47.58	3.84	6.10	51	163.5
## 6	25.96	12.35	2.12	75	9.9
## 7	24.42	14.97	2.20	76	14.2

## 9	18.95	19.46	1.89	82	4.9
## 10	14.51	23.52	1.44	81	4.0
## 12	21.62	11.24	1.90	75	16.9
## 13	20.16	3.38	2.12	79	9.6
## 14	30.57	6.89	2.24	70	40.9
## 15	18.99	15.78	1.84	78	18.4
## 17	16.88	23.81	1.85	80	4.2
## 18	34.40	5.74	2.76	74	18.3
## 19	42.95	4.54	5.01	57	89.5
## 20	28.53	6.90	2.32	67	44.6
## 21	35.23	7.28	3.31	67	41.4
## 22	16.35	20.52	1.26	76	6.7
## 23	33.75	5.63	2.71	66	53.3
## 24	24.56	10.81	1.82	74	14.4
## 25	25.75	7.03	2.03	77	8.0
## 26	13.53	26.11	1.51	74	12.1
## 27	45.66	3.88	5.78	56	102.4
## 28	44.20	3.87	6.21	53	104.3
## 29	31.23	7.67	2.93	65	39.7
## 30	43.08	4.89	4.94	53	94.9
## 31	16.37	20.82	1.66	82	5.3
## 32	30.17	7.05	2.38	72	22.2
## 33	40.07	5.74	4.54	48	128.6
## 34	48.52	3.80	6.49	51	149.8
## 35	21.38	13.80	1.84	79	9.1
## 36	17.95	13.42	1.66	76	14.0
## 37	28.03	9.19	2.35	78	17.6
## 38	42.17	4.50	4.85	62	77.6
## 39	42.37	5.13	5.05	58	96.0
## 40	30.61	9.07	NA	77	10.6
## 41	23.94	10.15	1.83	79	9.9
## 42	41.48	5.10	4.91	56	107.6
## 43	14.98	24.69	1.48	77	4.7
## 44	16.58	17.95	1.46	78	5.5
## 45	17.16	16.92	1.47	81	3.2
## 46	14.56	23.23	1.53	78	3.8
## 47	21.98	12.74	2.00	69	28.8
## 48	45.11	4.51	6.15	49	145.7
## 49	17.66	23.90	1.88	79	3.7
## 50	33.72	5.96	3.53	58	80.9
## 51	25.96	12.35	NA	74	12.6
## 52	30.53	8.97	2.55	73	27.1
## 53	30.29	9.21	2.62	76	23.3
## 54	31.25	8.62	2.85	73	21.0
## 55	30.62	9.64	2.24	72	15.9
## 56	38.95	4.53	5.04	54	100.3
## 57	43.10	3.73	4.88	61	51.8
## 58	15.69	23.92	1.62	76	3.6
## 59	43.29	5.17	4.77	60	68.3
## 60	28.88	8.38	2.64	70	22.4

## 61	16.42	25.90	1.85	81	2.9
## 62	18.26	23.82	1.98	82	4.1
## 63	38.49	7.38	4.18	62	62.0
## 64	45.90	3.72	5.79	58	72.9
## 65	17.62	19.47	1.82	72	19.9
## 66	13.17	26.72	1.40	81	4.1
## 67	38.59	5.40	3.99	64	72.0
## 68	14.60	25.41	1.51	81	4.8
## 69	26.96	9.72	2.22	74	13.5
## 70	40.80	6.56	3.91	69	32.0
## 71	42.46	5.03	5.09	55	101.2
## 72	41.55	5.06	5.05	50	129.1
## 73	36.77	5.18	2.64	63	35.2
## 74	35.35	6.70	3.28	63	75.6
## 75	35.72	6.41	3.10	74	22.9
## 76	14.62	23.41	1.38	75	6.2
## 77	20.71	17.62	2.11	82	2.3
## 78	29.43	8.10	2.53	65	56.3
## 79	29.27	7.86	2.40	69	31.0
## 80	23.68	7.82	1.91	73	17.6
## 81	40.51	4.95	4.15	69	34.4
## 82	21.54	16.59	2.00	81	4.0
## 83	27.53	15.15	2.92	82	4.2
## 84	14.04	26.97	1.45	82	3.8
## 85	27.78	10.98	2.31	75	16.8
## 86	13.12	31.92	1.39	83	3.0
## 87	34.13	5.30	3.39	74	19.1
## 88	25.46	10.04	2.52	67	18.7
## 89	42.37	4.25	4.54	60	72.9
## 90	30.10	8.84	3.01	67	59.9
## 91	24.90	3.80	2.65	80	11.0
## 92	30.21	6.34	3.03	69	26.6
## 93	35.61	5.76	3.20	68	71.8
## 94	14.57	24.24	1.57	74	8.7
## 95	21.64	12.03	1.50	74	9.3
## 96	36.75	6.31	3.15	50	99.6
## 97	43.06	4.76	4.95	59	74.8
## 98	29.45	6.96	2.47	65	15.4
## 99	15.13	20.57	1.49	74	5.4
## 100	17.46	19.15	1.65	82	2.2
## 101	42.72	4.45	4.59	66	58.2
## 102	45.44	4.92	5.55	58	71.0
## 103	26.65	8.21	1.99	74	8.5
## 104	29.03	6.65	2.31	77	10.5
## 105	47.14	4.29	6.85	51	128.0
## 106	14.98	22.87	1.37	80	6.8
## 107	30.10	8.84	NA	60	37.9
## 108	40.22	4.94	4.78	59	84.0
## 109	20.17	13.23	1.51	74	15.1
## 110	29.02	9.18	2.25	75	16.2

## 111	35.81	6.67	3.40	69	38.5
## 112	18.26	23.82	NA	82	3.8
## 113	27.05	5.80	2.45	68	27.5
## 114	19.01	18.58	1.69	76	5.9
## 115	27.85	7.61	2.65	72	31.1
## 116	45.38	5.01	5.34	53	89.7
## 117	25.28	8.15	1.98	65	52.3
## 118	36.59	5.38	3.17	65	38.7
## 120	35.58	7.65	2.50	68	41.6
## 121	17.21	23.02	1.76	81	4.1
## 122	20.26	19.01	2.10	81	5.7
## 123	33.37	6.59	2.59	73	24.4
## 124	49.99	4.26	7.58	56	113.5
## 125	44.23	4.49	6.02	53	123.7
## 126	30.61	9.07	NA	72	25.1
## 127	18.64	21.41	1.93	81	2.8
## 128	24.19	3.99	2.90	72	11.6
## 129	34.31	6.44	3.35	67	85.9
## 130	30.10	8.84	NA	72	20.8
## 131	28.65	10.13	2.52	77	18.5
## 132	38.37	4.79	3.90	63	63.0
## 133	32.78	8.01	2.93	75	22.0
## 134	29.18	9.12	2.48	77	18.2
## 135	34.53	6.21	3.11	69	29.8
## 136	14.91	20.48	1.39	76	5.0
## 137	14.92	24.39	1.33	80	3.6
## 138	13.28	1.73	2.06	82	7.4
## 139	15.25	16.58	1.29	81	3.8
## 141	15.05	20.66	1.39	74	12.2
## 142	15.45	18.60	1.51	69	10.3
## 143	43.56	3.94	4.73	60	55.0
## 144	25.96	12.35	NA	74	9.2
## 145	24.31	12.13	1.96	75	17.5
## 146	25.70	9.92	2.05	74	23.4
## 147	37.88	7.39	4.28	73	17.8
## 148	14.04	26.97	NA	83	3.3
## 149	41.60	4.76	4.22	63	53.2
## 150	29.69	4.59	2.76	76	8.6
## 151	43.54	4.57	5.02	61	59.6
## 152	16.45	20.52	1.37	74	6.6
## 153	21.95	10.05	2.23	74	13.1
## 154	41.74	4.41	4.86	47	181.6
## 155	16.48	15.13	1.27	82	2.9
## 156	15.00	18.60	1.37	76	7.5
## 157	14.16	23.16	1.49	80	3.1
## 158	40.37	5.10	4.17	70	31.1
## 159	47.35	4.46	6.77	50	147.4
## 160	29.53	8.44	2.44	58	44.6
## 161	42.28	5.26	5.10	54	104.0
## 162	15.20	22.86	1.47	82	4.5

## 163	25.15	12.40	2.35	75	9.6
## 164	41.48	4.99	4.56	62	73.1
## 165	27.83	9.55	2.32	72	20.8
## 166	38.05	5.34	3.48	50	79.7
## 167	16.71	25.32	1.93	82	2.9
## 168	14.79	23.25	1.51	83	4.3
## 169	35.35	6.09	3.04	75	15.1
## 170	35.75	4.80	3.81	68	58.3
## 171	18.47	13.96	1.43	74	13.2
## 172	16.89	17.56	1.44	75	7.4
## 173	46.33	5.16	6.11	64	56.7
## 174	41.89	4.44	4.75	56	95.5
## 175	37.33	7.96	3.86	72	12.8
## 177	23.22	10.49	2.04	76	16.1
## 178	26.00	10.56	2.08	76	14.2
## 179	28.65	6.30	2.38	63	52.8
## 180	30.61	9.07	NA	64	29.7
## 181	48.54	3.72	6.06	56	68.9
## 183	14.41	0.81	1.84	76	8.4
## 184	17.54	23.06	1.90	80	4.8
## 185	44.85	4.89	5.36	59	54.0
## 186	19.63	19.31	2.00	79	7.1
## 187	22.05	18.59	2.07	77	7.2
## 188	28.90	6.38	2.38	68	39.6
## 189	37.37	6.02	3.46	72	17.9
## 190	28.84	9.17	2.44	75	15.3
## 191	22.87	9.32	1.79	75	23.0
## 192	40.72	4.54	4.35	64	60.0
## 193	46.73	3.95	5.77	55	88.5
## 194	40.24	5.68	3.64	54	89.8
##	CellularSubscribers	LiteracyRate	GNI	PrimarySchoolEnrollment	Male
## 1	54.26	NA	1140		NA
## 2	96.39	NA	8820		NA
## 3	98.99	NA	8310		98.2
## 4	75.49	NA	NA		78.4
## 5	48.38	70.1	5230		93.1
## 6	196.41	99.0	17900		91.1
## 7	134.92	97.8	17130		NA
## 9	108.34	NA	38110		96.9
## 10	154.78	NA	42050		NA
## 12	86.06	NA	NA		NA
## 13	127.96	91.9	NA		NA
## 14	56.06	56.8	1940		NA
## 15	127.01	NA	NA		NA
## 17	116.61	NA	39190		98.9
## 18	69.96	NA	6090		NA
## 19	85.33	42.4	1620		NA
## 20	65.58	NA	5570		88.3
## 21	82.82	NA	4890		91.2
## 22	84.52	97.9	9190		86.5

## 23	142.82	84.5	14550	NA
## 24	124.26	NA	11420	NA
## 25	109.17	95.2	NA	NA
## 26	140.68	NA	14160	99.3
## 27	45.27	NA	1300	60.7
## 28	22.33	67.2	610	NA
## 29	96.17	NA	2230	96.4
## 30	52.35	NA	2330	99.6
## 31	79.73	NA	39660	NA
## 32	79.19	84.3	3980	94.6
## 33	40.65	56.0	810	81.3
## 34	31.80	34.5	1360	NA
## 35	129.71	NA	16330	94.3
## 36	73.19	94.3	8390	NA
## 37	98.45	93.4	9560	91.7
## 38	28.71	74.9	1110	NA
## 39	93.84	NA	3240	92.3
## 40	NA	NA	NA	97.6
## 41	92.20	96.2	11860	NA
## 42	86.06	56.2	1710	NA
## 43	116.37	98.8	18760	94.8
## 44	11.69	99.8	NA	100.0
## 45	97.71	98.3	NA	99.1
## 46	123.44	NA	24370	NA
## 47	4.09	NA	NA	NA
## 48	23.09	66.8	340	NA
## 49	128.47	NA	41900	94.8
## 50	21.32	NA	NA	NA
## 51	164.02	NA	13000	NA
## 52	87.22	89.5	9420	95.5
## 53	104.55	91.9	8510	NA
## 54	101.08	72.0	6120	NA
## 55	133.54	84.5	6640	95.2
## 56	59.15	93.9	25620	56.5
## 57	4.47	67.8	580	37.2
## 58	138.98	99.8	20850	97.7
## 59	16.67	NA	1110	84.8
## 60	83.72	NA	4610	NA
## 61	166.02	NA	37670	97.7
## 62	94.79	NA	35910	99.1
## 63	117.32	88.4	13740	NA
## 64	78.89	50.0	1750	68.2
## 65	102.31	99.7	5350	NA
## 66	132.30	NA	40230	NA
## 67	84.78	67.3	1810	NA
## 68	106.48	97.2	25100	98.8
## 69	NA	NA	10350	NA
## 70	140.38	75.2	4760	98.6
## 71	44.02	41.0	1020	85.2
## 72	56.18	54.2	1240	76.7

## 73	69.94	NA	NA	82.4
## 74	41.49	NA	1180	NA
## 75	103.97	84.8	3820	94.8
## 76	117.30	99.0	20310	97.8
## 77	106.08	NA	31020	98.8
## 78	72.00	NA	3590	NA
## 79	103.09	NA	4500	NA
## 80	74.93	NA	NA	NA
## 81	78.12	78.2	3750	NA
## 82	108.41	NA	34180	99.4
## 83	121.66	NA	27110	97.0
## 84	157.93	98.9	32400	99.6
## 85	108.12	86.6	NA	83.4
## 86	104.95	NA	35330	NA
## 87	118.20	92.6	5930	90.8
## 88	155.74	99.7	11250	NA
## 89	67.49	87.4	1710	NA
## 90	13.64	NA	3300	NA
## 91	175.09	NA	NA	NA
## 92	116.40	NA	2180	95.5
## 93	87.16	NA	2580	98.1
## 94	102.94	99.8	17700	95.0
## 95	78.65	NA	14470	93.5
## 96	56.17	89.6	2050	72.2
## 97	49.17	60.8	540	NA
## 98	155.70	89.2	NA	NA
## 99	151.30	99.7	19640	95.6
## 100	148.27	NA	64260	93.6
## 101	40.65	NA	950	NA
## 102	25.69	74.8	870	NA
## 103	127.04	93.1	15650	NA
## 104	165.72	NA	7430	96.5
## 105	68.32	31.1	1040	70.6
## 106	124.86	NA	NA	93.3
## 107	NA	NA	NA	NA
## 108	93.60	58.0	2400	72.8
## 109	99.04	88.5	14330	NA
## 110	82.38	93.1	15390	99.2
## 111	NA	NA	3580	NA
## 112	89.73	NA	NA	NA
## 113	105.08	97.4	4290	99.6
## 114	NA	98.4	13700	NA
## 115	113.26	NA	4880	NA
## 116	32.83	56.1	970	94.6
## 117	2.57	92.3	NA	NA
## 118	96.39	88.8	6560	83.8
## 120	43.81	60.3	1260	NA
## 121	NA	NA	43140	NA
## 122	109.19	NA	NA	99.3
## 123	82.15	NA	3730	93.2

## 124	29.52	NA	720	64.2
## 125	58.58	61.3	2290	60.1
## 126	NA	NA	NA	NA
## 127	115.62	NA	61460	99.1
## 128	168.97	NA	NA	NA
## 129	61.61	NA	2870	81.3
## 130	74.94	NA	11080	NA
## 131	188.60	94.1	14510	99.1
## 132	34.22	60.6	2570	NA
## 133	99.40	93.9	5390	84.4
## 134	110.41	NA	9440	97.8
## 135	99.30	NA	4140	NA
## 136	130.97	99.5	20430	96.9
## 137	115.39	95.2	24440	99.1
## 138	123.11	96.3	86440	95.7
## 139	108.50	NA	30370	99.3
## 141	109.16	97.7	15120	87.9
## 142	179.31	99.6	20560	NA
## 143	40.63	71.1	1270	NA
## 144	NA	NA	16470	85.8
## 145	123.00	NA	11220	90.2
## 146	120.52	NA	10440	NA
## 147	NA	98.8	4270	93.2
## 148	111.75	NA	NA	NA
## 149	68.26	89.2	2080	NA
## 150	191.24	86.6	24700	96.7
## 151	73.25	NA	1940	75.9
## 152	125.39	97.9	11540	94.7
## 153	145.71	91.8	25140	NA
## 154	35.63	42.1	840	NA
## 155	150.24	95.9	59380	NA
## 156	109.35	NA	22130	NA
## 157	106.56	99.7	26510	97.7
## 158	49.77	NA	2350	87.7
## 159	6.85	NA	NA	NA
## 160	126.83	NA	10710	NA
## 161	NA	NA	NA	NA
## 162	113.22	97.7	31400	99.7
## 163	87.05	91.2	5520	93.9
## 164	56.14	71.1	2120	NA
## 165	178.88	94.7	NA	NA
## 166	63.70	87.4	5930	NA
## 167	118.57	NA	42200	99.7
## 168	131.43	NA	52570	98.9
## 169	63.17	83.4	NA	NA
## 170	90.64	99.7	2300	99.5
## 171	111.63	NA	8360	NA
## 172	107.24	97.3	11090	97.3
## 173	53.23	58.3	NA	86.2
## 174	50.45	NA	1040	NA

## 175	52.63	NA	5000	NA
## 177	116.93	NA	9030	NA
## 178	88.70	NA	16940	99.5
## 179	68.77	99.6	8690	NA
## 180	21.63	NA	NA	NA
## 181	48.38	73.2	1310	89.7
## 183	148.62	NA	47890	NA
## 184	130.75	NA	36010	99.8
## 185	55.53	73.2	1500	NA
## 186	92.72	NA	48820	95.4
## 187	140.75	98.1	14640	NA
## 188	91.65	99.4	3420	93.3
## 189	55.76	82.6	4330	NA
## 190	97.78	NA	12430	94.7
## 191	143.39	93.2	3250	NA
## 192	47.05	63.9	2170	85.5
## 193	60.59	71.2	1490	91.4
## 194	72.13	92.2	NA	NA
##	PrimarySchoolEnrollmentFemale			
## 1		NA		
## 2		NA		
## 3		96.4		
## 4		79.4		
## 5		78.2		
## 6		84.5		
## 7		NA		
## 9		97.5		
## 10		NA		
## 12		NA		
## 13		NA		
## 14		NA		
## 15		NA		
## 17		99.2		
## 18		NA		
## 19		NA		
## 20		91.5		
## 21		91.5		
## 22		88.4		
## 23		NA		
## 24		NA		
## 25		NA		
## 26		99.7		
## 27		55.9		
## 28		NA		
## 29		95.4		
## 30		87.4		
## 31		NA		
## 32		92.4		
## 33		60.6		
## 34		NA		

## 35	94.4
## 36	NA
## 37	91.3
## 38	NA
## 39	89.3
## 40	99.3
## 41	NA
## 42	NA
## 43	97.0
## 44	99.7
## 45	99.5
## 46	NA
## 47	NA
## 48	NA
## 49	96.9
## 50	NA
## 51	NA
## 52	90.4
## 53	NA
## 54	NA
## 55	95.5
## 56	56.0
## 57	32.5
## 58	97.0
## 59	79.5
## 60	NA
## 61	97.9
## 62	99.3
## 63	NA
## 64	70.4
## 65	NA
## 66	NA
## 67	NA
## 68	99.3
## 69	NA
## 70	97.5
## 71	72.1
## 72	73.3
## 73	85.9
## 74	NA
## 75	97.0
## 76	98.3
## 77	99.2
## 78	NA
## 79	NA
## 80	NA
## 81	NA
## 82	100.0
## 83	97.8
## 84	98.5

## 85	81.4
## 86	NA
## 87	90.7
## 88	NA
## 89	NA
## 90	NA
## 91	NA
## 92	95.1
## 93	95.4
## 94	96.8
## 95	92.9
## 96	75.3
## 97	NA
## 98	NA
## 99	95.8
## 100	95.7
## 101	NA
## 102	NA
## 103	NA
## 104	96.5
## 105	60.8
## 106	94.3
## 107	NA
## 108	76.0
## 109	NA
## 110	99.9
## 111	NA
## 112	NA
## 113	98.5
## 114	NA
## 115	NA
## 116	89.4
## 117	NA
## 118	88.5
## 120	NA
## 121	NA
## 122	99.6
## 123	94.5
## 124	52.0
## 125	54.8
## 126	NA
## 127	99.2
## 128	NA
## 129	66.5
## 130	NA
## 131	98.2
## 132	NA
## 133	83.9
## 134	98.5
## 135	NA

## 136	96.7
## 137	99.7
## 138	96.6
## 139	98.4
## 141	87.3
## 142	NA
## 143	NA
## 144	86.2
## 145	89.2
## 146	NA
## 147	97.1
## 148	NA
## 149	NA
## 150	96.5
## 151	80.2
## 152	94.4
## 153	NA
## 154	NA
## 155	NA
## 156	NA
## 157	97.3
## 158	87.3
## 159	NA
## 160	NA
## 161	NA
## 162	99.8
## 163	94.4
## 164	NA
## 165	NA
## 166	NA
## 167	99.0
## 168	99.5
## 169	NA
## 170	96.0
## 171	NA
## 172	99.2
## 173	85.6
## 174	NA
## 175	NA
## 177	NA
## 178	98.3
## 179	NA
## 180	NA
## 181	92.3
## 183	NA
## 184	99.6
## 185	NA
## 186	96.1
## 187	NA
## 188	91.0

```
## 189          NA
## 190        95.1
## 191          NA
## 192        70.5
## 193        93.9
## 194          NA
```

#TASK4:Filtering data using is.na()

#We want to find all the rows having null values from the LiteracyRate column

```
null_values_data <- health_data[which(is.na(health_data$LiteracyRate)),]
null_values_data
```

##	Country	Region	Population
## 1	Afghanistan	Eastern Mediterranean	29825
## 2	Albania	Europe	3162
## 3	Algeria	Africa	38482
## 4	Andorra	Europe	78
## 9	Australia	Western Pacific	23050
## 10	Austria	Europe	8464
## 11	Azerbaijan	Europe	9309
## 12	Bahamas	Americas	372
## 15	Barbados	Americas	283
## 16	Belarus	Europe	9405
## 17	Belgium	Europe	11060
## 18	Belize	Americas	324
## 20	Bhutan	South-East Asia	742
## 21	Bolivia (Plurinational State of)	Americas	10496
## 24	Brazil	Americas	199000
## 26	Bulgaria	Europe	7278
## 27	Burkina Faso	Africa	16460
## 29	Cambodia	Western Pacific	14865
## 30	Cameroon	Africa	21700
## 31	Canada	Americas	34838
## 35	Chile	Americas	17465
## 39	Congo	Africa	4337
## 40	Cook Islands	Western Pacific	21
## 46	Czech Republic	Europe	10660
## 47	Democratic People's Republic of Korea	South-East Asia	24763
## 49	Denmark	Europe	5598
## 50	Djibouti	Eastern Mediterranean	860
## 51	Dominica	Americas	72
## 59	Ethiopia	Africa	91729
## 60	Fiji	Western Pacific	875
## 61	Finland	Europe	5408
## 62	France	Europe	63937
## 66	Germany	Europe	82800
## 69	Grenada	Americas	105
## 73	Guyana	Americas	795
## 74	Haiti	Americas	10174
## 77	Iceland	Europe	326

## 78	India	South-East Asia	1240000
## 79	Indonesia	South-East Asia	247000
## 80	Iran (Islamic Republic of)	Eastern Mediterranean	76424
## 82	Ireland	Europe	4576
## 83	Israel	Europe	7644
## 86	Japan	Western Pacific	127000
## 90	Kiribati	Western Pacific	101
## 91	Kuwait	Eastern Mediterranean	3250
## 92	Kyrgyzstan	Europe	5474
## 93	Lao People's Democratic Republic	Western Pacific	6646
## 95	Lebanon	Eastern Mediterranean	4647
## 100	Luxembourg	Europe	524
## 101	Madagascar	Africa	22294
## 104	Maldives	South-East Asia	338
## 106	Malta	Europe	428
## 107	Marshall Islands	Western Pacific	53
## 111	Micronesia (Federated States of)	Western Pacific	103
## 112	Monaco	Europe	38
## 115	Morocco	Eastern Mediterranean	32521
## 119	Nauru	Western Pacific	10
## 121	Netherlands	Europe	16714
## 122	New Zealand	Western Pacific	4460
## 123	Nicaragua	Americas	5992
## 124	Niger	Africa	17157
## 126	Niue	Western Pacific	1
## 127	Norway	Europe	4994
## 128	Oman	Eastern Mediterranean	3314
## 129	Pakistan	Eastern Mediterranean	179000
## 130	Palau	Western Pacific	21
## 134	Peru	Americas	29988
## 135	Philippines	Western Pacific	96707
## 139	Republic of Korea	Western Pacific	49003
## 144	Saint Kitts and Nevis	Americas	54
## 145	Saint Lucia	Americas	181
## 146	Saint Vincent and the Grenadines	Americas	109
## 148	San Marino	Europe	31
## 151	Senegal	Africa	13726
## 156	Slovakia	Europe	5446
## 158	Solomon Islands	Western Pacific	550
## 159	Somalia	Eastern Mediterranean	10195
## 160	South Africa	Africa	52386
## 161	South Sudan	Eastern Mediterranean	10838
## 167	Sweden	Europe	9511
## 168	Switzerland	Europe	7997
## 171	Thailand	South-East Asia	66785
## 174	Togo	Africa	6643
## 175	Tonga	Western Pacific	105
## 177	Tunisia	Eastern Mediterranean	10875
## 178	Turkey	Europe	73997
## 180	Tuvalu	Western Pacific	10

## 183	United Arab Emirates Eastern Mediterranean					9206
## 184	United Kingdom Europe					62783
## 186	United States of America Americas					318000
## 190	Venezuela (Bolivarian Republic of) Americas					29955
##	Under15	Over60	FertilityRate	LifeExpectancy	ChildMortality	
## 1	47.42	3.82	5.40	60	98.5	
## 2	21.33	14.93	1.75	74	16.7	
## 3	27.42	7.17	2.83	73	20.0	
## 4	15.20	22.86	NA	82	3.2	
## 9	18.95	19.46	1.89	82	4.9	
## 10	14.51	23.52	1.44	81	4.0	
## 11	22.25	8.24	1.96	71	35.2	
## 12	21.62	11.24	1.90	75	16.9	
## 15	18.99	15.78	1.84	78	18.4	
## 16	15.10	19.31	1.47	71	5.2	
## 17	16.88	23.81	1.85	80	4.2	
## 18	34.40	5.74	2.76	74	18.3	
## 20	28.53	6.90	2.32	67	44.6	
## 21	35.23	7.28	3.31	67	41.4	
## 24	24.56	10.81	1.82	74	14.4	
## 26	13.53	26.11	1.51	74	12.1	
## 27	45.66	3.88	5.78	56	102.4	
## 29	31.23	7.67	2.93	65	39.7	
## 30	43.08	4.89	4.94	53	94.9	
## 31	16.37	20.82	1.66	82	5.3	
## 35	21.38	13.80	1.84	79	9.1	
## 39	42.37	5.13	5.05	58	96.0	
## 40	30.61	9.07	NA	77	10.6	
## 46	14.56	23.23	1.53	78	3.8	
## 47	21.98	12.74	2.00	69	28.8	
## 49	17.66	23.90	1.88	79	3.7	
## 50	33.72	5.96	3.53	58	80.9	
## 51	25.96	12.35	NA	74	12.6	
## 59	43.29	5.17	4.77	60	68.3	
## 60	28.88	8.38	2.64	70	22.4	
## 61	16.42	25.90	1.85	81	2.9	
## 62	18.26	23.82	1.98	82	4.1	
## 66	13.17	26.72	1.40	81	4.1	
## 69	26.96	9.72	2.22	74	13.5	
## 73	36.77	5.18	2.64	63	35.2	
## 74	35.35	6.70	3.28	63	75.6	
## 77	20.71	17.62	2.11	82	2.3	
## 78	29.43	8.10	2.53	65	56.3	
## 79	29.27	7.86	2.40	69	31.0	
## 80	23.68	7.82	1.91	73	17.6	
## 82	21.54	16.59	2.00	81	4.0	
## 83	27.53	15.15	2.92	82	4.2	
## 86	13.12	31.92	1.39	83	3.0	
## 90	30.10	8.84	3.01	67	59.9	
## 91	24.90	3.80	2.65	80	11.0	

## 92	30.21	6.34	3.03	69	26.6
## 93	35.61	5.76	3.20	68	71.8
## 95	21.64	12.03	1.50	74	9.3
## 100	17.46	19.15	1.65	82	2.2
## 101	42.72	4.45	4.59	66	58.2
## 104	29.03	6.65	2.31	77	10.5
## 106	14.98	22.87	1.37	80	6.8
## 107	30.10	8.84	NA	60	37.9
## 111	35.81	6.67	3.40	69	38.5
## 112	18.26	23.82	NA	82	3.8
## 115	27.85	7.61	2.65	72	31.1
## 119	30.10	8.84	NA	71	37.1
## 121	17.21	23.02	1.76	81	4.1
## 122	20.26	19.01	2.10	81	5.7
## 123	33.37	6.59	2.59	73	24.4
## 124	49.99	4.26	7.58	56	113.5
## 126	30.61	9.07	NA	72	25.1
## 127	18.64	21.41	1.93	81	2.8
## 128	24.19	3.99	2.90	72	11.6
## 129	34.31	6.44	3.35	67	85.9
## 130	30.10	8.84	NA	72	20.8
## 134	29.18	9.12	2.48	77	18.2
## 135	34.53	6.21	3.11	69	29.8
## 139	15.25	16.58	1.29	81	3.8
## 144	25.96	12.35	NA	74	9.2
## 145	24.31	12.13	1.96	75	17.5
## 146	25.70	9.92	2.05	74	23.4
## 148	14.04	26.97	NA	83	3.3
## 151	43.54	4.57	5.02	61	59.6
## 156	15.00	18.60	1.37	76	7.5
## 158	40.37	5.10	4.17	70	31.1
## 159	47.35	4.46	6.77	50	147.4
## 160	29.53	8.44	2.44	58	44.6
## 161	42.28	5.26	5.10	54	104.0
## 167	16.71	25.32	1.93	82	2.9
## 168	14.79	23.25	1.51	83	4.3
## 171	18.47	13.96	1.43	74	13.2
## 174	41.89	4.44	4.75	56	95.5
## 175	37.33	7.96	3.86	72	12.8
## 177	23.22	10.49	2.04	76	16.1
## 178	26.00	10.56	2.08	76	14.2
## 180	30.61	9.07	NA	64	29.7
## 183	14.41	0.81	1.84	76	8.4
## 184	17.54	23.06	1.90	80	4.8
## 186	19.63	19.31	2.00	79	7.1
## 190	28.84	9.17	2.44	75	15.3
##	CellularSubscribers	LiteracyRate	GNI	PrimarySchoolEnrollmentMale	
## 1	54.26	NA	1140		NA
## 2	96.39	NA	8820		NA
## 3	98.99	NA	8310		98.2

## 4	75.49	NA NA	78.4
## 9	108.34	NA 38110	96.9
## 10	154.78	NA 42050	NA
## 11	108.75	NA 8960	85.3
## 12	86.06	NA NA	NA
## 15	127.01	NA NA	NA
## 16	111.88	NA 14460	NA
## 17	116.61	NA 39190	98.9
## 18	69.96	NA 6090	NA
## 20	65.58	NA 5570	88.3
## 21	82.82	NA 4890	91.2
## 24	124.26	NA 11420	NA
## 26	140.68	NA 14160	99.3
## 27	45.27	NA 1300	60.7
## 29	96.17	NA 2230	96.4
## 30	52.35	NA 2330	99.6
## 31	79.73	NA 39660	NA
## 35	129.71	NA 16330	94.3
## 39	93.84	NA 3240	92.3
## 40	NA	NA NA	97.6
## 46	123.44	NA 24370	NA
## 47	4.09	NA NA	NA
## 49	128.47	NA 41900	94.8
## 50	21.32	NA NA	NA
## 51	164.02	NA 13000	NA
## 59	16.67	NA 1110	84.8
## 60	83.72	NA 4610	NA
## 61	166.02	NA 37670	97.7
## 62	94.79	NA 35910	99.1
## 66	132.30	NA 40230	NA
## 69	NA	NA 10350	NA
## 73	69.94	NA NA	82.4
## 74	41.49	NA 1180	NA
## 77	106.08	NA 31020	98.8
## 78	72.00	NA 3590	NA
## 79	103.09	NA 4500	NA
## 80	74.93	NA NA	NA
## 82	108.41	NA 34180	99.4
## 83	121.66	NA 27110	97.0
## 86	104.95	NA 35330	NA
## 90	13.64	NA 3300	NA
## 91	175.09	NA NA	NA
## 92	116.40	NA 2180	95.5
## 93	87.16	NA 2580	98.1
## 95	78.65	NA 14470	93.5
## 100	148.27	NA 64260	93.6
## 101	40.65	NA 950	NA
## 104	165.72	NA 7430	96.5
## 106	124.86	NA NA	93.3
## 107	NA	NA NA	NA

## 111	NA	NA 3580	NA
## 112	89.73	NA NA	NA
## 115	113.26	NA 4880	NA
## 119	65.00	NA NA	NA
## 121	NA	NA 43140	NA
## 122	109.19	NA NA	99.3
## 123	82.15	NA 3730	93.2
## 124	29.52	NA 720	64.2
## 126	NA	NA NA	NA
## 127	115.62	NA 61460	99.1
## 128	168.97	NA NA	NA
## 129	61.61	NA 2870	81.3
## 130	74.94	NA 11080	NA
## 134	110.41	NA 9440	97.8
## 135	99.30	NA 4140	NA
## 139	108.50	NA 30370	99.3
## 144	NA	NA 16470	85.8
## 145	123.00	NA 11220	90.2
## 146	120.52	NA 10440	NA
## 148	111.75	NA NA	NA
## 151	73.25	NA 1940	75.9
## 156	109.35	NA 22130	NA
## 158	49.77	NA 2350	87.7
## 159	6.85	NA NA	NA
## 160	126.83	NA 10710	NA
## 161	NA	NA NA	NA
## 167	118.57	NA 42200	99.7
## 168	131.43	NA 52570	98.9
## 171	111.63	NA 8360	NA
## 174	50.45	NA 1040	NA
## 175	52.63	NA 5000	NA
## 177	116.93	NA 9030	NA
## 178	88.70	NA 16940	99.5
## 180	21.63	NA NA	NA
## 183	148.62	NA 47890	NA
## 184	130.75	NA 36010	99.8
## 186	92.72	NA 48820	95.4
## 190	97.78	NA 12430	94.7

##	PrimarySchoolEnrollmentFemale
## 1	NA
## 2	NA
## 3	96.4
## 4	79.4
## 9	97.5
## 10	NA
## 11	84.1
## 12	NA
## 15	NA
## 16	NA
## 17	99.2

## 18	NA
## 20	91.5
## 21	91.5
## 24	NA
## 26	99.7
## 27	55.9
## 29	95.4
## 30	87.4
## 31	NA
## 35	94.4
## 39	89.3
## 40	99.3
## 46	NA
## 47	NA
## 49	96.9
## 50	NA
## 51	NA
## 59	79.5
## 60	NA
## 61	97.9
## 62	99.3
## 66	NA
## 69	NA
## 73	85.9
## 74	NA
## 77	99.2
## 78	NA
## 79	NA
## 80	NA
## 82	100.0
## 83	97.8
## 86	NA
## 90	NA
## 91	NA
## 92	95.1
## 93	95.4
## 95	92.9
## 100	95.7
## 101	NA
## 104	96.5
## 106	94.3
## 107	NA
## 111	NA
## 112	NA
## 115	NA
## 119	NA
## 121	NA
## 122	99.6
## 123	94.5
## 124	52.0

```
## 126          NA
## 127        99.2
## 128          NA
## 129        66.5
## 130          NA
## 134        98.5
## 135          NA
## 139        98.4
## 144        86.2
## 145        89.2
## 146          NA
## 148          NA
## 151        80.2
## 156          NA
## 158        87.3
## 159          NA
## 160          NA
## 161          NA
## 167        99.0
## 168        99.5
## 171          NA
## 174          NA
## 175          NA
## 177          NA
## 178        98.3
## 180          NA
## 183          NA
## 184        99.6
## 186        96.1
## 190        95.1
```

#TASK5:Take backup of data

```
backup_health_data <- health_data
```

#TASK6:Handling null values task

#TASK 6.1:Removing records from data

```
health_data <- health_data[complete.cases(health_data),]
```

#TASK 6.2:Replace with mean

```
health_data$LiteracyRate[is.na(health_data$LiteracyRate)] <-
mean(health_data$LiteracyRate, na.rm = TRUE)
```

#TASK 6.3:Replace with median

```
health_data$LiteracyRate[is.na(health_data$LiteracyRate)] <-
median(health_data$LiteracyRate, na.rm = TRUE)
```

#TASK6.4:Check for null values (Check finally)

```
null_values_after_handling <- sum(is.na(health_data))
null_values_after_handling
```

```
## [1] 0

# checking data type of data
str(health_data)

## 'data.frame':    50 obs. of  13 variables:
##  $ Country          : chr  "Angola" "Antigua and Barbuda"
##    "Bosnia and Herzegovina" "Cape Verde" ...
##  $ Region           : chr  "Africa" "Americas" "Europe"
##    "Africa" ...
##  $ Population        : int   20821 89 3834 494 4525 47704 4307
##    10277 6297 736 ...
##  $ Under15           : num   47.6 26 16.4 30.2 40.1 ...
##  $ Over60            : num   3.84 12.35 20.52 7.05 5.74 ...
##  $ FertilityRate      : num   6.1 2.12 1.26 2.38 4.54 2.35 1.48
##    2.55 2.24 5.04 ...
##  $ LifeExpectancy    : int   51 75 76 72 48 78 77 73 72 54 ...
##  $ ChildMortality    : num   163.5 9.9 6.7 22.2 128.6 ...
##  $ CellularSubscribers : num   48.4 196.4 84.5 79.2 40.6 ...
##  $ LiteracyRate       : num   70.1 99 97.9 84.3 56 93.4 98.8 89.5
##    84.5 93.9 ...
##  $ GNI                : num   5230 17900 9190 3980 810 ...
##  $ PrimarySchoolEnrollmentMale : num   93.1 91.1 86.5 94.6 81.3 91.7 94.8
##    95.5 95.2 56.5 ...
##  $ PrimarySchoolEnrollmentFemale: num   78.2 84.5 88.4 92.4 60.6 91.3 97
##    90.4 95.5 56 ...
```

Analysis and Visualization of the data

```
#TASK 7:Show correlation of data
numeric_health_data <-
health_data[,c("Population", 'Under15', 'Over60', 'FertilityRate', 'LifeExpectancy', 'ChildMortality', 'CellularSubscribers', 'LiteracyRate', 'GNI', 'PrimarySchoolEnrollmentMale', 'PrimarySchoolEnrollmentFemale')]
numeric_health_data
```

	Population	Under15	Over60	FertilityRate	LifeExpectancy	ChildMortality
## 5	20821	47.58	3.84	6.10	51	163.5
## 6	89	25.96	12.35	2.12	75	9.9
## 22	3834	16.35	20.52	1.26	76	6.7
## 32	494	30.17	7.05	2.38	72	22.2
## 33	4525	40.07	5.74	4.54	48	128.6
## 37	47704	28.03	9.19	2.35	78	17.6
## 43	4307	14.98	24.69	1.48	77	4.7
## 52	10277	30.53	8.97	2.55	73	27.1
## 55	6297	30.62	9.64	2.24	72	15.9
## 56	736	38.95	4.53	5.04	54	100.3
## 57	6131	43.10	3.73	4.88	61	51.8
## 58	1291	15.69	23.92	1.62	76	3.6
## 64	1791	45.90	3.72	5.79	58	72.9
## 68	11125	14.60	25.41	1.51	81	4.8

## 70	15083	40.80	6.56	3.91	69	32.0
## 71	11451	42.46	5.03	5.09	55	101.2
## 72	1664	41.55	5.06	5.05	50	129.1
## 75	7936	35.72	6.41	3.10	74	22.9
## 76	9976	14.62	23.41	1.38	75	6.2
## 84	60885	14.04	26.97	1.45	82	3.8
## 87	7009	34.13	5.30	3.39	74	19.1
## 94	2060	14.57	24.24	1.57	74	8.7
## 96	2052	36.75	6.31	3.15	50	99.6
## 99	3028	15.13	20.57	1.49	74	5.4
## 105	14854	47.14	4.29	6.85	51	128.0
## 108	3796	40.22	4.94	4.78	59	84.0
## 110	121000	29.02	9.18	2.25	75	16.2
## 113	2796	27.05	5.80	2.45	68	27.5
## 116	25203	45.38	5.01	5.34	53	89.7
## 118	2259	36.59	5.38	3.17	65	38.7
## 125	169000	44.23	4.49	6.02	53	123.7
## 131	3802	28.65	10.13	2.52	77	18.5
## 133	6687	32.78	8.01	2.93	75	22.0
## 136	38211	14.91	20.48	1.39	76	5.0
## 137	10604	14.92	24.39	1.33	80	3.6
## 138	2051	13.28	1.73	2.06	82	7.4
## 140	3514	16.52	16.72	1.47	71	17.6
## 141	21755	15.05	20.66	1.39	74	12.2
## 150	28288	29.69	4.59	2.76	76	8.6
## 152	9553	16.45	20.52	1.37	74	6.6
## 157	2068	14.16	23.16	1.49	80	3.1
## 162	46755	15.20	22.86	1.47	82	4.5
## 163	21098	25.15	12.40	2.35	75	9.6
## 170	8009	35.75	4.80	3.81	68	58.3
## 172	2106	16.89	17.56	1.44	75	7.4
## 181	36346	48.54	3.72	6.06	56	68.9
## 182	45530	14.18	20.76	1.45	71	10.7
## 188	28541	28.90	6.38	2.38	68	39.6
## 192	23852	40.72	4.54	4.35	64	60.0
## 193	14075	46.73	3.95	5.77	55	88.5
##	CellularSubscribers	LiteracyRate	GNI	PrimarySchoolEnrollmentMale		
## 5	48.38	70.1	5230		93.1	
## 6	196.41	99.0	17900		91.1	
## 22	84.52	97.9	9190		86.5	
## 32	79.19	84.3	3980		94.6	
## 33	40.65	56.0	810		81.3	
## 37	98.45	93.4	9560		91.7	
## 43	116.37	98.8	18760		94.8	
## 52	87.22	89.5	9420		95.5	
## 55	133.54	84.5	6640		95.2	
## 56	59.15	93.9	25620		56.5	
## 57	4.47	67.8	580		37.2	
## 58	138.98	99.8	20850		97.7	
## 64	78.89	50.0	1750		68.2	

## 68	106.48	97.2	25100	98.8
## 70	140.38	75.2	4760	98.6
## 71	44.02	41.0	1020	85.2
## 72	56.18	54.2	1240	76.7
## 75	103.97	84.8	3820	94.8
## 76	117.30	99.0	20310	97.8
## 84	157.93	98.9	32400	99.6
## 87	118.20	92.6	5930	90.8
## 94	102.94	99.8	17700	95.0
## 96	56.17	89.6	2050	72.2
## 99	151.30	99.7	19640	95.6
## 105	68.32	31.1	1040	70.6
## 108	93.60	58.0	2400	72.8
## 110	82.38	93.1	15390	99.2
## 113	105.08	97.4	4290	99.6
## 116	32.83	56.1	970	94.6
## 118	96.39	88.8	6560	83.8
## 125	58.58	61.3	2290	60.1
## 131	188.60	94.1	14510	99.1
## 133	99.40	93.9	5390	84.4
## 136	130.97	99.5	20430	96.9
## 137	115.39	95.2	24440	99.1
## 138	123.11	96.3	86440	95.7
## 140	104.80	98.5	3640	90.1
## 141	109.16	97.7	15120	87.9
## 150	191.24	86.6	24700	96.7
## 152	125.39	97.9	11540	94.7
## 157	106.56	99.7	26510	97.7
## 162	113.22	97.7	31400	99.7
## 163	87.05	91.2	5520	93.9
## 170	90.64	99.7	2300	99.5
## 172	107.24	97.3	11090	97.3
## 181	48.38	73.2	1310	89.7
## 182	122.98	99.7	7040	90.8
## 188	91.65	99.4	3420	93.3
## 192	47.05	63.9	2170	85.5
## 193	60.59	71.2	1490	91.4

PrimarySchoolEnrollmentFemale

## 5	78.2
## 6	84.5
## 22	88.4
## 32	92.4
## 33	60.6
## 37	91.3
## 43	97.0
## 52	90.4
## 55	95.5
## 56	56.0
## 57	32.5
## 58	97.0

```
## 64          70.4
## 68          99.3
## 70          97.5
## 71          72.1
## 72          73.3
## 75          97.0
## 76          98.3
## 84          98.5
## 87          90.7
## 94          96.8
## 96          75.3
## 99          95.8
## 105         60.8
## 108         76.0
## 110         99.9
## 113         98.5
## 116         89.4
## 118         88.5
## 125         54.8
## 131         98.2
## 133         83.9
## 136         96.7
## 137         99.7
## 138         96.6
## 140         90.1
## 141         87.3
## 150         96.5
## 152         94.4
## 157         97.3
## 162         99.8
## 163         94.4
## 170         96.0
## 172         99.2
## 181         92.3
## 182         91.5
## 188         91.0
## 192         70.5
## 193         93.9
```

```
correlation_matrix <- cor(numeric_health_data)
correlation_matrix
```

```
##          Population    Under15    Over60
FertilityRate
## Population    1.00000000  0.08869404 -0.0302817
0.1255234
## Under15      0.08869404  1.00000000 -0.8641711
0.9439490
## Over60      -0.03028170 -0.86417114  1.00000000 -
0.7740002
```


## FertilityRate	0.12552336	0.94394904	-0.7740002	
1.0000000				
## LifeExpectancy	-0.04076296	-0.83262351	0.6511111	-
0.8733160				
## ChildMortality	0.11207370	0.81941553	-0.6457055	
0.8893141				
## CellularSubscribers	-0.09534305	-0.60907380	0.4613104	-
0.6469807				
## LiteracyRate	-0.07742949	-0.79799468	0.6066862	-
0.8610433				
## GNI	-0.02439652	-0.60069464	0.3413351	-
0.4686225				
## PrimarySchoolEnrollmentMale	-0.07251804	-0.53672052	0.4421306	-
0.5952212				
## PrimarySchoolEnrollmentFemale	-0.08392584	-0.61146706	0.4969884	-
0.6749036				
##	LifeExpectancy	ChildMortality		
CellularSubscribers				
## Population	-0.04076296	0.1120737		-
0.09534305				
## Under15	-0.83262351	0.8194155		-
0.60907380				
## Over60	0.65111107	-0.6457055		
0.46131039				
## FertilityRate	-0.87331599	0.8893141		-
0.64698071				
## LifeExpectancy	1.00000000	-0.9488224		
0.72078446				
## ChildMortality	-0.94882245	1.00000000		-
0.69215472				
## CellularSubscribers	0.72078446	-0.6921547		
1.00000000				
## LiteracyRate	0.78525040	-0.7942652		
0.59742815				
## GNI	0.56907332	-0.4543238		
0.45357064				
## PrimarySchoolEnrollmentMale	0.62759327	-0.5678088		
0.57739558				
## PrimarySchoolEnrollmentFemale	0.71170343	-0.6954943		
0.63801577				
##	LiteracyRate	GNI		
## Population	-0.07742949	-0.02439652		
## Under15	-0.79799468	-0.60069464		
## Over60	0.60668617	0.34133515		
## FertilityRate	-0.86104333	-0.46862251		
## LifeExpectancy	0.78525040	0.56907332		
## ChildMortality	-0.79426521	-0.45432380		
## CellularSubscribers	0.59742815	0.45357064		
## LiteracyRate	1.00000000	0.45985101		
## GNI	0.45985101	1.00000000		

```
## PrimarySchoolEnrollmentMale      0.52757378  0.30502062
## PrimarySchoolEnrollmentFemale    0.64979869  0.35253186
##                                PrimarySchoolEnrollmentMale
## Population                        -0.07251804
## Under15                          -0.53672052
## Over60                           0.44213059
## FertilityRate                    -0.59522124
## LifeExpectancy                   0.62759327
## ChildMortality                   -0.56780876
## CellularSubscribers              0.57739558
## LiteracyRate                     0.52757378
## GNI                              0.30502062
## PrimarySchoolEnrollmentMale      1.00000000
## PrimarySchoolEnrollmentFemale    0.93935208
##                                PrimarySchoolEnrollmentFemale
## Population                        -0.08392584
## Under15                          -0.61146706
## Over60                           0.49698844
## FertilityRate                    -0.67490356
## LifeExpectancy                   0.71170343
## ChildMortality                   -0.69549430
## CellularSubscribers              0.63801577
## LiteracyRate                     0.64979869
## GNI                              0.35253186
## PrimarySchoolEnrollmentMale      0.93935208
## PrimarySchoolEnrollmentFemale    1.00000000
```

#TASK 8: Show a group of histograms to depict characteristics of people in different regions

library(dplyr)

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
sum_by_region <- health_data %>%
```

```
  group_by(Region) %>%
```

```
  summarize(SumLiteracyRate = sum(LiteracyRate), sumPopulation =
```

```
sum(Population),
```

```
            sumFertilityRate = sum(FertilityRate), sumLifeExpectancy =
```

```
sum(LifeExpectancy),
```

```
            sumChildMortality = sum(ChildMortality), sumGNI = sum(GNI) )
```

```
sum_by_region
```

```
## # A tibble: 6 × 7
##   Region          SumLiteracyRate sumPopulation sumFertilityRate
sumLifeExpectancy
##   <chr>              <dbl>          <int>          <dbl>
<int>
## 1 Africa            1047.          315198          80.0
891
## 2 Americas           808.          218875          24.0
668
## 3 Eastern Medi...    339.           61200          12.6
296
## 4 Europe             1873.         313152          30.8
1434
## 5 South-East A...    91.2           21098           2.35
75
## 6 Western Paci...    97.4           2796            2.45
68
## # i 2 more variables: sumChildMortality <dbl>, sumGNI <dbl>

par(mfrow = c(2,3))

barplot(height = sum_by_region$SumLiteracyRate, names.arg =
sum_by_region$Region,
        main = "Sum of Literacy Rate by Region",
        xlab = "Region", ylab = "Literacy Rate", col = "pink")

barplot(height = sum_by_region$sumPopulation, names.arg =
sum_by_region$Region,
        main = "Sum of Population by Region",
        xlab = "Region", ylab = "Population", col = "blue")

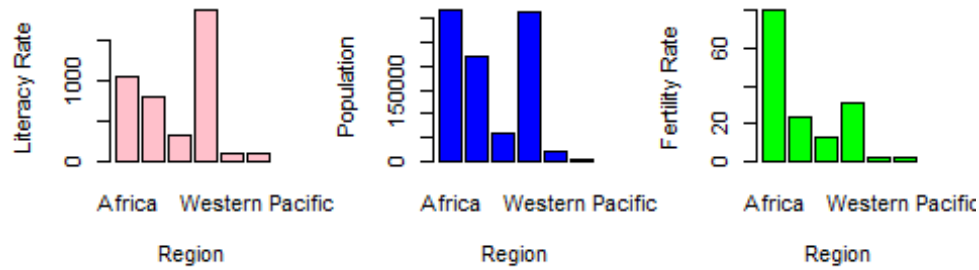
barplot(height = sum_by_region$sumFertilityRate, names.arg =
sum_by_region$Region,
        main = "Sum of Fertility Rate by Region",
        xlab = "Region", ylab = "Fertility Rate", col = "green")

barplot(height = sum_by_region$sumLifeExpectancy, names.arg =
sum_by_region$Region,
        main = "Sum of Life Expectancy by Region",
        xlab = "Region", ylab = "Life Expectancy", col = "yellow")

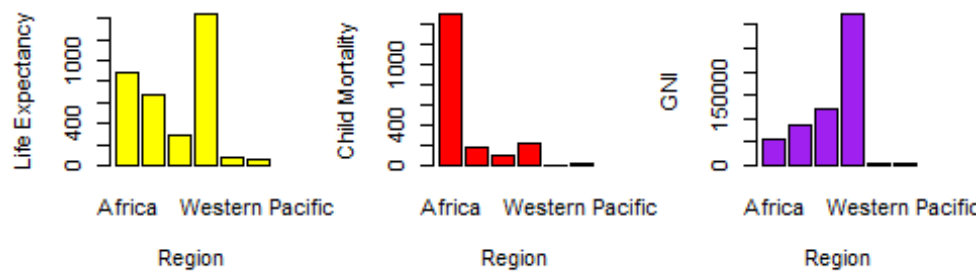
barplot(height = sum_by_region$sumChildMortality, names.arg =
sum_by_region$Region,
        main = "Sum of Child Mortality by Region",
        xlab = "Region", ylab = "Child Mortality", col = "red")

barplot(height = sum_by_region$sumGNI, names.arg = sum_by_region$Region,
        main = "Sum of GNI by Region",
        xlab = "Region", ylab = "GNI", col = "purple")
```

um of Literacy Rate by ReSum of Population by ReSum of Fertility Rate by Re



m of Life Expectancy by Sum of Child Mortality by R Sum of GNI by Region

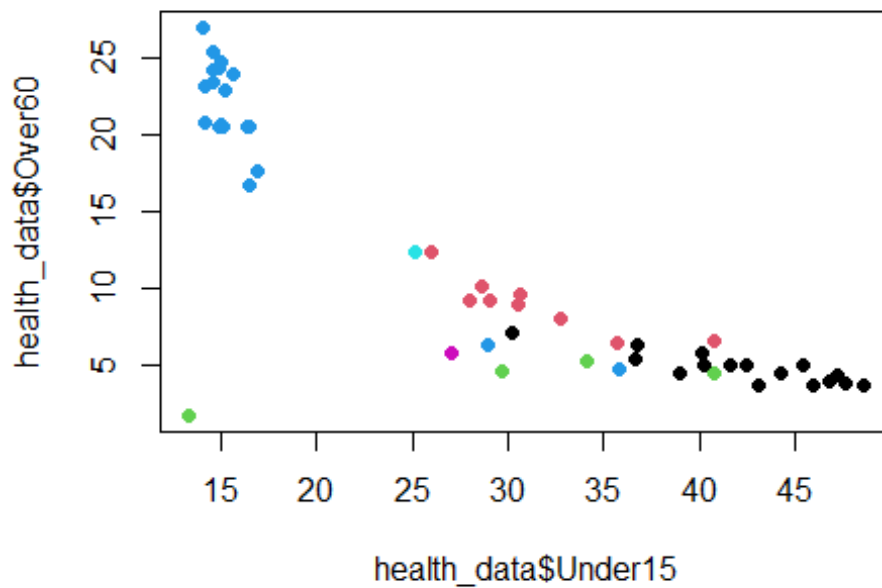


```
par(mfrow = c(1, 1))
```

#TASK 9: Show a scatter plot Under15 vs Over60, with colouring according to region

```
health_data$Region <- as.factor(health_data$Region)
health_data$Region <- as.numeric(health_data$Region)
```

```
scatter_plot <- plot(health_data$Under15, health_data$Over60, col =
health_data$Region, pch = 16)
```

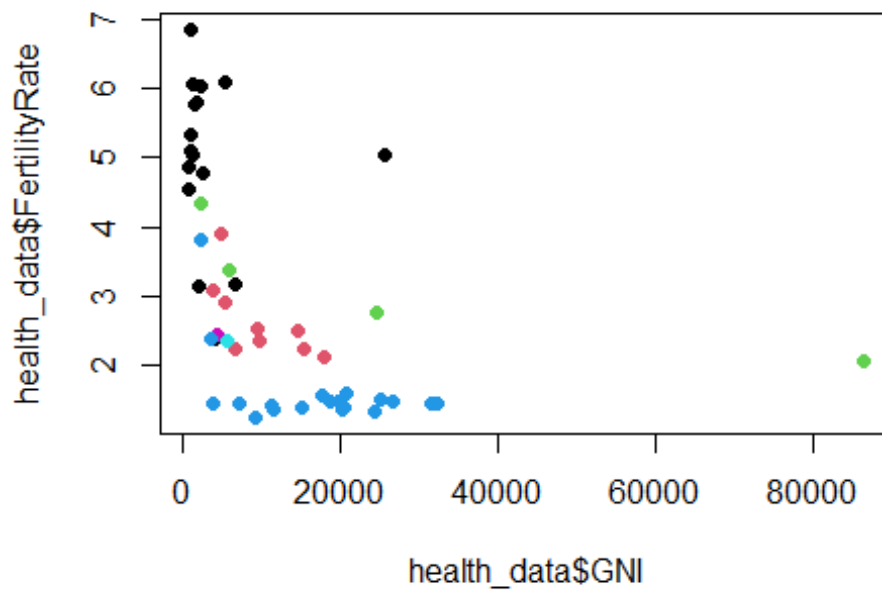


```
scatter_plot
```

```
## NULL
```

```
#TASK 10:Show a scatter plot GNI vs Fertility rate, with coloring according  
to region
```

```
scatter_plot <- plot(health_data$GNI, health_data$FertilityRate, col =  
health_data$Region, pch = 16)
```



```
scatter_plot
```

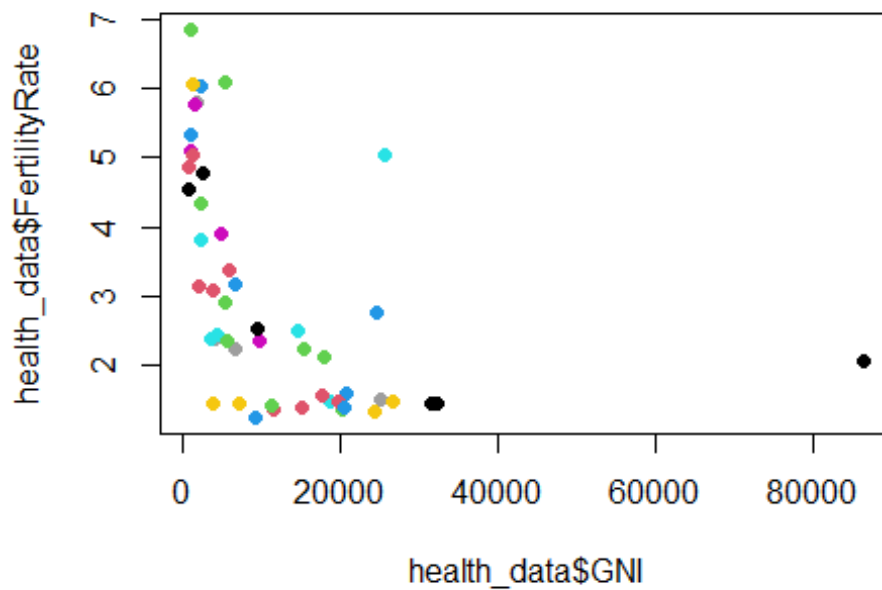
```
## NULL
```

```
#TASK 11:Show a scatter plot GNI vs Fertility rate, with coloring according  
to life expectancy
```

```
health_data$LifeExpectancy <- as.factor(health_data$LifeExpectancy)
```

```
health_data$LifeExpectancy <- as.numeric(health_data$LifeExpectancy)
```

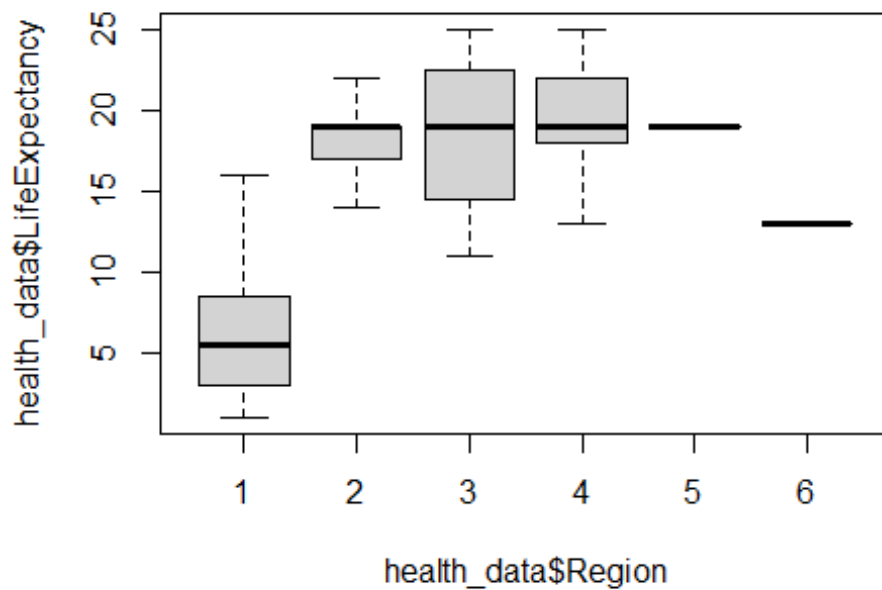
```
scatter_plot <- plot(health_data$GNI, health_data$FertilityRate, col =  
health_data$LifeExpectancy, pch = 16)
```



```
scatter_plot
```

```
## NULL
```

```
#TASK 12:Plot a box-plot showing life expectancy of different regions
boxplot_life_expectancy <- boxplot(health_data$LifeExpectancy ~
health_data$Region)
```



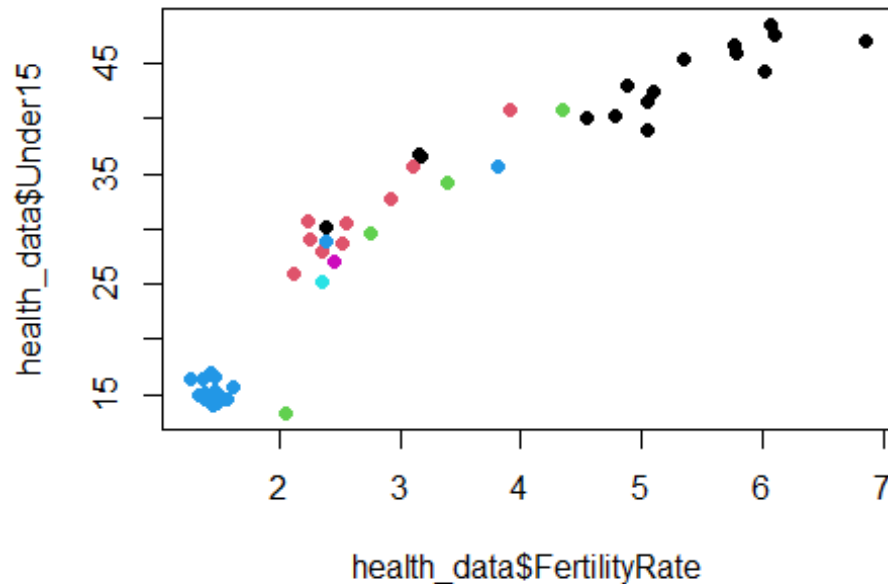
```
boxplot_life_expectancy
```

```
## $stats
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]  1.0  14 11.0  13  19  13
## [2,]  3.0  17 14.5  18  19  13
## [3,]  5.5  19 19.0  19  19  13
## [4,]  8.5  19 22.5  22  19  13
## [5,] 16.0  22 25.0  25  19  13
##
## $n
## [1] 16  9  4 19  1  1
##
## $conf
##      [,1]      [,2] [,3]      [,4] [,5] [,6]
## [1,] 3.3275 17.94667 12.68 17.55009  19  13
## [2,] 7.6725 20.05333 25.32 20.44991  19  13
##
## $out
## numeric(0)
##
## $group
## numeric(0)
##
## $names
## [1] "1" "2" "3" "4" "5" "6"
```


#TASK 13: Is the fertility rate of a region a good predictor of the percentage of the population under 15?

#(hint: show a scatter plot with coloring according to region)

```
scatter_plot_fertility_under15 <- plot(health_data$FertilityRate,  
health_data$Under15, col = health_data$Region, pch = 16)
```



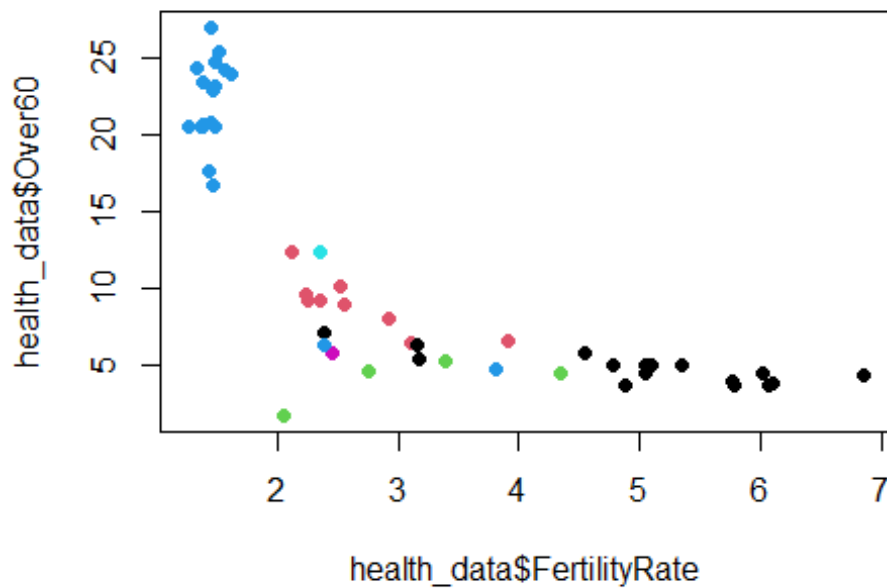
```
scatter_plot_fertility_under15
```

```
## NULL
```

#TASK 14: Is the fertility rate of a region a good predictor of the percentage of the population over 60?

#(hint: show a scatter plot with coloring according to region)

```
scatter_plot_fertility_over60 <- plot(health_data$FertilityRate,  
health_data$Over60, col = health_data$Region, pch = 16)
```



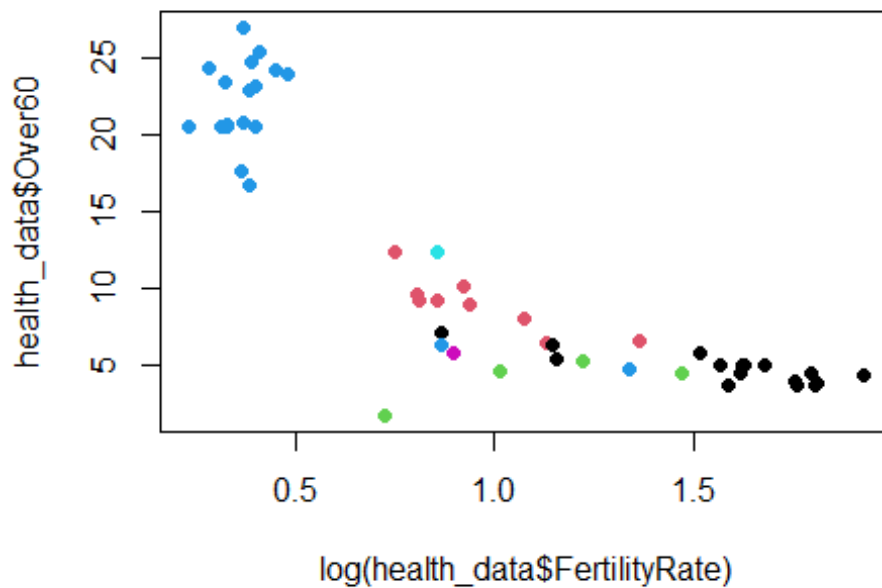
```
scatter_plot_fertility_over60
```

```
## NULL
```

```
#TASK 15:Try a log transformation showing log(Fertility rate) vs population Over60
```

```
 #(Hint: scatter plot with coloring according to region)
```

```
scatter_plot_log_fertility_over60 <- plot(log(health_data$FertilityRate),  
health_data$Over60, col = health_data$Region, pch = 16)
```

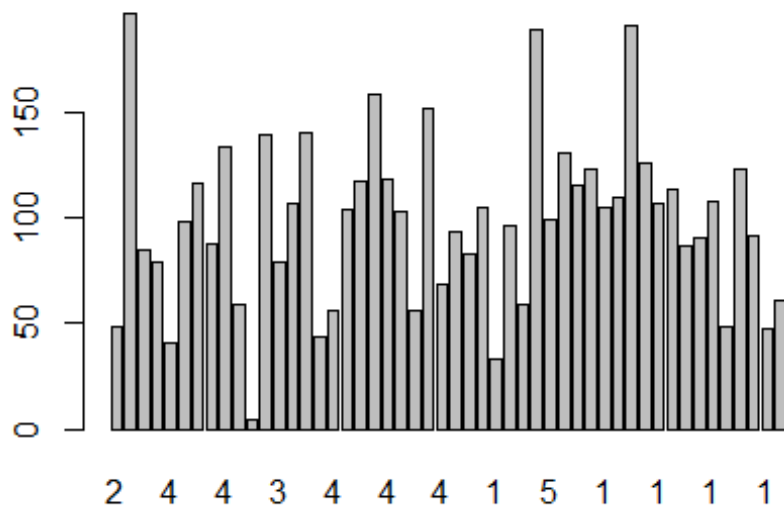


```
scatter_plot_log_fertility_over60
```

```
## NULL
```

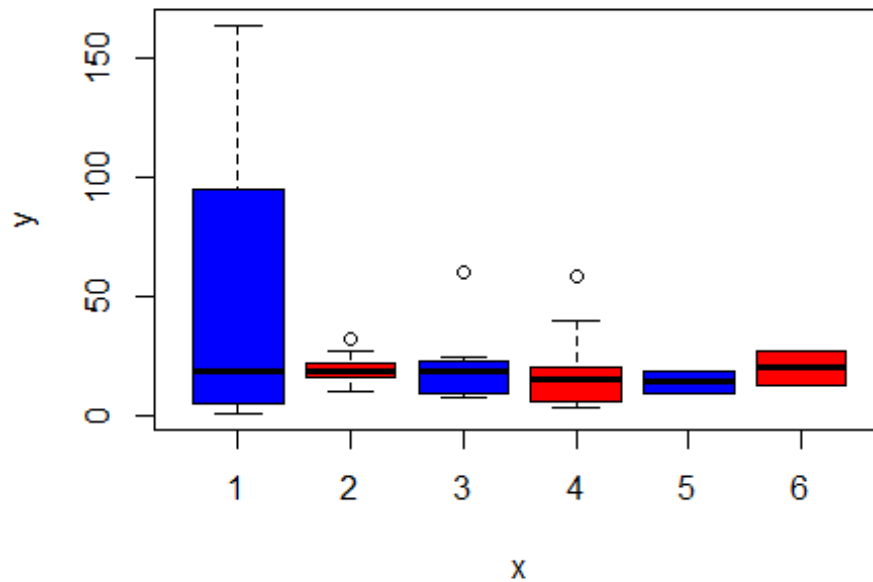
#TASK 16: Plot an ordered bar chart to show that regions team does that max cellular subscribers

```
max_subscribers_region <- barplot(health_data$CellularSubscribers,
                                  names.arg =
health_data$Region[order(health_data$CellularSubscribers,
decreasing = TRUE)])
```



max_subscribers_region

```
##      [,1]
## [1,]  0.7
## [2,]  1.9
## [3,]  3.1
## [4,]  4.3
## [5,]  5.5
## [6,]  6.7
## [7,]  7.9
## [8,]  9.1
## [9,] 10.3
## [10,] 11.5
## [11,] 12.7
## [12,] 13.9
## [13,] 15.1
## [14,] 16.3
## [15,] 17.5
## [16,] 18.7
## [17,] 19.9
## [18,] 21.1
## [19,] 22.3
## [20,] 23.5
## [21,] 24.7
## [22,] 25.9
## [23,] 27.1
## [24,] 28.3
```

stacked_plot

```
## $stats
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]  1.00  9.90  7.40  3.1  9.6 13.00
## [2,]  5.50 16.20  9.80  6.2  9.6 13.00
## [3,] 19.10 18.75 18.55 15.0 14.3 20.25
## [4,] 94.65 22.00 22.50 20.0 19.0 27.50
## [5,] 163.50 27.10 25.00 39.6 19.0 27.50
##
## $n
## [1] 32 18  8 38  2  2
##
## $conf
##      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,] -5.800235 16.59002 11.4556 11.46292  3.79805  4.050184
## [2,] 44.000235 20.90998 25.6444 18.53708 24.80195 36.449816
##
## $out
## [1] 32.0 60.0 58.3
##
## $group
## [1] 2 3 4
##
## $names
## [1] "1" "2" "3" "4" "5" "6"
```