Assignment 1

Shubhang Periwal 19201104

10/8/2019

```
# question 1
setwd("D://r programming")
data <- read.csv("EurostatCrime2017.csv", header = TRUE, row.names = 1)

# question 2 number of rows and columns
dim(data)
```

```
## [1] 41 11
```

```
# structure of data set
str(data)
```

```
## 'data.frame':    41 obs. of  11 variables:
##  $ Intentional_homicide                                : num  1.7 1.34 0.62 1.06 0.89 2.2 0.86 0.7
##  $ Attempted_intentional_homicide                      : num  8.47 0.44 0.72 3.69 2.18 1.22 0.27 1
##  $ Assault                                             : num  611 39.6 45.1 33.1 166.1 ...
##  $ Kidnapping                                          : num  10.31 1.44 0.16 NA 5.6 ...
##  $ Sexual.violence                                     : num  63.22 9.19 13.37 83.41 42.19 ...
##  $ Robbery                                             : num  167 21.9 15 35.5 47.1 ...
##  $ Burglary                                            : num  NA 125 228 955 443 ...
##  $ Burglary_of_private_residential_premises            : num  NA NA 68.4 702.6 141.2 ...
##  $ Theft                                               : num  NA 452 632 3721 1401 ...
##  $ Theft_of_a_motorized_land_vehicle                   : num  NA 33.36 201.84 3.79 65.58 ...
##  $ Unlawful_acts_involving_controlled_drugs_or_precursors: num  506.6 70.2 52.9 481.6 400.6 ...
```

```
# q3 i) adding a new column
columns = c("Theft", "Theft_of_a_motorized_land_vehicle", "Burglary",
    "Burglary_of_private_residential_premises")
data$All_Theft <- rowSums(data[, columns], na.rm = TRUE)
# q3 ii) Removing columns
data <- data[, !(names(data) %in% columns)]
# q4 Listing countries with missing data
country_list <- row.names(data[rowSums(is.na(data)) > 0, ])
# There are 16 countries with NA values q5 Removing countries
# with missind data frames
data <- na.omit(data)
# q6
dim(data)
```

```
## [1] 25  8
```

```
# 25 rows and 8 columns are remaining
```

R Markdown

```
# 3 most common crimes in Ireland
irish_rop <- order(-data["Ireland", ])
colnames(data[irish_rop[1:3]])
```

```
## [1] "All_Theft"
## [2] "Unlawful_acts_involving_controlled_drugs_or_precursors"
## [3] "Assault"
```

```
# Top 3 are
#'All_Theft'
#'Unlawful_acts_involving_controlled_drugs_or_precursors'
#'Assault'
# Highest overall record of offences
data$Total <- rowSums(data)
colnames(data)
```

```
## [1] "Intentional_homicide"
## [2] "Attempted_intentional_homicide"
## [3] "Assault"
## [4] "Kidnapping"
## [5] "Sexual.violence"
## [6] "Robbery"
## [7] "Unlawful_acts_involving_controlled_drugs_or_precursors"
## [8] "All_Theft"
## [9] "Total"
```

```
row.names(data[which.max(data$Total), ])
```

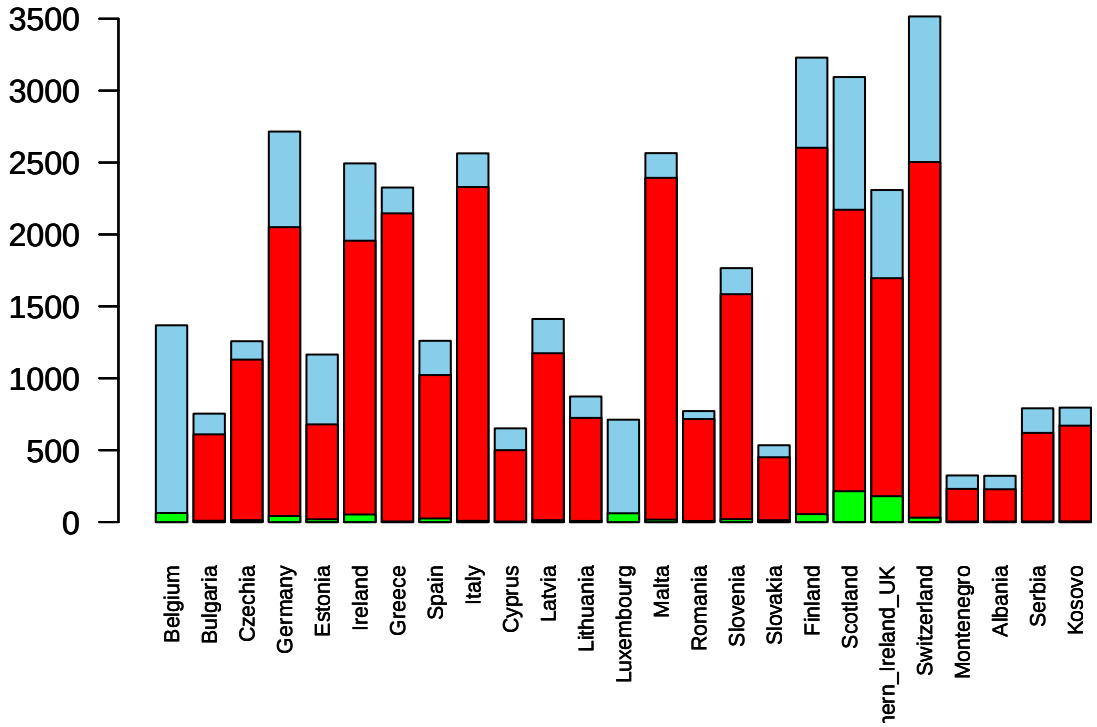
```
## [1] "Switzerland"
```

```
# The result is Switzerland
```

Including Plots

Creativity

```
barplot(names.arg = rownames(data), cex.names = 0.7, horiz = F,
    las = 2, col = "skyblue", data$Total, )
barplot(names.arg = rownames(data), cex.names = 0.7, horiz = F,
    las = 2, col = "red", data$All_Theft, add = TRUE)
barplot(names.arg = rownames(data), cex.names = 0.7, horiz = F,
    las = 2, col = "green", data$Sexual.violence, add = TRUE)
```




```
# From the bar plots we can see that thefts account for
# majority of the crime in Europe, where the minimum amount
# of theft is in Belgium and Luxembourg compared to other
# regions. We can also analyse that Scotland and UK have
# relatively high amounts of sexual violences compared to its
# other counter part countries.
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.