Working with Data Structures

Kaitlyn Watson-Group 9

M2 ICA2

Introduction

Today we will investigate a data set from (http://www.gapminder.org), a site which contains a wealth of data and visualizations related to health, wealth, population, etc. of countries of the world. Typical datasets from (http://www.gapminder.org) are somewhat messy, with a lot of missing observations, so today we will work with data extracted from the more complete data sets by Jenny Bryan at UBC. First we load the data into R.This dataset can be extracted from the "gapminder" package in CRAN. As such initially we need to install and call the package. This dataset can aslo be as a .RData file in the same working directory as the .Rmd file and can be called from there.

```
#install.packages("gapminder")
library("gapminder")
gapminder <- system.file("extdata", "gapminder.tsv", package = "gapminder")
gapminder <- read.delim(gapminder)
ls(gapminder)
```

```
[1] "continent" "country"   "gdpPercap" "lifeExp"   "pop"       "year"
```

```
head(gapminder)
```

```
      country continent year lifeExp      pop gdpPercap
1 Afghanistan      Asia 1952  28.801  8425333  779.4453
2 Afghanistan      Asia 1957  30.332  9240934  820.8530
3 Afghanistan      Asia 1962  31.997 10267083  853.1007
4 Afghanistan      Asia 1967  34.020 11537966  836.1971
5 Afghanistan      Asia 1972  36.088 13079460  739.9811
6 Afghanistan      Asia 1977  38.438 14880372  786.1134
```

Function `ls` returns a vector of character strings giving the names of the objects in the specified environment.

Dataframe

The `str` function reports on the structure of an object in R. It's often useful to use `str` when working with a new dataset.

- If the data is large, never display the whole data in the .Rmd file. It will takes a long time
- to generate the html output.

```
str(gapminder)
```

```
'data.frame':   1704 obs. of  6 variables:
 $ country  : chr  "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...
 $ continent: chr  "Asia" "Asia" "Asia" "Asia" ...
 $ year     : int  1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...
 $ lifeExp  : num  28.8 30.3 32 34 36.1 ...
 $ pop      : int  8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957 16317921 22227
 $ gdpPercap: num  779 821 853 836 740 ...
```

We see that the `gapminder` data frame has six variables, including `country` (a factor), `year` (an integer), `pop` (numeric), `continent` (a factor), `lifeExp` (numeric), and `gdpPercap` (numeric).

How would you extract the third row of the data frame?

```
#Extract the third row here
gapminder [3, ]
```

```
      country continent year lifeExp      pop gdpPercap
3 Afghanistan      Asia 1962  31.997 10267083  853.1007
```

How would you extract the first 50 elements of the variable `year`?

```
#Extract the first 50 elements of the variable year here
gapminder[1:50, 3 ]
```

```
 [1] 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 2002 2007 1952 1957 1962
[16] 1967 1972 1977 1982 1987 1992 1997 2002 2007 1952 1957 1962 1967 1972 1977
[31] 1982 1987 1992 1997 2002 2007 1952 1957 1962 1967 1972 1977 1982 1987 1992
[46] 1997 2002 2007 1952 1957
```

There are a lot of repeated values in `year`. The `unique` function reports on the unique values in a variable.

```
#Display the unique values in the variable year here
unique(gapminder$year)
```

```
 [1] 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 2002 2007
```

Next we will work with two particular years, 1952 and 2002, from the dataset. We will use the `subset` function to extract data frames containing data from these two years. Here is how to extract the data for 1952:

```
gapminder1952 <- subset(gapminder, year == 1952)
str(gapminder1952)
```

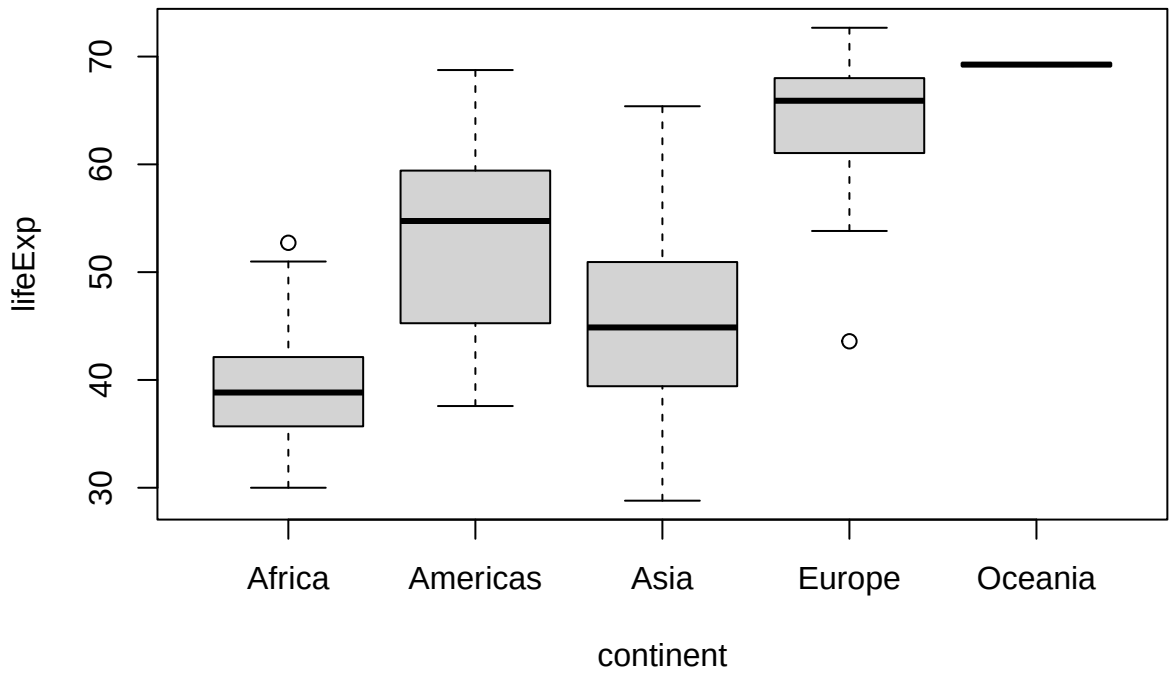
```
'data.frame':   142 obs. of  6 variables:
 $ country  : chr  "Afghanistan" "Albania" "Algeria" "Angola" ...
 $ continent: chr  "Asia" "Europe" "Africa" "Africa" ...
 $ year     : int  1952 1952 1952 1952 1952 1952 1952 1952 1952 1952 ...
 $ lifeExp  : num  28.8 55.2 43.1 30 62.5 ...
 $ pop      : int  8425333 1282697 9279525 4232095 17876956 8691212 6927772 120447 46886859 8730405 ...
 $ gdpPercap: num  779 1601 2449 3521 5911 ...
```

Note that to specify that a variable is equal to a particular value, we use ==. Next extract the data for 2002, and assign the resulting data frame to an object named gapminder2002.

```
#Create the data frame here
gapminder2002 <- subset(gapminder, year == 2002)
str(gapminder2002)
```

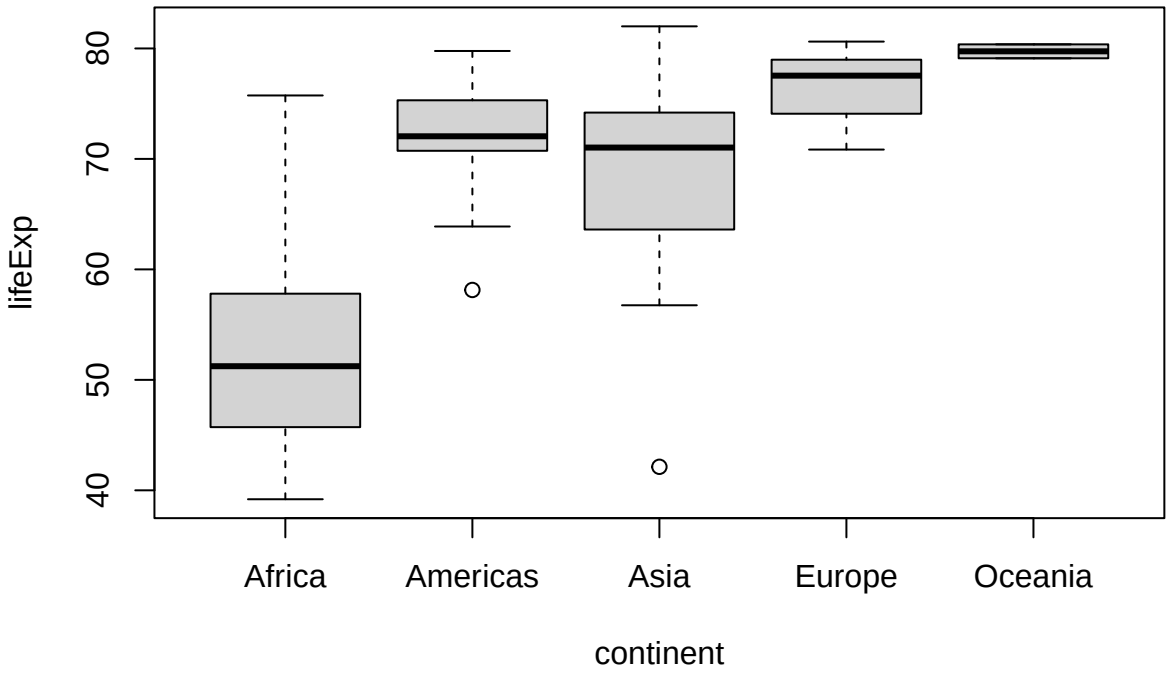
```
'data.frame':   142 obs. of  6 variables:
 $ country  : chr  "Afghanistan" "Albania" "Algeria" "Angola" ...
 $ continent: chr  "Asia" "Europe" "Africa" "Africa" ...
 $ year     : int  2002 2002 2002 2002 2002 2002 2002 2002 2002 2002 ...
 $ lifeExp  : num  42.1 75.7 71 41 74.3 ...
 $ pop      : int  25268405 3508512 31287142 10866106 38331121 19546792 8148312 656397 135656790 103119
 $ gdpPercap: num  727 4604 5288 2773 8798 ...
```

5. Create comparison boxplots for between continent and life expectancy for year 2002. An example from the 1952 data is given below.



Draw similar boxplots for 2002.

```
#Draw the boxplots here
boxplot(lifeExp ~ continent, data = gapminder2002)
```



6. From the boxplots for 2002, what do you notice? Which continent has the largest median life expectancy? Which continent has the largest inter-quartile range with regards to life expectancy?

Africa has the lowest life expectancy range in comparison to the rest of the continents. Oceana has the largest median life expectancy. Africa has the largest inter quartile range with regards to life expectancy.

7. Compute a few summary statistics for life expectancy and GDP in 1952. Compare these with the same statistics from 2002. Should we compare raw GDP numbers across time such as this?

No, because inflation is not accounted for in raw GDP and therefore, the numbers would be skewed. When compared GDP have larger average life expectancy and GDP in 2002 than in 1952.

```
summary(gapminder1952$lifeExp)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  28.80   39.06   45.14   49.06   59.77   72.67
```

```
summary(gapminder1952$gdpPercap)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
   298.9    864.8   1968.5   3725.3   3913.5 108382.4
```

```
summary (gapminder2002$lifeExp)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  39.19   55.52   70.83   65.69   75.46   82.00
```

```
summary (gapminder2002$gdpPercap)
```

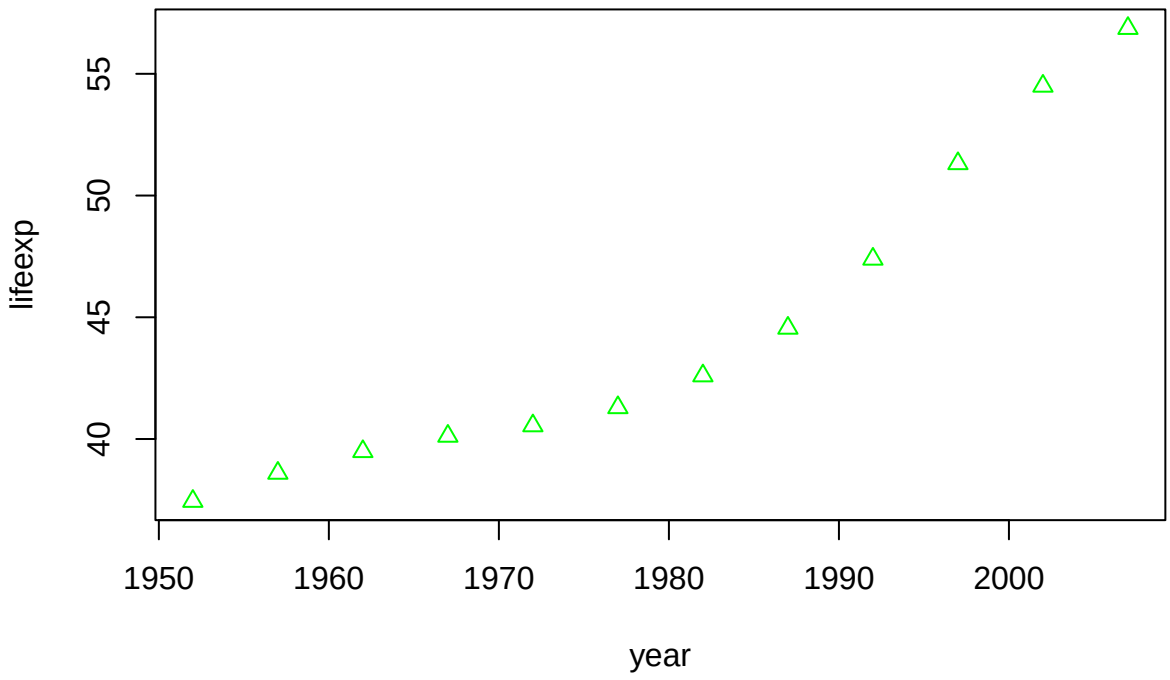
```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  241.2  1409.6  5319.8  9917.9 13359.5 44684.0
```

8. Choose a country of interest. Create a data frame which only contains data from that country. Draw scatter plots of life expectancy, and of the GDP, both versus the year. To add some customization such as color, connecting the points, and changing the point style see https://www.statmethods.net/advgraphs/parameters.html.

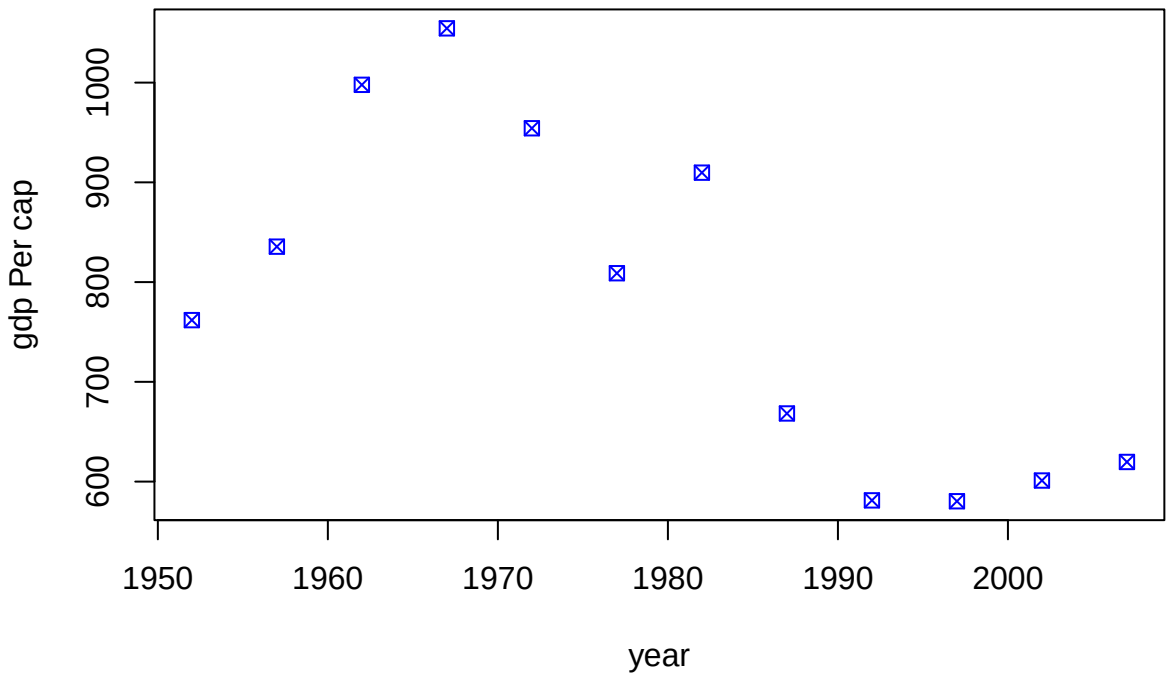
```
gapminderNiger <-subset(gapminder, country == "Niger")
gapminderNiger
```

```
     country continent year lifeExp      pop gdpPercap
1117   Niger    Africa 1952  37.444  3379468  761.8794
1118   Niger    Africa 1957  38.598  3692184  835.5234
1119   Niger    Africa 1962  39.487  4076008  997.7661
1120   Niger    Africa 1967  40.118  4534062 1054.3849
1121   Niger    Africa 1972  40.546  5060262  954.2092
1122   Niger    Africa 1977  41.291  5682086  808.8971
1123   Niger    Africa 1982  42.598  6437188  909.7221
1124   Niger    Africa 1987  44.555  7332638  668.3000
1125   Niger    Africa 1992  47.391  8392818  581.1827
1126   Niger    Africa 1997  51.313  9666252  580.3052
1127   Niger    Africa 2002  54.496 11140655  601.0745
1128   Niger    Africa 2007  56.867 12894865  619.6769
```

```
plot(gapminderNiger$year, gapminderNiger$lifeExp,pch=2,col="green", xlab = "year", ylab="lifeexp")
```




```
plot(gapminderNiger$year, gapminderNiger$gdpPercap, pch=7,col="blue",xlab = "year", ylab="gdp Per cap")
```



9. Next calculate the median life expectancy, for both 1952 and 2002.

```
#Calculate the median life expectancies here

gapminderNiger1952 <- subset(gapminderNiger, year == 1952)
str(gapminderNiger1952)
```

```
'data.frame':   1 obs. of  6 variables:
 $ country  : chr "Niger"
 $ continent: chr "Africa"
 $ year     : int 1952
 $ lifeExp  : num 37.4
 $ pop      : int 3379468
 $ gdpPercap: num 762
```

```
gapminderNiger2002 <- subset(gapminderNiger, year == 2002)
str(gapminderNiger2002)
```

```
'data.frame':   1 obs. of  6 variables:
 $ country  : chr "Niger"
 $ continent: chr "Africa"
 $ year     : int 2002
 $ lifeExp  : num 54.5
 $ pop      : int 11140655
 $ gdpPercap: num 601
```

```
median(gapminderNiger1952$lifeExp)
```

```
[1] 37.444
```

```
median(gapminderNiger2002$lifeExp)
```

```
[1] 54.496
```

What do you notice? The median is the same value as the single observation for Nigeria for those years. The life expectancy had increased from 1952-2002. 10. Next calculate the mean and median GDP for both years.

```
#Calculate the mean and median GDPs here.

mean(gapminderNiger1952$gdpPercap)
```

```
[1] 761.8794
```

```
median(gapminderNiger1952$gdpPercap)
```

```
[1] 761.8794
```

```
mean(gapminderNiger2002$gdpPercap)
```

```
[1] 601.0745
```

```
median(gapminderNiger2002$gdpPercap)
```

```
[1] 601.0745
```

Again, what do you notice?

The mean and median were the same for each year due to limited data, however the GDP increased from 1952 to 2002.

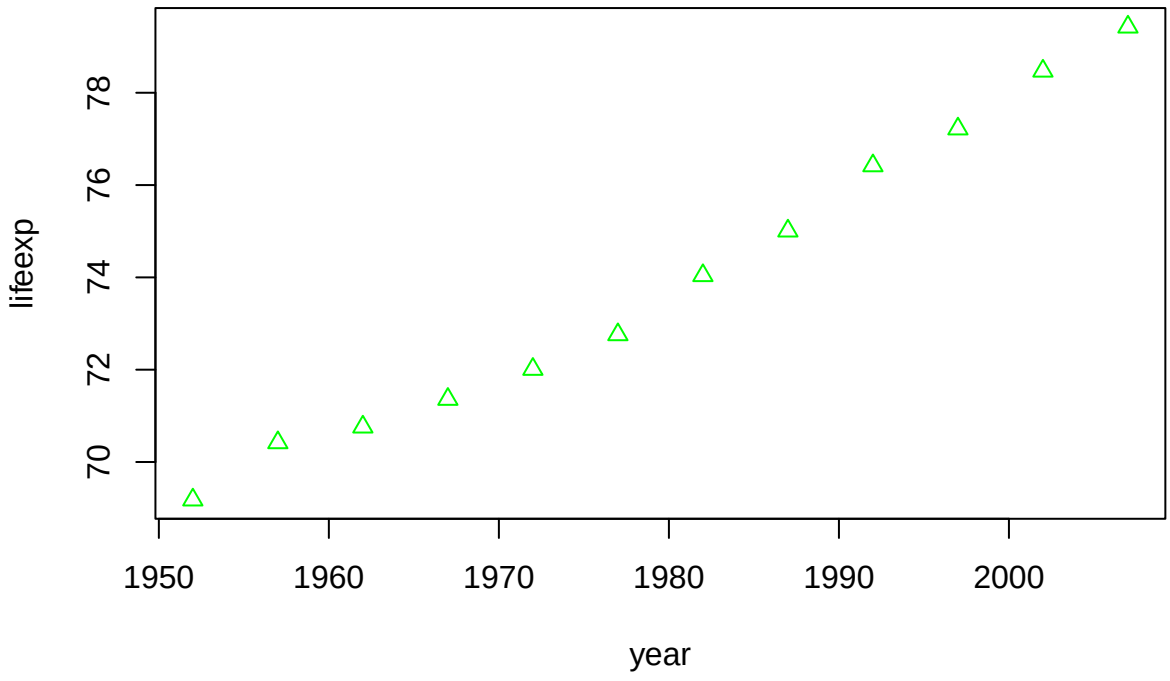
Choose a country of interest. Create a data frame which only contains data from that country, and then draw scatter plots of the life expectancy, and of the GDP, both versus the year.

```
#Create the data frame and draw the scatter plots here.
gapminderUK <-subset(gapminder, country == "United Kingdom")
gapminderUK
```

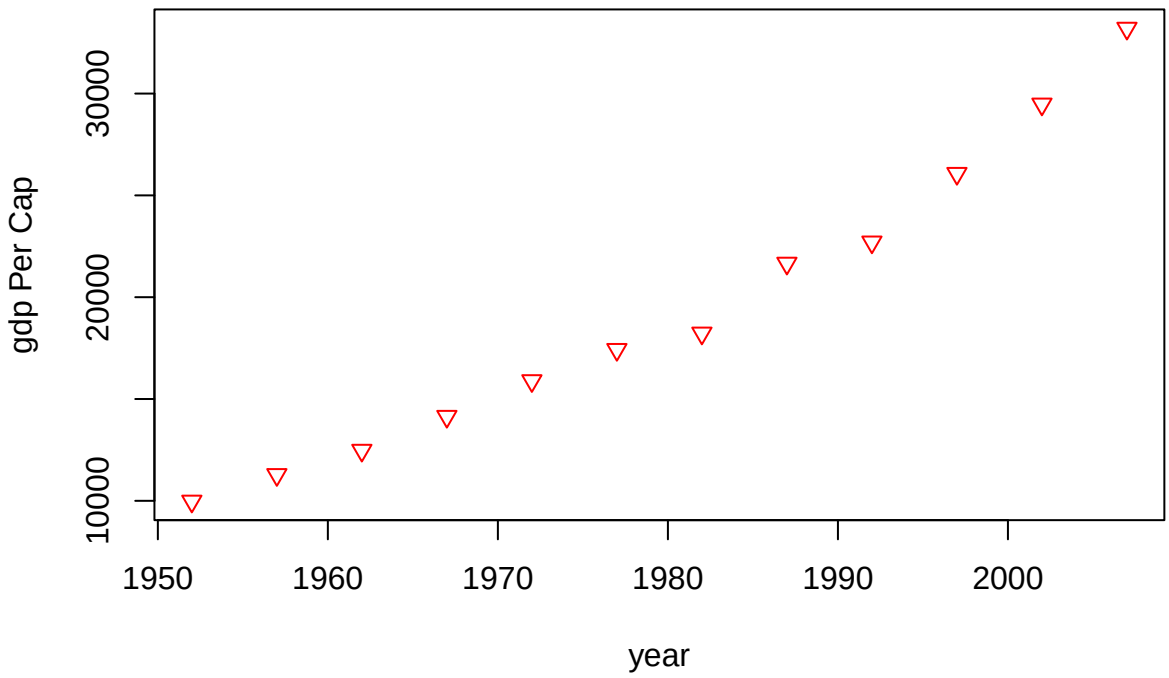
```
            country continent year lifeExp      pop gdpPercap
1597 United Kingdom    Europe 1952  69.180 50430000  9979.508
1598 United Kingdom    Europe 1957  70.420 51430000 11283.178
1599 United Kingdom    Europe 1962  70.760 53292000 12477.177
1600 United Kingdom    Europe 1967  71.360 54959000 14142.851
1601 United Kingdom    Europe 1972  72.010 56079000 15895.116
1602 United Kingdom    Europe 1977  72.760 56179000 17428.748
```

```
1603 United Kingdom    Europe 1982  74.040 56339704 18232.425
1604 United Kingdom    Europe 1987  75.007 56981620 21664.788
1605 United Kingdom    Europe 1992  76.420 57866349 22705.093
1606 United Kingdom    Europe 1997  77.218 58808266 26074.531
1607 United Kingdom    Europe 2002  78.471 59912431 29478.999
1608 United Kingdom    Europe 2007  79.425 60776238 33203.261
```

```
plot(gapminderUK$year, gapminderUK$lifeExp,pch=2,col="green", xlab = "year", ylab="lifeexp")
```




```
plot(gapminderUK$year, gapminderUK$gdpPercap,pch=6,col="red", xlab = "year", ylab="gdp Per Cap")
```



Logical subsetting

Subsetting with logical vectors is an essential skill. When a vector, say x, is subset with a logical vector, the components of x are returned wherever a TRUE value component exists in the logical vector. Below are some examples. Think about what is happening in each example. Recall that we can combine conditions with the operators `&` and `|` which represent "and" and "or".

```
mean(gapminder$pop[gapminder$country == "France"])
```

```
[1] 52952564
```

```
unique(gapminder$country[gapminder$continent == "Africa"])
```

```
 [1] "Algeria"                  "Angola"
 [3] "Benin"                    "Botswana"
 [5] "Burkina Faso"             "Burundi"
 [7] "Cameroon"                 "Central African Republic"
 [9] "Chad"                     "Comoros"
[11] "Congo, Dem. Rep."         "Congo, Rep."
[13] "Cote d'Ivoire"            "Djibouti"
[15] "Egypt"                    "Equatorial Guinea"
[17] "Eritrea"                  "Ethiopia"
[19] "Gabon"                    "Gambia"
```

```
[21] "Ghana"                     "Guinea"
[23] "Guinea-Bissau"             "Kenya"
[25] "Lesotho"                   "Liberia"
[27] "Libya"                     "Madagascar"
[29] "Malawi"                    "Mali"
[31] "Mauritania"                "Mauritius"
[33] "Morocco"                   "Mozambique"
[35] "Namibia"                   "Niger"
[37] "Nigeria"                   "Reunion"
[39] "Rwanda"                    "Sao Tome and Principe"
[41] "Senegal"                   "Sierra Leone"
[43] "Somalia"                   "South Africa"
[45] "Sudan"                     "Swaziland"
[47] "Tanzania"                  "Togo"
[49] "Tunisia"                   "Uganda"
[51] "Zambia"                    "Zimbabwe"
```

```
gapminder$country[(gapminder$pop > 150000000) & (gapminder$year == 1992)]
```

```
[1] "Brazil"        "China"         "India"         "Indonesia"
[5] "United States"
```

Operator `%in%` returns a logical vector indicating if there is a match or not for its left operand. Consider the example below

```
x <- 5:10
y <- c(3, 5, 6, 9, 12, 15)

x %in% y
```

```
[1]  TRUE  TRUE FALSE FALSE  TRUE FALSE
```

```
y %in% x
```

```
[1] FALSE  TRUE  TRUE  TRUE FALSE FALSE
```

Subsetting also can be used to change values of existing `R` objects as in the following example. Remove the chunk option `eval = FALSE` to see the example's result in your knitted HTML file.

```
dd <-  data.frame(x = c("dog", "cat", "oink", "pig", "oink", "cat", "dog"),
                  y = c("dog", "cat", "cat", "pig", "cow", "dog", "dog"),
                  stringsAsFactors = FALSE)
dd
```

```
     x   y
1  dog dog
2  cat cat
3 oink cat
4  pig pig
5 oink cow
6  cat dog
7  dog dog
```

```
dd$x[dd$x == "oink"] <- "pig"
dd
```

```
    x   y
1 dog dog
2 cat cat
3 pig cat
4 pig pig
5 pig cow
6 cat dog
7 dog dog
```

```
dd$same <- rep("no", dim(dd)[1])
dd
```

```
    x   y same
1 dog dog   no
2 cat cat   no
3 pig cat   no
4 pig pig   no
5 pig cow   no
6 cat dog   no
7 dog dog   no
```

```
dd$same[dd$x == dd$y] <- "yes"
dd
```

```
    x   y same
1 dog dog  yes
2 cat cat  yes
3 pig cat   no
4 pig pig  yes
5 pig cow   no
6 cat dog   no
7 dog dog  yes
```

Exercises

1. Extract the population values of all countries whose life expectancy is more than 70 years for the year 1967.

```
gapminder1967 <-subset(gapminder, year == 1967)
gapminder1967LE <- subset (gapminder1967, lifeExp >70)
gapminder1967LE
```

```
            country continent year lifeExp       pop gdpPercap
64        Australia   Oceania 1967   71.10  11872264 14526.125
76          Austria    Europe 1967   70.14   7376998 12834.602
112         Belgium    Europe 1967   70.94   9556500 13149.041
184        Bulgaria    Europe 1967   70.42   8310226  5577.003
```

```
244           Canada  Americas 1967   72.13  20819767 16076.588
400   Czech Republic    Europe 1967   70.38   9835109 11399.445
412          Denmark    Europe 1967   72.96   4838800 15937.211
532           France    Europe 1967   71.55  49569000 12999.918
568          Germany    Europe 1967   70.80  76368453 14745.626
592           Greece    Europe 1967   71.00   8716441  8513.097
688          Iceland    Europe 1967   73.73    198676 13319.896
748          Ireland    Europe 1967   71.08   2900100  7655.569
760           Israel      Asia 1967   70.75   2693585  8393.741
772            Italy    Europe 1967   71.06  52667100 10022.401
796            Japan      Asia 1967   71.43 100825279  9847.789
1084     Netherlands    Europe 1967   73.82  12596822 15363.251
1096     New Zealand   Oceania 1967   71.52   2728150 14463.919
1144          Norway    Europe 1967   74.08   3786019 16361.876
1252     Puerto Rico  Americas 1967   71.10   2648961  6929.278
1372 Slovak Republic    Europe 1967   70.98   4442238  8412.902
1420           Spain    Europe 1967   71.44  32850275  7993.512
1468          Sweden    Europe 1967   74.16   7867931 15258.297
1480     Switzerland    Europe 1967   72.77   6063000 22966.144
1600  United Kingdom    Europe 1967   71.36  54959000 14142.851
1612   United States  Americas 1967   70.76 198712000 19530.366
```

2. For the year 2007, how many countries had a life expectancy of at least 75?

```
gapminder2007 <-subset(gapminder, year == 2007)
gapminder2007LE <- subset (gapminder2007, lifeExp >= 75)
gapminder2007LE
```

```
             country continent year lifeExp       pop gdpPercap
24           Albania    Europe 2007  76.423   3600523  5937.030
60         Argentina  Americas 2007  75.320  40301927 12779.380
72         Australia   Oceania 2007  81.235  20434176 34435.367
84           Austria    Europe 2007  79.829   8199783 36126.493
96           Bahrain      Asia 2007  75.635    708573 29796.048
120          Belgium    Europe 2007  79.441  10392226 33692.605
252           Canada  Americas 2007  80.653  33390141 36319.235
288            Chile  Americas 2007  78.553  16284741 13171.639
360       Costa Rica  Americas 2007  78.782   4133884  9645.061
384          Croatia    Europe 2007  75.748   4493312 14619.223
396             Cuba  Americas 2007  78.273  11416987  8948.103
408   Czech Republic    Europe 2007  76.486  10228744 22833.309
420          Denmark    Europe 2007  78.332   5468120 35278.419
528          Finland    Europe 2007  79.313   5238460 33207.084
540           France    Europe 2007  80.657  61083916 30470.017
576          Germany    Europe 2007  79.406  82400996 32170.374
600           Greece    Europe 2007  79.483  10706290 27538.412
672 Hong Kong, China      Asia 2007  82.208   6980412 39724.979
696          Iceland    Europe 2007  81.757    301931 36180.789
756          Ireland    Europe 2007  78.885   4109086 40675.996
768           Israel      Asia 2007  80.745   6426679 25523.277
780            Italy    Europe 2007  80.546  58147733 28569.720
804            Japan      Asia 2007  82.603 127467972 31656.068
852      Korea, Rep.      Asia 2007  78.623  49044790 23348.140
```

```
864            Kuwait      Asia 2007  77.588   2505559 47306.990
996            Mexico  Americas 2007  76.195 108700891 11977.575
1092      Netherlands    Europe 2007  79.762  16570613 36797.933
1104      New Zealand   Oceania 2007  80.204   4115771 25185.009
1152           Norway    Europe 2007  80.196   4627926 49357.190
1164             Oman      Asia 2007  75.640   3204897 22316.193
1188           Panama  Americas 2007  75.537   3242173  9809.186
1236           Poland    Europe 2007  75.563  38518241 15389.925
1248         Portugal    Europe 2007  78.098  10642836 20509.648
1260      Puerto Rico  Americas 2007  78.746   3942491 19328.709
1272          Reunion    Africa 2007  76.442    798094  7670.123
1368        Singapore      Asia 2007  79.972   4553009 47143.180
1392         Slovenia    Europe 2007  77.926   2009245 25768.258
1428            Spain    Europe 2007  80.941  40448191 28821.064
1476           Sweden    Europe 2007  80.884   9031088 33859.748
1488      Switzerland    Europe 2007  81.701   7554661 37506.419
1512           Taiwan      Asia 2007  78.400  23174294 28718.277
1608   United Kingdom    Europe 2007  79.425  60776238 33203.261
1620    United States  Americas 2007  78.242 301139947 42951.653
1632          Uruguay  Americas 2007  76.384   3447496 10611.463
```

44 countries have a life expectancy of at least 75.

3. Add a variable called `G8` to the `gapminder` data frame, which will be equal to 1 or 0 depending on whether the country is in the G8 group of nations: France, Germany, Italy, the United Kingdom, Japan, the United States, Canada, and Russia.

```
gapminder$G8 <- 0
G8country <- c("France", "Germany", "Italy", "United Kingdom", "Japan", "United States", "Canada")
gapminder$G8 [gapminder$country %in% G8country] <- 1
unique(gapminder$country [gapminder$G8 ==1])
```

```
[1] "Canada"         "France"         "Germany"        "Italy"
[5] "Japan"          "United Kingdom" "United States"
```

4. Create a plot of your choice that involves countries of the G8.

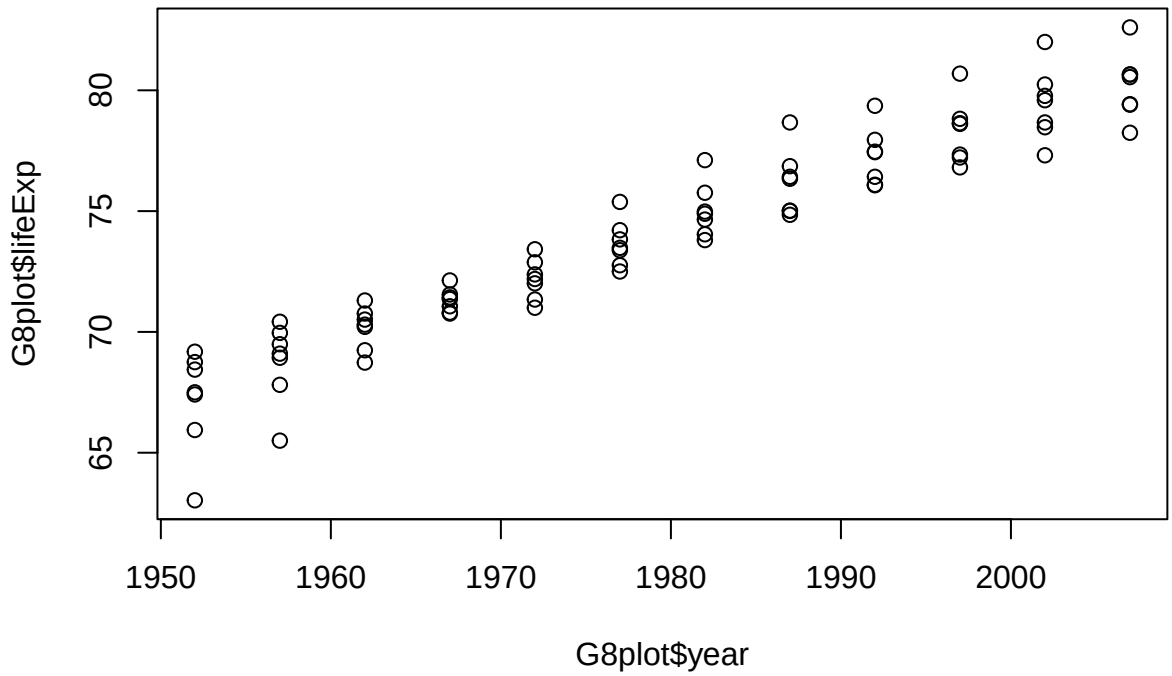
```
G8plot <- subset(gapminder, subset= (G8==1))
G8plot
```

```
           country continent year lifeExp       pop gdpPercap G8
241         Canada  Americas 1952  68.750  14785584 11367.161  1
242         Canada  Americas 1957  69.960  17010154 12489.950  1
243         Canada  Americas 1962  71.300  18985849 13462.486  1
244         Canada  Americas 1967  72.130  20819767 16076.588  1
245         Canada  Americas 1972  72.880  22284500 18970.571  1
246         Canada  Americas 1977  74.210  23796400 22090.883  1
247         Canada  Americas 1982  75.760  25201900 22898.792  1
248         Canada  Americas 1987  76.860  26549700 26626.515  1
249         Canada  Americas 1992  77.950  28523502 26342.884  1
250         Canada  Americas 1997  78.610  30305843 28954.926  1
251         Canada  Americas 2002  79.770  31902268 33328.965  1
```

252	Canada	Americas	2007	80.653	33390141	36319.235	1
529	France	Europe	1952	67.410	42459667	7029.809	1
530	France	Europe	1957	68.930	44310863	8662.835	1
531	France	Europe	1962	70.510	47124000	10560.486	1
532	France	Europe	1967	71.550	49569000	12999.918	1
533	France	Europe	1972	72.380	51732000	16107.192	1
534	France	Europe	1977	73.830	53165019	18292.635	1
535	France	Europe	1982	74.890	54433565	20293.897	1
536	France	Europe	1987	76.340	55630100	22066.442	1
537	France	Europe	1992	77.460	57374179	24703.796	1
538	France	Europe	1997	78.640	58623428	25889.785	1
539	France	Europe	2002	79.590	59925035	28926.032	1
540	France	Europe	2007	80.657	61083916	30470.017	1
565	Germany	Europe	1952	67.500	69145952	7144.114	1
566	Germany	Europe	1957	69.100	71019069	10187.827	1
567	Germany	Europe	1962	70.300	73739117	12902.463	1
568	Germany	Europe	1967	70.800	76368453	14745.626	1
569	Germany	Europe	1972	71.000	78717088	18016.180	1
570	Germany	Europe	1977	72.500	78160773	20512.921	1
571	Germany	Europe	1982	73.800	78335266	22031.533	1
572	Germany	Europe	1987	74.847	77718298	24639.186	1
573	Germany	Europe	1992	76.070	80597764	26505.303	1
574	Germany	Europe	1997	77.340	82011073	27788.884	1
575	Germany	Europe	2002	78.670	82350671	30035.802	1
576	Germany	Europe	2007	79.406	82400996	32170.374	1
769	Italy	Europe	1952	65.940	47666000	4931.404	1
770	Italy	Europe	1957	67.810	49182000	6248.656	1
771	Italy	Europe	1962	69.240	50843200	8243.582	1
772	Italy	Europe	1967	71.060	52667100	10022.401	1
773	Italy	Europe	1972	72.190	54365564	12269.274	1
774	Italy	Europe	1977	73.480	56059245	14255.985	1
775	Italy	Europe	1982	74.980	56535636	16537.483	1
776	Italy	Europe	1987	76.420	56729703	19207.235	1
777	Italy	Europe	1992	77.440	56840847	22013.645	1
778	Italy	Europe	1997	78.820	57479469	24675.024	1
779	Italy	Europe	2002	80.240	57926999	27968.098	1
780	Italy	Europe	2007	80.546	58147733	28569.720	1
793	Japan	Asia	1952	63.030	86459025	3216.956	1
794	Japan	Asia	1957	65.500	91563009	4317.694	1
795	Japan	Asia	1962	68.730	95831757	6576.649	1
796	Japan	Asia	1967	71.430	100825279	9847.789	1
797	Japan	Asia	1972	73.420	107188273	14778.786	1
798	Japan	Asia	1977	75.380	113872473	16610.377	1
799	Japan	Asia	1982	77.110	118454974	19384.106	1
800	Japan	Asia	1987	78.670	122091325	22375.942	1
801	Japan	Asia	1992	79.360	124329269	26824.895	1
802	Japan	Asia	1997	80.690	125956499	28816.585	1
803	Japan	Asia	2002	82.000	127065841	28604.592	1
804	Japan	Asia	2007	82.603	127467972	31656.068	1
1597	United Kingdom	Europe	1952	69.180	50430000	9979.508	1
1598	United Kingdom	Europe	1957	70.420	51430000	11283.178	1
1599	United Kingdom	Europe	1962	70.760	53292000	12477.177	1
1600	United Kingdom	Europe	1967	71.360	54959000	14142.851	1
1601	United Kingdom	Europe	1972	72.010	56079000	15895.116	1

```
1602 United Kingdom    Europe 1977  72.760  56179000 17428.748  1
1603 United Kingdom    Europe 1982  74.040  56339704 18232.425  1
1604 United Kingdom    Europe 1987  75.007  56981620 21664.788  1
1605 United Kingdom    Europe 1992  76.420  57866349 22705.093  1
1606 United Kingdom    Europe 1997  77.218  58808266 26074.531  1
1607 United Kingdom    Europe 2002  78.471  59912431 29478.999  1
1608 United Kingdom    Europe 2007  79.425  60776238 33203.261  1
1609  United States  Americas 1952  68.440 157553000 13990.482  1
1610  United States  Americas 1957  69.490 171984000 14847.127  1
1611  United States  Americas 1962  70.210 186538000 16173.146  1
1612  United States  Americas 1967  70.760 198712000 19530.366  1
1613  United States  Americas 1972  71.340 209896000 21806.036  1
1614  United States  Americas 1977  73.380 220239000 24072.632  1
1615  United States  Americas 1982  74.650 232187835 25009.559  1
1616  United States  Americas 1987  75.020 242803533 29884.350  1
1617  United States  Americas 1992  76.090 256894189 32003.932  1
1618  United States  Americas 1997  76.810 272911760 35767.433  1
1619  United States  Americas 2002  77.310 287675526 39097.100  1
1620  United States  Americas 2007  78.242 301139947 42951.653  1
```

```
plot(x=G8plot$year, y=G8plot$lifeExp)
```



References

1. https://www.gapminder.org/data/
2. https://www.statmethods.net/advgraphs/parameters.html