Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

ADVANCED DATA VISUALIZATION TECHNIQUES

NAME : MOTHISHWARAN C.

REG No : 19MID0017

COURSE CODE : CSI3005

SLOT : L35+L36

FACULTY : DR. LYDIA JANE G

GGPlots

Answer the following questions using the provided Expectancy.csv file.

Create all plots using ggplot2 library. Add title, x axis and y axis label to all the plots. Apply color and fill attribute as necessary.

```
library(ggplot2)

df_expectancy <- read.csv("Expectancy.csv")
head(df_expectancy)

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

tail(df_expectancy)
```

```
##         X  country continent year lifeExp      pop gdpPercap
## 1699 1699 Zimbabwe    Africa 1982  60.363  7636524  788.8550
## 1700 1700 Zimbabwe    Africa 1987  62.351  9216418  706.1573
## 1701 1701 Zimbabwe    Africa 1992  60.377 10704340  693.4208
## 1702 1702 Zimbabwe    Africa 1997  46.809 11404948  792.4500
## 1703 1703 Zimbabwe    Africa 2002  39.989 11926563  672.0386
## 1704 1704 Zimbabwe    Africa 2007  43.487 12311143  469.7093

str(df_expectancy)

## 'data.frame':    1704 obs. of  7 variables:
##  $ X        : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ 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 128
81816 13867957 16317921 22227415 ...
##  $ gdpPercap: num  779 821 853 836 740 ...

summary(df_expectancy)

##        X            country           continent              year
##  Min.   :   1.0   Length:1704        Length:1704        Min.   :1952
##  1st Qu.: 426.8   Class :character   Class :character   1st Qu.:1966
##  Median : 852.5   Mode  :character   Mode  :character   Median :1980
##  Mean   : 852.5                                        Mean   :1980
##  3rd Qu.:1278.2                                        3rd Qu.:1993
##  Max.   :1704.0                                        Max.   :2007
##     lifeExp          pop             gdpPercap
##  Min.   :23.60   Min.   :6.001e+04   Min.   :   241.2
##  1st Qu.:48.20   1st Qu.:2.794e+06   1st Qu.:  1202.1
##  Median :60.71   Median :7.024e+06   Median :  3531.8
##  Mean   :59.47   Mean   :2.960e+07   Mean   :  7215.3
##  3rd Qu.:70.85   3rd Qu.:1.959e+07   3rd Qu.:  9325.5
##  Max.   :82.60   Max.   :1.319e+09   Max.   :113523.1

lapply(df_expectancy,function(x) { length(which(is.na(x)))})

## $X
## [1] 0
##
## $country
## [1] 0
##
## $continent
## [1] 0
##
## $year
## [1] 0
```

```
##
## $lifeExp
## [1] 0
##
## $pop
## [1] 0
##
## $gdpPercap
## [1] 0
```

```
head(df_expectancy)
```

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

```
unique(df_expectancy[c("country")])
```

```
##                       country
## 1                 Afghanistan
## 13                    Albania
## 25                    Algeria
## 37                     Angola
## 49                  Argentina
## 61                  Australia
## 73                    Austria
## 85                    Bahrain
## 97                 Bangladesh
## 109                   Belgium
## 121                     Benin
## 133                   Bolivia
## 145    Bosnia and Herzegovina
## 157                  Botswana
## 169                    Brazil
## 181                  Bulgaria
## 193              Burkina Faso
## 205                   Burundi
## 217                  Cambodia
## 229                  Cameroon
## 241                    Canada
## 253  Central African Republic
## 265                      Chad
## 277                     Chile
## 289                     China
## 301                  Colombia
## 313                   Comoros
## 325          Congo, Dem. Rep.
```

```
## 337               Congo, Rep.
## 349                Costa Rica
## 361             Cote d'Ivoire
## 373                   Croatia
## 385                      Cuba
## 397            Czech Republic
## 409                   Denmark
## 421                  Djibouti
## 433        Dominican Republic
## 445                   Ecuador
## 457                     Egypt
## 469               El Salvador
## 481         Equatorial Guinea
## 493                   Eritrea
## 505                  Ethiopia
## 517                   Finland
## 529                    France
## 541                     Gabon
## 553                    Gambia
## 565                   Germany
## 577                     Ghana
## 589                    Greece
## 601                 Guatemala
## 613                    Guinea
## 625             Guinea-Bissau
## 637                     Haiti
## 649                  Honduras
## 661          Hong Kong, China
## 673                   Hungary
## 685                   Iceland
## 697                     India
## 709                 Indonesia
## 721                      Iran
## 733                      Iraq
## 745                   Ireland
## 757                    Israel
## 769                     Italy
## 781                   Jamaica
## 793                     Japan
## 805                    Jordan
## 817                     Kenya
## 829          Korea, Dem. Rep.
## 841               Korea, Rep.
## 853                    Kuwait
## 865                   Lebanon
## 877                   Lesotho
## 889                   Liberia
## 901                     Libya
## 913                Madagascar
## 925                    Malawi
```

```
## 937                  Malaysia
## 949                      Mali
## 961                Mauritania
## 973                 Mauritius
## 985                    Mexico
## 997                  Mongolia
## 1009               Montenegro
## 1021                  Morocco
## 1033               Mozambique
## 1045                  Myanmar
## 1057                  Namibia
## 1069                    Nepal
## 1081              Netherlands
## 1093              New Zealand
## 1105                Nicaragua
## 1117                    Niger
## 1129                  Nigeria
## 1141                   Norway
## 1153                     Oman
## 1165                 Pakistan
## 1177                   Panama
## 1189                 Paraguay
## 1201                     Peru
## 1213              Philippines
## 1225                   Poland
## 1237                 Portugal
## 1249              Puerto Rico
## 1261                  Reunion
## 1273                  Romania
## 1285                   Rwanda
## 1297    Sao Tome and Principe
## 1309             Saudi Arabia
## 1321                  Senegal
## 1333                   Serbia
## 1345             Sierra Leone
## 1357                Singapore
## 1369          Slovak Republic
## 1381                 Slovenia
## 1393                  Somalia
## 1405             South Africa
## 1417                    Spain
## 1429                Sri Lanka
## 1441                    Sudan
## 1453                Swaziland
## 1465                   Sweden
## 1477              Switzerland
## 1489                    Syria
## 1501                   Taiwan
## 1513                 Tanzania
## 1525                 Thailand
```

```
## 1537                    Togo
## 1549      Trinidad and Tobago
## 1561                  Tunisia
## 1573                   Turkey
## 1585                   Uganda
## 1597           United Kingdom
## 1609            United States
## 1621                  Uruguay
## 1633                Venezuela
## 1645                  Vietnam
## 1657       West Bank and Gaza
## 1669              Yemen, Rep.
## 1681                   Zambia
## 1693                 Zimbabwe
```

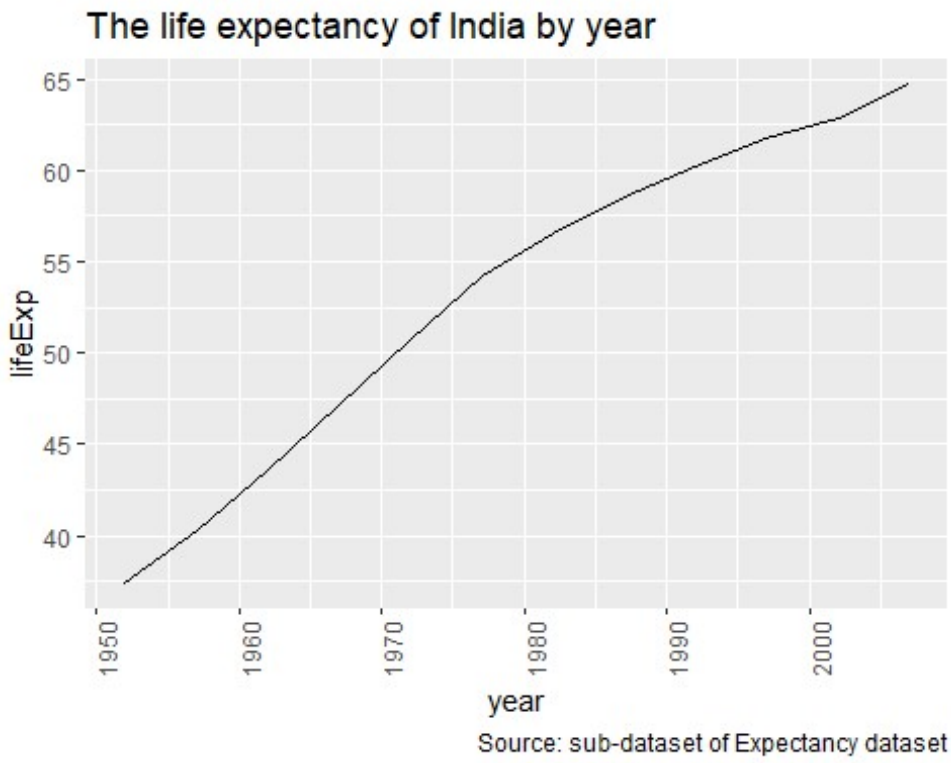
1. Choose a country and plot the life expectancy by year (both scatter and line plot)

```
df_india <- subset(df_expectancy,country =="India")
head(df_india)
```

```
##       X country continent year lifeExp       pop gdpPercap
## 697 697   India      Asia 1952  37.373 372000000  546.5657
## 698 698   India      Asia 1957  40.249 409000000  590.0620
## 699 699   India      Asia 1962  43.605 454000000  658.3472
## 700 700   India      Asia 1967  47.193 506000000  700.7706
## 701 701   India      Asia 1972  50.651 567000000  724.0325
## 702 702   India      Asia 1977  54.208 634000000  813.3373
```

```
scatterplot<-ggplot(data=df_india, aes(x=year, y=lifeExp))+geom_point()
scatterplot+labs(
    title = "The life expectancy of India by year",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```


```
lineplot<-ggplot(data=df_india, aes(x=year, y=lifeExp))+geom_line()
lineplot+labs(
    title = "The life expectancy of India by year",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```



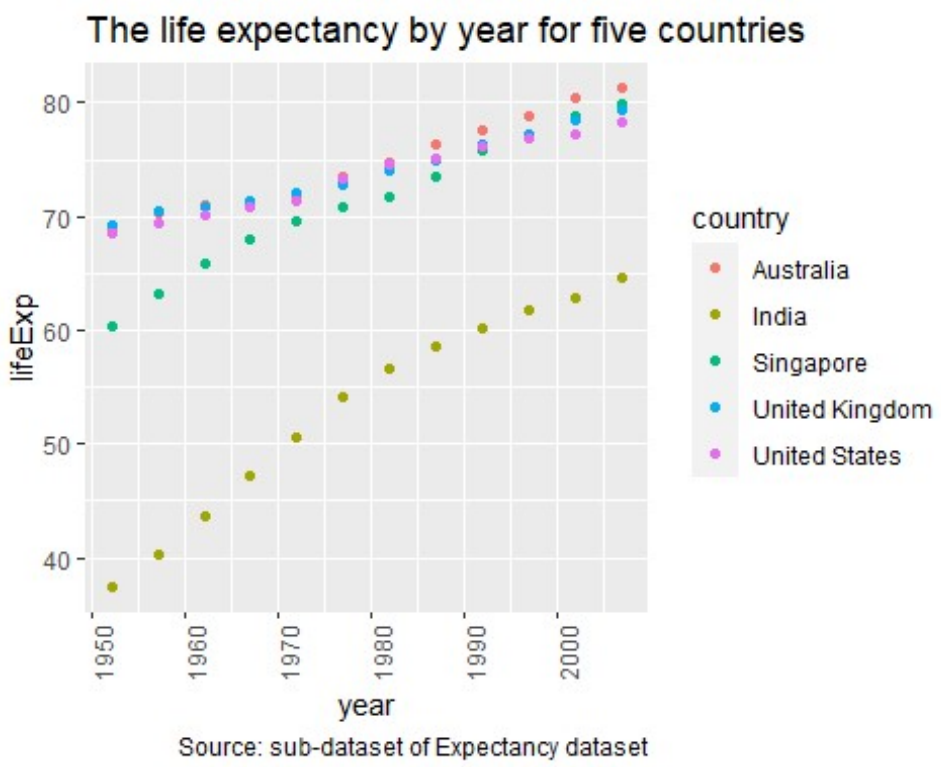

```
scatterplot<-ggplot(data=df_india, aes(x=year, y=lifeExp))+geom_point()+geom_
line()
scatterplot+labs(
    title = "The life expectancy of India by year",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```

2. Plot the life expectancy by year for five countries. (both scatter and line plot)

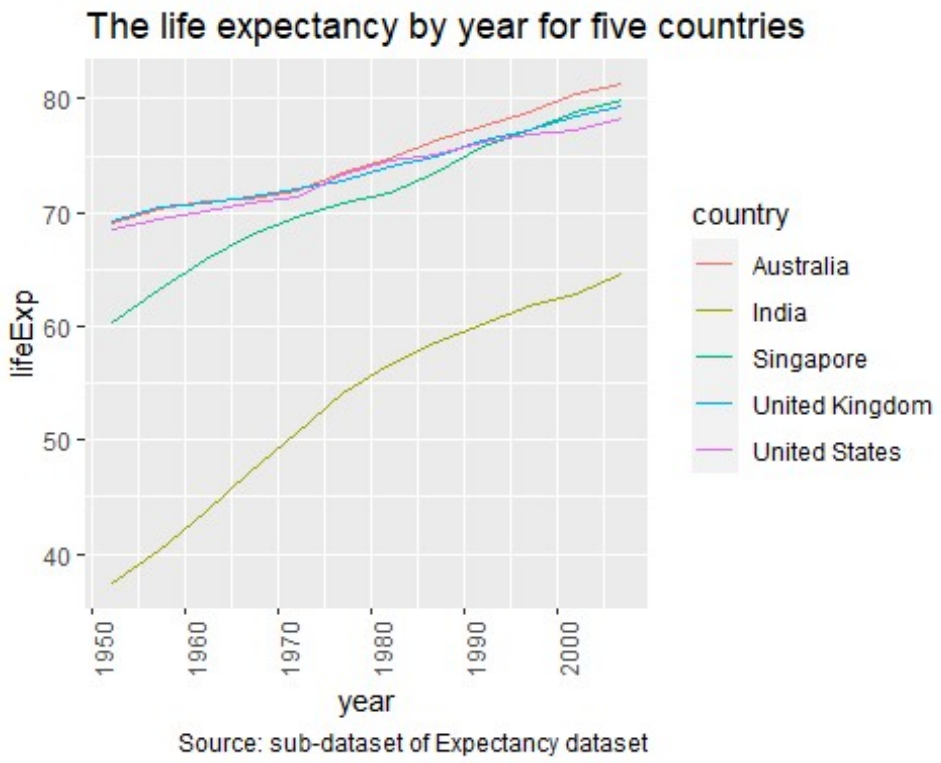
```
df_5countries<-df_expectancy[df_expectancy$country %in% c("India", "United States", "United Kingdom", "Australia", "Singapore"), ]
head(df_5countries)
```

```
##     X   country continent year lifeExp      pop gdpPercap
## 61 61 Australia   Oceania 1952   69.12  8691212  10039.60
## 62 62 Australia   Oceania 1957   70.33  9712569  10949.65
## 63 63 Australia   Oceania 1962   70.93 10794968  12217.23
## 64 64 Australia   Oceania 1967   71.10 11872264  14526.12
## 65 65 Australia   Oceania 1972   71.93 13177000  16788.63
## 66 66 Australia   Oceania 1977   73.49 14074100  18334.20
```

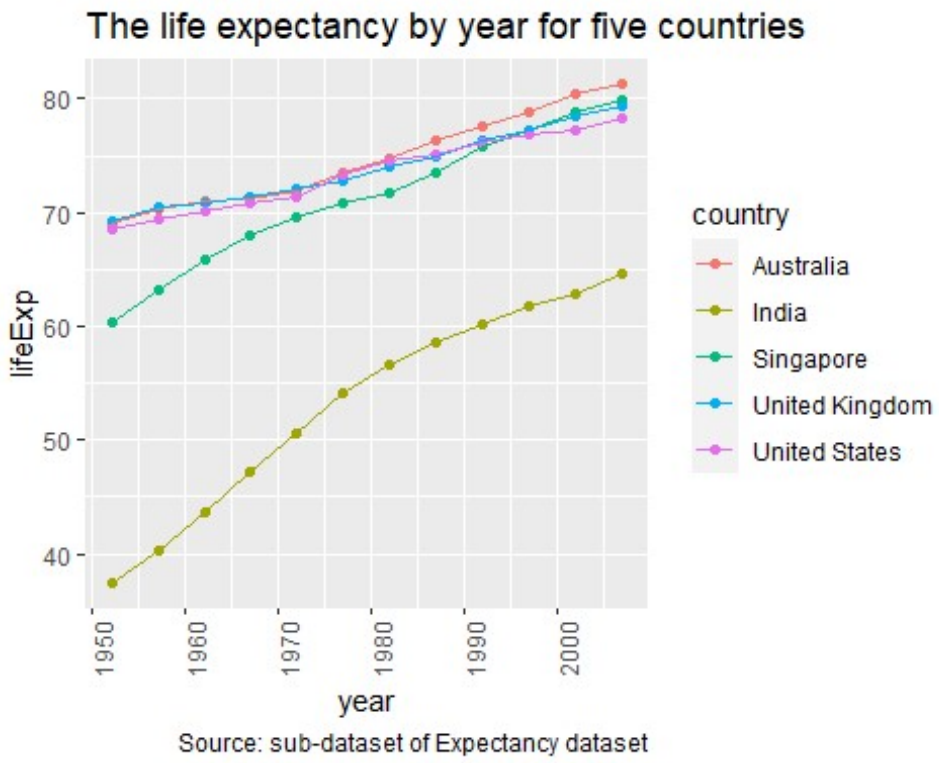
```
scatterplot<-ggplot(data=df_5countries, aes(x=year, y=lifeExp, col=country))+
geom_point()
scatterplot+labs(
    title = "The life expectancy by year for five countries",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```




```
lineplot<-ggplot(data=df_5countries, aes(x=year, y=lifeExp,color=country))+ge
om_line()
lineplot+labs(
    title = "The life expectancy by year for five countries",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```

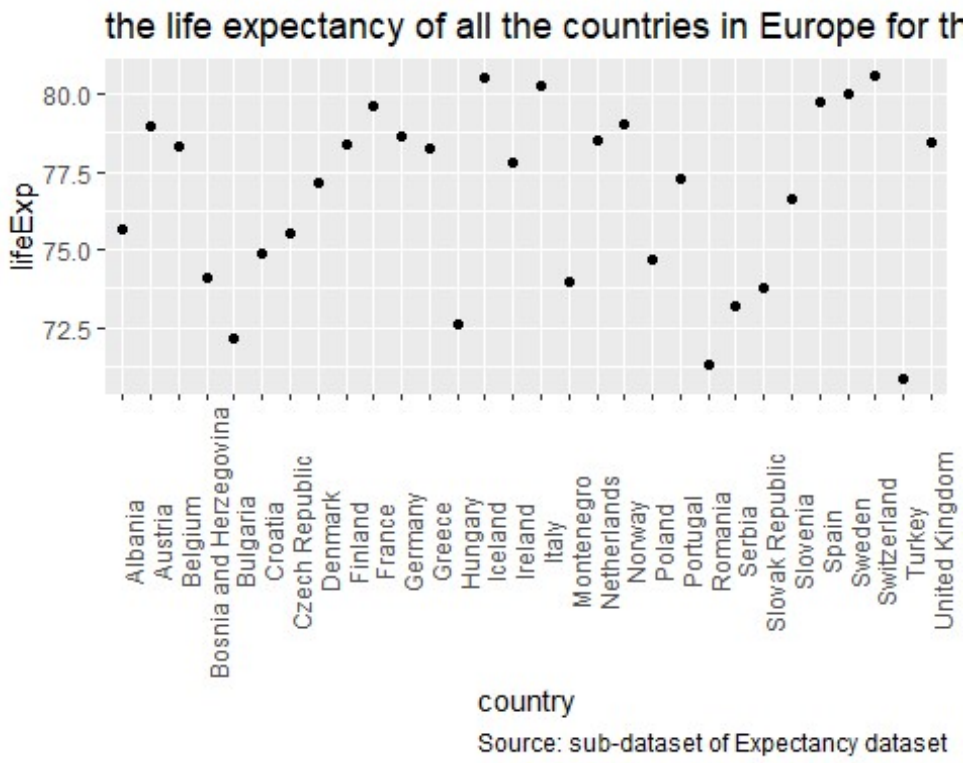



```
plot<-ggplot(data=df_5countries, aes(x=year, y=lifeExp,color=country))+geom_p
oint()+geom_line()
plot+labs(
    title = "The life expectancy by year for five countries",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```

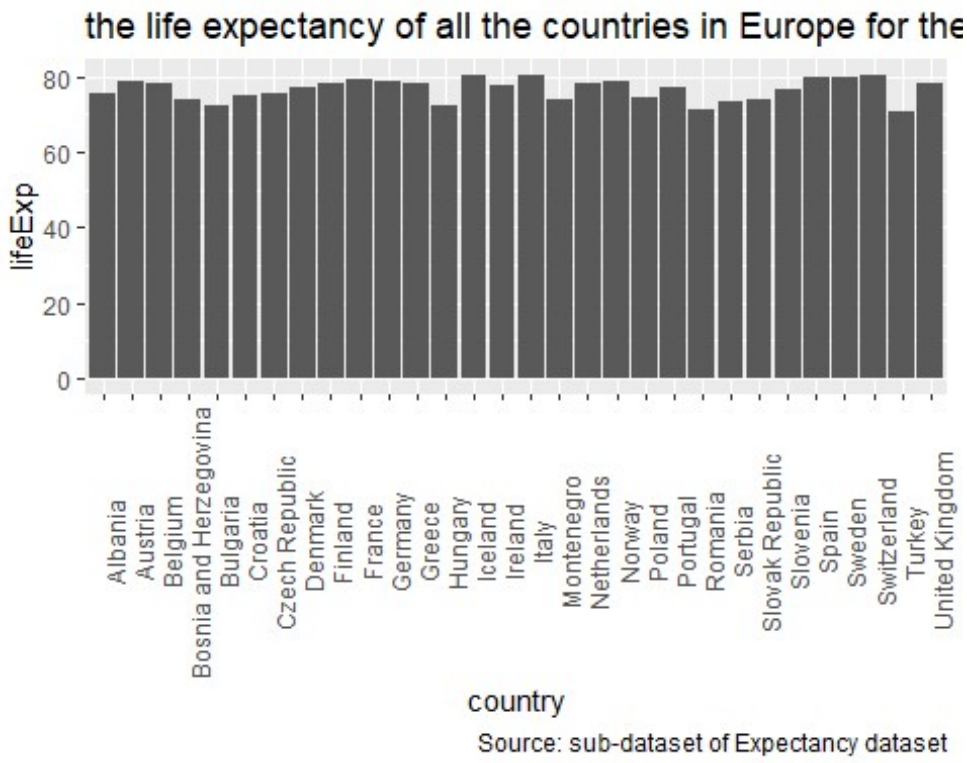


3. Plot the life expectancy of all the countries in Europe for any particular year.

```
df_europe2002 <- subset(df_expectancy,continent =="Europe"&year==2002)
View(df_europe2002)

scatterplot<-ggplot(data=df_europe2002, aes(x=country, y=lifeExp))+geom_point
()
scatterplot+labs(
    title = "the life expectancy of all the countries in Europe for the year
2002",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```




```
plot<-ggplot(data=df_europe2002, aes(x=country, y=lifeExp))+geom_bar(stat="id
entity",las=2)

## Warning: Ignoring unknown parameters: las

plot+labs(
    title = "the life expectancy of all the countries in Europe for the year
2002",
    caption = "Source: sub-dataset of Expectancy dataset"
)+theme(axis.text.x = element_text(angle=90))
```


the life expectancy of all the countries in Europe for the
lifeExp
80
60
40
20
0
Albania
Austria
Belgium
Bosnia and Herzegovina
Bulgaria
Croatia
Czech Republic
Denmark
Finland
France
Germany
Greece
Hungary
Iceland
Ireland
Italy
Montenegro
Netherlands
Norway
Poland
Portugal
Romania
Serbia
Slovak Republic
Slovenia
Spain
Sweden
Switzerland
Turkey
United Kingdom
country
Source: sub-dataset of Expectancy dataset