Demography Analysis

Abhinav Kumar

2025-09-19

```
getwd()
```

```
## [1] "/Users/abhinav/Downloads/essentials/projects/R prog/Demography Analysis"
```

```
setwd("/Users/abhinav/Downloads/essentials/projects/R prog/Demography Analysis")
getwd()
```

```
## [1] "/Users/abhinav/Downloads/essentials/projects/R prog/Demography Analysis"
```

```
stats1 <- read.csv("S5-Demographic-Data.csv")
str(stats1)
```

```
## 'data.frame':    195 obs. of  5 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Birth.rate    : num  10.2 35.3 46 12.9 11 ...
##  $ Internet.users: num  78.9 5.9 19.1 57.2 88 ...
##  $ Income.Group  : chr  "High income" "Low income" "Upper middle income" "Upper middle income" ...
```

```
Countries_2012_Dataset <- c("Aruba","Afghanistan","Angola","Albania","United Arab Emirates","Argentina"
Codes_2012_Dataset <- c("ABW","AFG","AGO","ALB","ARE","ARG","ARM","ATG","AUS","AUT","AZE","BDI","BEL","
Regions_2012_Dataset <- c("The Americas","Asia","Africa","Europe","Middle East","The Americas","Asia","
```

```
head(stats1)
```

```
##           Country.Name Country.Code Birth.rate Internet.users
## 1                Aruba          ABW     10.244           78.9
## 2          Afghanistan          AFG     35.253            5.9
## 3               Angola          AGO     45.985           19.1
## 4              Albania          ALB     12.877           57.2
## 5 United Arab Emirates          ARE     11.044           88.0
## 6            Argentina          ARG     17.716           59.9
##          Income.Group
## 1         High income
## 2          Low income
## 3 Upper middle income
## 4 Upper middle income
## 5         High income
## 6         High income
```

```
summary(stats1)
```

```
##  Country.Name       Country.Code         Birth.rate    Internet.users
##  Length:195         Length:195         Min.   : 7.90   Min.   : 0.90
##  Class :character   Class :character   1st Qu.:12.12   1st Qu.:14.52
##  Mode  :character   Mode  :character   Median :19.68   Median :41.00
```

```
##                                         Mean   :21.47   Mean   :42.08
##                                         3rd Qu.:29.76   3rd Qu.:66.22
##                                         Max.   :49.66   Max.   :96.55
##  Income.Group
##  Length:195
##  Class :character
##  Mode  :character
##
##
##
```

```
stats1$Country.Name[81]
```

```
## [1] "India"
```

```
stats1$Income.Group <- factor(stats1$Income.Group)
str(stats1$Income.Group)
```

```
##  Factor w/ 4 levels "High income",..: 1 2 4 4 1 1 3 1 1 1 ...
```

```
stats1[c(44, 56), ]
```

```
##    Country.Name Country.Code Birth.rate Internet.users Income.Group
## 44       Cyprus          CYP     11.436        65.4548  High income
## 56     Ethiopia          ETH     32.925         1.9000   Low income
```

```
is.data.frame(stats1[c(44, 56), ])
```

```
## [1] TRUE
```

```
is.data.frame(stats1[,5, drop=F])
```

```
## [1] TRUE
```

```
stats1$Calc <- stats1$Birth.rate * stats1$Internet.users
str(stats1)
```

```
## 'data.frame':    195 obs. of  6 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Birth.rate    : num  10.2 35.3 46 12.9 11 ...
##  $ Internet.users: num  78.9 5.9 19.1 57.2 88 ...
##  $ Income.Group  : Factor w/ 4 levels "High income",..: 1 2 4 4 1 1 3 1 1 1 ...
##  $ Calc          : num  808 208 878 737 972 ...
```

```
stats1$Calc <- NULL
str(stats1)
```

```
## 'data.frame':    195 obs. of  5 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Birth.rate    : num  10.2 35.3 46 12.9 11 ...
##  $ Internet.users: num  78.9 5.9 19.1 57.2 88 ...
##  $ Income.Group  : Factor w/ 4 levels "High income",..: 1 2 4 4 1 1 3 1 1 1 ...
```

```
stats1[stats1$Country.Name == "India", ]
```

```
##    Country.Name Country.Code Birth.rate Internet.users        Income.Group
## 81        India          IND     20.291           15.1 Lower middle income
```

```
stats1[stats1$Birth.rate > 44.2, ]
```

```
##     Country.Name Country.Code Birth.rate Internet.users        Income.Group
## 3         Angola          AGO     45.985           19.1 Upper middle income
## 128        Niger          NER     49.661            1.7          Low income
## 168         Chad          TCD     45.745            2.3          Low income
```

```
stats1[stats1$Internet.users < 6, ]
```

```
##                 Country.Name Country.Code Birth.rate Internet.users
## 2                Afghanistan          AFG     35.253           5.90
## 12                   Burundi          BDI     44.151           1.30
## 14                     Benin          BEN     36.440           4.90
## 30  Central African Republic          CAF     34.076           3.50
## 53                   Eritrea          ERI     34.800           0.90
## 56                  Ethiopia          ETH     32.925           1.90
## 65                    Guinea          GIN     37.337           1.60
## 67             Guinea-Bissau          GNB     37.503           3.10
## 100                  Liberia          LBR     35.521           3.20
## 105                  Lesotho          LSO     28.738           5.00
## 112               Madagascar          MDG     34.686           3.00
## 116                     Mali          MLI     44.138           3.50
## 118                  Myanmar          MMR     18.119           1.60
## 121               Mozambique          MOZ     39.705           5.40
## 124                   Malawi          MWI     39.459           5.05
## 128                    Niger          NER     49.661           1.70
## 155             Sierra Leone          SLE     36.729           1.70
## 157                  Somalia          SOM     43.891           1.50
## 168                     Chad          TCD     45.745           2.30
## 169                     Togo          TGO     36.080           4.50
## 173              Timor-Leste          TLS     35.755           1.10
## 178                 Tanzania          TZA     39.518           4.40
## 193         Congo, Dem. Rep.          COD     42.394           2.20
##            Income.Group
## 2           Low income
## 12          Low income
## 14          Low income
## 30          Low income
## 53          Low income
## 56          Low income
## 65          Low income
## 67          Low income
## 100         Low income
## 105 Lower middle income
## 112         Low income
## 116         Low income
## 118 Lower middle income
## 121         Low income
## 124         Low income
## 128         Low income
## 155         Low income
## 157         Low income
## 168         Low income
## 169         Low income
```

```
## 173 Lower middle income
## 178          Low income
## 193          Low income
```

```
stats1[stats1$Income.Group == "High income", ]
```

```
##              Country.Name Country.Code Birth.rate Internet.users Income.Group
## 1                   Aruba          ABW     10.244       78.90000  High income
## 5    United Arab Emirates          ARE     11.044       88.00000  High income
## 6               Argentina          ARG     17.716       59.90000  High income
## 8     Antigua and Barbuda          ATG     16.447       63.40000  High income
## 9               Australia          AUS     13.200       83.00000  High income
## 10                Austria          AUT      9.400       80.61880  High income
## 13                Belgium          BEL     11.200       82.17020  High income
## 18                Bahrain          BHR     15.040       90.00004  High income
## 19           Bahamas, The          BHS     15.339       72.00000  High income
## 23                Bermuda          BMU     10.400       95.30000  High income
## 26               Barbados          BRB     12.188       73.00000  High income
## 27     Brunei Darussalam          BRN     16.405       64.50000  High income
## 31                 Canada          CAN     10.900       85.80000  High income
## 32            Switzerland          CHE     10.200       86.34000  High income
## 33                  Chile          CHL     13.385       66.50000  High income
## 43         Cayman Islands          CYM     12.500       74.10000  High income
## 44                 Cyprus          CYP     11.436       65.45480  High income
## 45         Czech Republic          CZE     10.200       74.11040  High income
## 46                Germany          DEU      8.500       84.17000  High income
## 48                Denmark          DNK     10.000       94.62970  High income
## 54                  Spain          ESP      9.100       71.63500  High income
## 55                Estonia          EST     10.300       79.40000  High income
## 57                Finland          FIN     10.700       91.51440  High income
## 59                 France          FRA     12.300       81.91980  High income
## 62         United Kingdom          GBR     12.200       89.84410  High income
## 68      Equatorial Guinea          GNQ     35.362       16.40000  High income
## 69                 Greece          GRC      8.500       59.86630  High income
## 71              Greenland          GRL     14.500       65.80000  High income
## 73                   Guam          GUM     17.389       65.40000  High income
## 75   Hong Kong SAR, China          HKG      7.900       74.20000  High income
## 77                Croatia          HRV      9.400       66.74760  High income
## 79                Hungary          HUN      9.200       72.64390  High income
## 82                Ireland          IRL     15.000       78.24770  High income
## 85                Iceland          ISL     13.400       96.54680  High income
## 86                 Israel          ISR     21.300       70.80000  High income
## 87                  Italy          ITA      8.500       58.45930  High income
## 90                  Japan          JPN      8.200       89.71000  High income
## 96            Korea, Rep.          KOR      8.600       84.77000  High income
## 97                 Kuwait          KWT     20.575       75.46000  High income
## 103         Liechtenstein          LIE      9.200       93.80000  High income
## 106             Lithuania          LTU     10.100       68.45290  High income
## 107            Luxembourg          LUX     11.300       93.77650  High income
## 108                Latvia          LVA     10.200       75.23440  High income
## 109      Macao SAR, China          MAC     11.256       65.80000  High income
## 117                 Malta          MLT      9.500       68.91380  High income
## 127         New Caledonia          NCL     17.000       66.00000  High income
## 131           Netherlands          NLD     10.200       93.95640  High income
## 132                Norway          NOR     11.600       95.05340  High income
```

```
## 134           New Zealand          NZL     13.120       82.78000  High income
## 135                  Oman          OMN     20.419       66.45000  High income
## 141                Poland          POL      9.600       62.84920  High income
## 142           Puerto Rico          PRI     10.800       73.90000  High income
## 143              Portugal          PRT      7.900       62.09560  High income
## 145      French Polynesia          PYF     16.393       56.80000  High income
## 146                 Qatar          QAT     11.940       85.30000  High income
## 148    Russian Federation          RUS     13.200       67.97000  High income
## 150          Saudi Arabia          SAU     20.576       60.50000  High income
## 153             Singapore          SGP      9.300       81.00000  High income
## 162       Slovak Republic          SVK     10.100       77.88260  High income
## 163              Slovenia          SVN     10.200       72.67560  High income
## 164                Sweden          SWE     11.800       94.78360  High income
## 166            Seychelles          SYC     18.600       50.40000  High income
## 175   Trinidad and Tobago          TTO     14.590       63.80000  High income
## 181               Uruguay          URY     14.374       57.69000  High income
## 182         United States          USA     12.500       84.20000  High income
## 185         Venezuela, RB          VEN     19.842       54.90000  High income
## 186 Virgin Islands (U.S.)          VIR     10.700       45.30000  High income
```

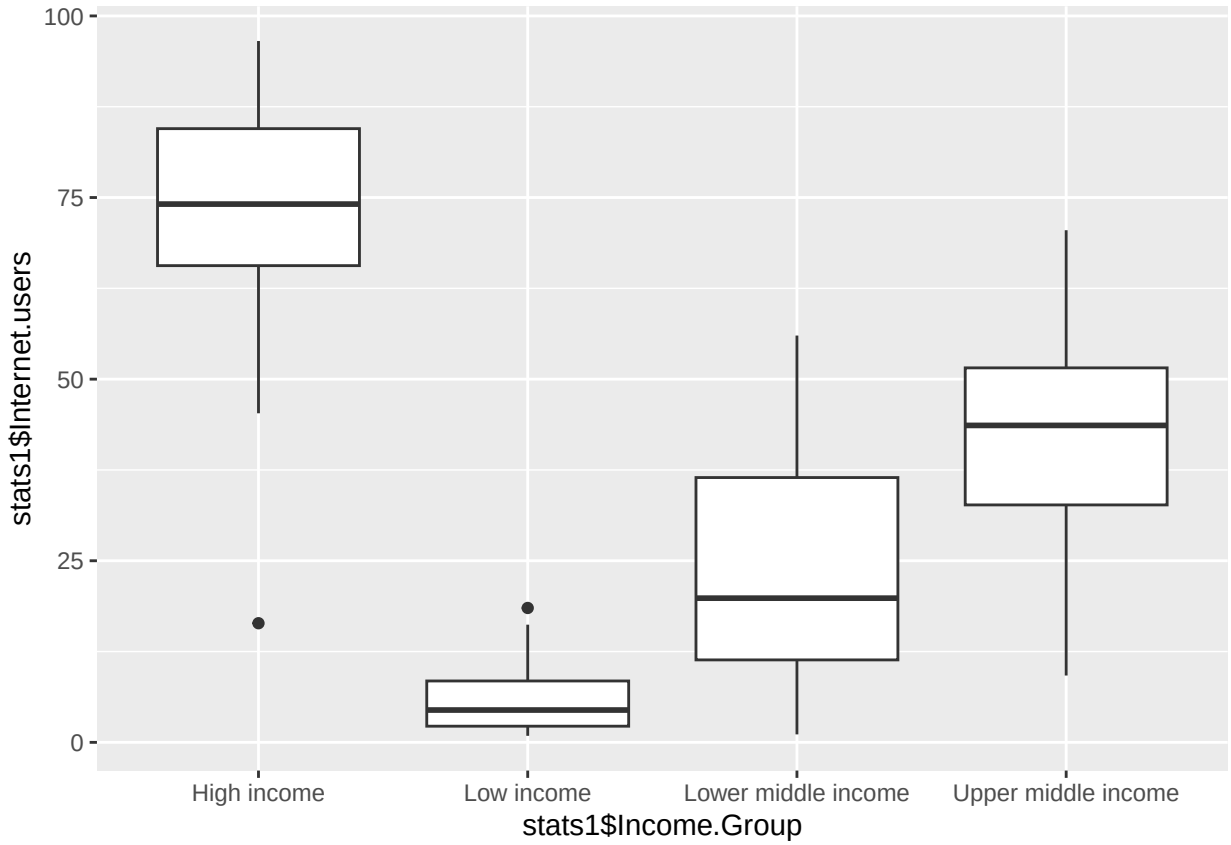
```
library(ggplot2)
```

```
qplot(data = stats1, x = stats1$Income.Group, y = stats1$Internet.users, geom="boxplot")
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

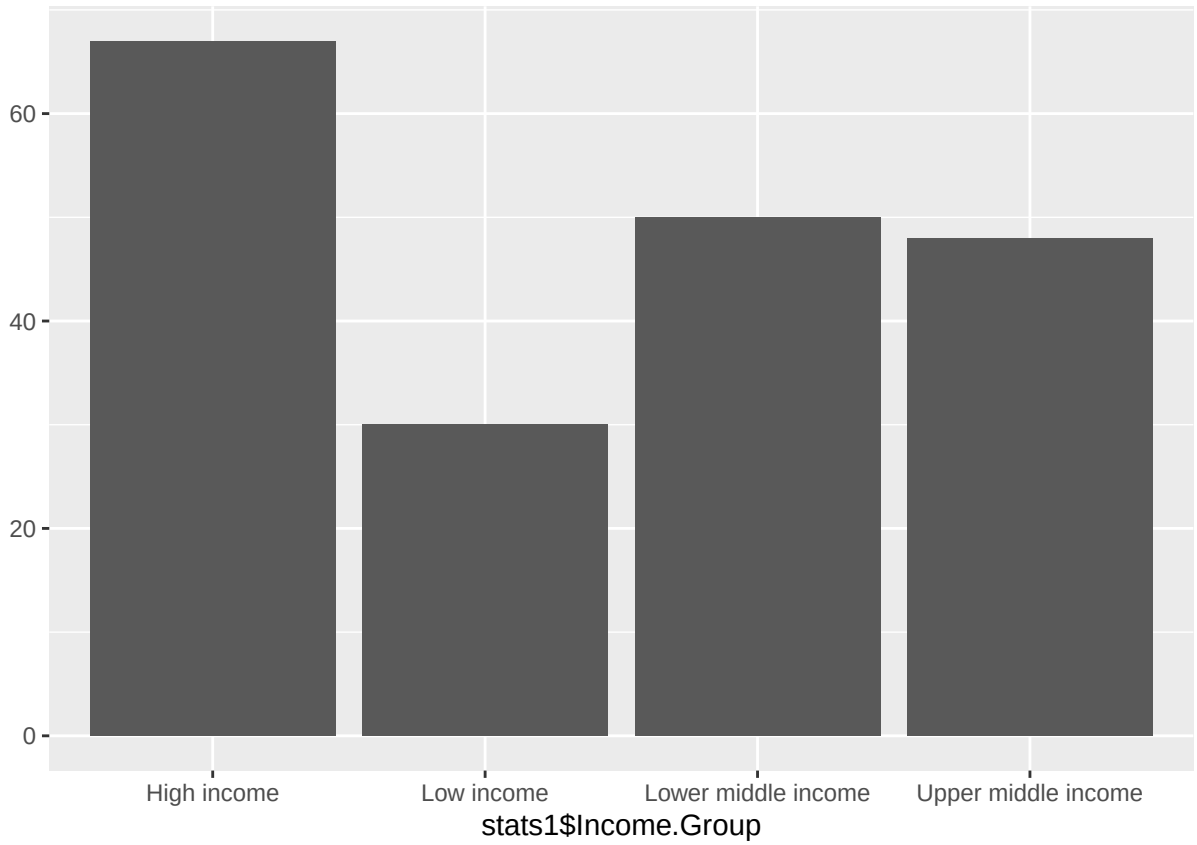
```
## Warning: Use of `stats1$Income.Group` is discouraged.
## i Use `Income.Group` instead.
```

```
## Warning: Use of `stats1$Internet.users` is discouraged.
## i Use `Internet.users` instead.
```




```
qplot(data = stats1, x = stats1$Income.Group)
```

```
## Warning: Use of `stats1$Income.Group` is discouraged.
## i Use `Income.Group` instead.
```

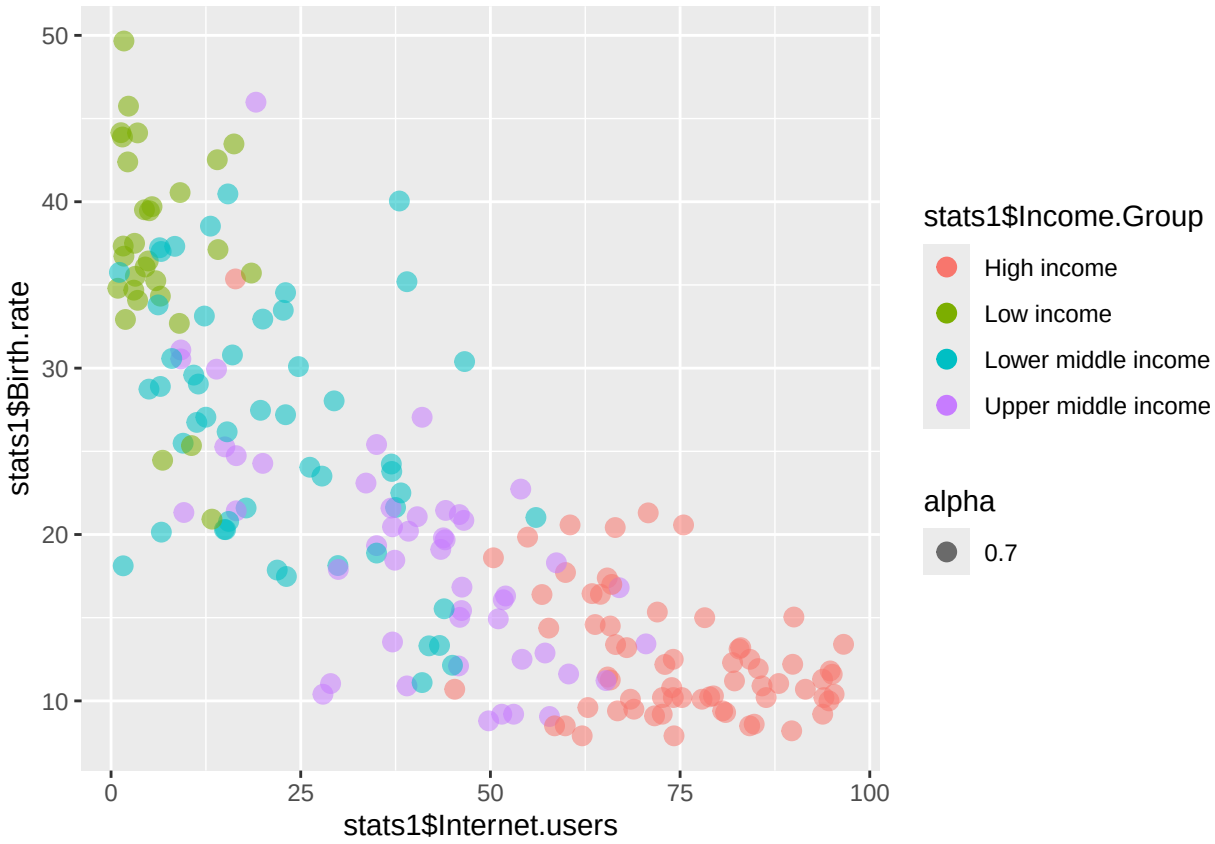



```
qplot(data = stats1, x = stats1$Internet.users, y = stats1$Birth.rate,
      color = stats1$Income.Group, alpha = 0.7, size = I(3))
```

```
## Warning: Use of `stats1$Internet.users` is discouraged.
## i Use `Internet.users` instead.
```

```
## Warning: Use of `stats1$Birth.rate` is discouraged.
## i Use `Birth.rate` instead.
```

```
## Warning: Use of `stats1$Income.Group` is discouraged.
## i Use `Income.Group` instead.
```




```
stats2 <- data.frame(Country.Code = Codes_2012_Dataset, Country.Name = Countries_2012_Dataset, Region =
stats2$Region <- factor(stats2$Region)
```

```
str(stats2)
```

```
## 'data.frame':    195 obs. of  3 variables:
##  $ Country.Code: chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Country.Name: chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Region      : Factor w/ 6 levels "Africa","Asia",..: 6 2 1 3 4 6 2 6 5 3 ...
```

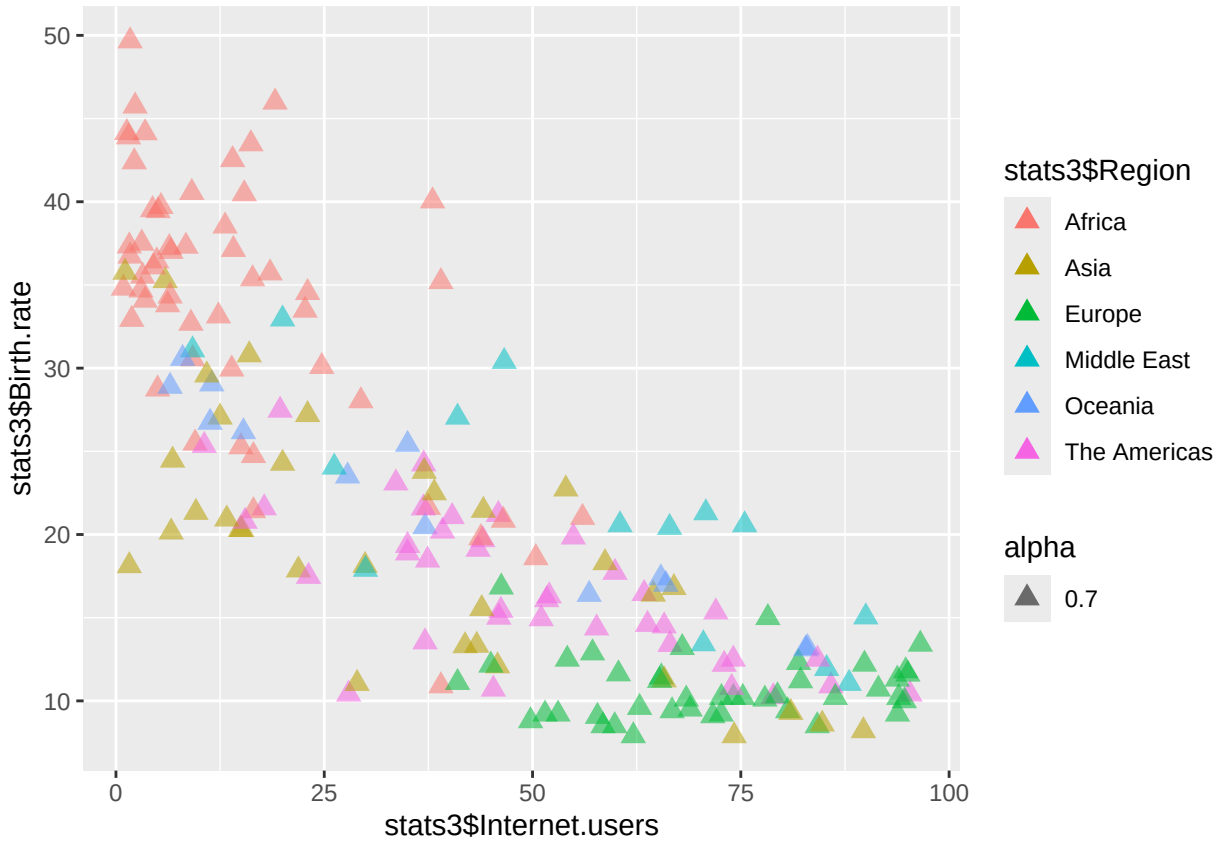
```
stats3 <- merge(stats1, stats2, by.x = "Country.Code", by.y = "Country.Code")
```

```
qplot(data = stats3, x = stats3$Internet.users, y = stats3$Birth.rate,
      color = stats3$Region, alpha = 0.7, size = I(3), shape = I(17))
```

```
## Warning: Use of `stats3$Internet.users` is discouraged.
## i Use `Internet.users` instead.
```

```
## Warning: Use of `stats3$Birth.rate` is discouraged.
## i Use `Birth.rate` instead.
```

```
## Warning: Use of `stats3$Region` is discouraged.
## i Use `Region` instead.
```




```
Country_Code <- c("ABW","AFG","AGO","ALB","ARE","ARG","ARM","ATG","AUS","AUT","AZE","BDI","BEL","BEN","
Life_Expectancy_At_Birth_1960 <- c(65.5693658536586,32.328512195122,32.9848292682927,62.2543658536585,5
Life_Expectancy_At_Birth_2013 <- c(75.3286585365854,60.0282682926829,51.8661707317073,77.537243902439,7
```

```
stats4 <- read.csv("S5-Homework-Data.csv")
```

```
head(stats4)
```

```
##           Country.Name Country.Code       Region Year Fertility.Rate
## 1                Aruba          ABW The Americas 1960          4.820
## 2          Afghanistan          AFG         Asia 1960          7.450
## 3               Angola          AGO       Africa 1960          7.379
## 4              Albania          ALB       Europe 1960          6.186
## 5 United Arab Emirates          ARE  Middle East 1960          6.928
## 6            Argentina          ARG The Americas 1960          3.109
```

```
str(stats4)
```

```
## 'data.frame':    374 obs. of  5 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Region        : chr  "The Americas" "Asia" "Africa" "Europe" ...
##  $ Year          : int  1960 1960 1960 1960 1960 1960 1960 1960 1960 1960 ...
##  $ Fertility.Rate: num  4.82 7.45 7.38 6.19 6.93 ...
```

```
stats1960 <- stats4[stats4$Year == "1960", ]
stats2013 <- stats4[stats4$Year == "2013", ]
```

```
str(stats1960)
```

```
## 'data.frame':    187 obs. of  5 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Region        : chr  "The Americas" "Asia" "Africa" "Europe" ...
##  $ Year          : int  1960 1960 1960 1960 1960 1960 1960 1960 1960 1960 ...
##  $ Fertility.Rate: num  4.82 7.45 7.38 6.19 6.93 ...
```

```
str(stats2013)
```

```
## 'data.frame':    187 obs. of  5 variables:
##  $ Country.Name  : chr  "Aruba" "Afghanistan" "Angola" "Albania" ...
##  $ Country.Code  : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ Region        : chr  "The Americas" "Asia" "Africa" "Europe" ...
##  $ Year          : int  2013 2013 2013 2013 2013 2013 2013 2013 2013 2013 ...
##  $ Fertility.Rate: num  1.67 5.05 6.17 1.77 1.8 ...
```

```
stats5 <- data.frame(Country.Code = Country_Code, LifeExpectancy1960 = Life_Expectancy_At_Birth_1960,  L
```

```
str(stats5)
```

```
## 'data.frame':    187 obs. of  3 variables:
##  $ Country.Code      : chr  "ABW" "AFG" "AGO" "ALB" ...
##  $ LifeExpectancy1960: num  65.6 32.3 33 62.3 52.2 ...
##  $ LifeExpectancy2013: num  75.3 60 51.9 77.5 77.2 ...
```

```
a <- merge(stats1960, stats5, by.x = "Country.Code", by.y = "Country.Code")
a$LifeExpectancy2013 <- NULL
stats6 <- a
```

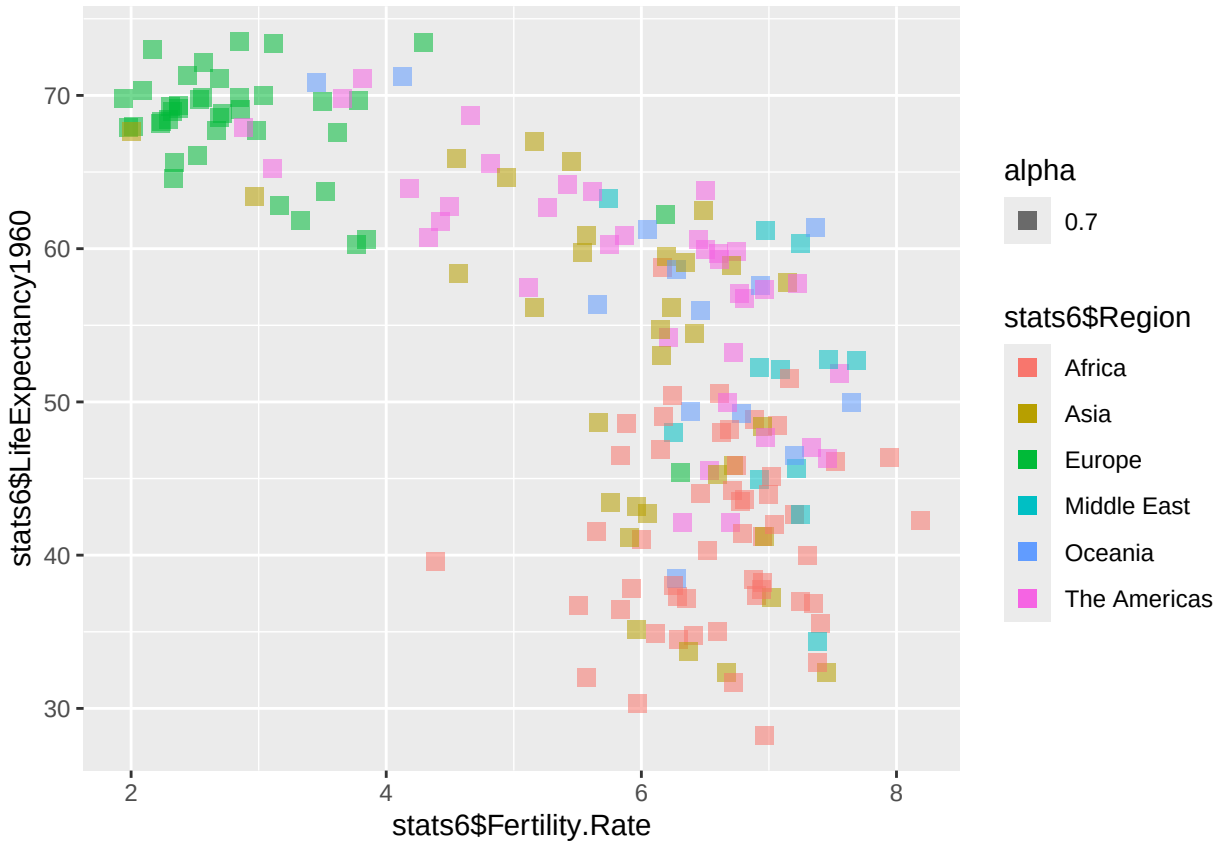
```
b <- merge(stats2013, stats5, by.x = "Country.Code", by.y = "Country.Code")
b$LifeExpectancy1960 <- NULL
stats7 <- b
```

```
qplot(data = stats6, x = stats6$Fertility.Rate, y = stats6$LifeExpectancy1960,
      color = stats6$Region, alpha = 0.7, size = I(3), shape = I(15))
```

```
## Warning: Use of `stats6$Fertility.Rate` is discouraged.
## i Use `Fertility.Rate` instead.
```

```
## Warning: Use of `stats6$LifeExpectancy1960` is discouraged.
## i Use `LifeExpectancy1960` instead.
```

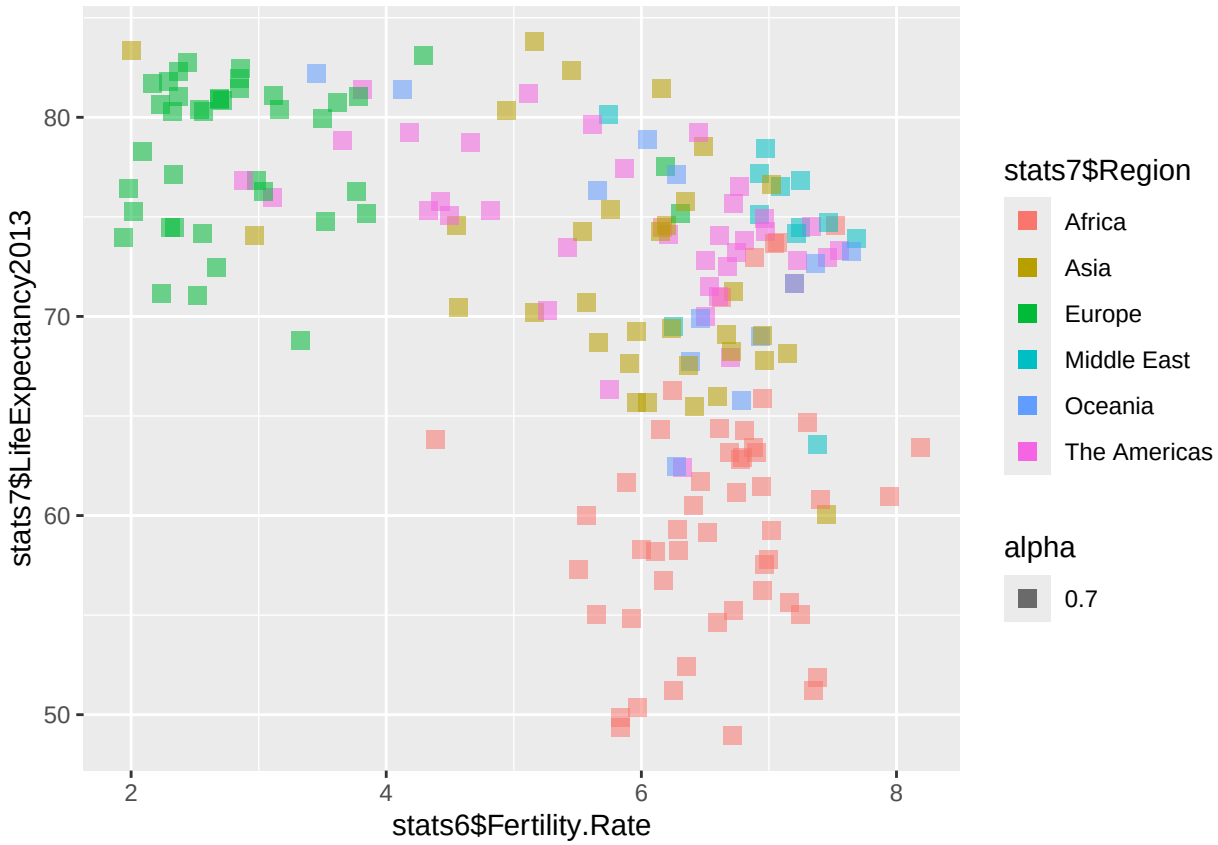
```
## Warning: Use of `stats6$Region` is discouraged.
## i Use `Region` instead.
```




```
qplot(data = stats7, x = stats6$Fertility.Rate, y = stats7$LifeExpectancy2013,
      color = stats7$Region, alpha = 0.7, size = I(3), shape = I(15))
```

```
## Warning: Use of `stats7$LifeExpectancy2013` is discouraged.
## i Use `LifeExpectancy2013` instead.
```

```
## Warning: Use of `stats7$Region` is discouraged.
## i Use `Region` instead.
```


stats7$LifeExpectancy2013
80
70
60
50
2
4
6
8
stats6$Fertility.Rate
stats7$Region
Africa
Asia
Europe
Middle East
Oceania
The Americas
alpha
0.7