Case Study HDI and CPI of Countries

Ahmet Selimhan ŞAHİN

21 11 2021

Case Study

Installation of required packages

```
library(rmarkdown)
library(knitr)
library(rworldmap)
library(ggplot2)
library(dplyr)
```

```
load("F:/ders/3rd_class/IE421 Data Science for Engineers/week6/countryData.Rdata")
```

```
# col names
names(countryData)
```

```
## [1] "Country"  "HDI.Rank" "HDI"      "CPI"      "Region"
```

```
# dimensions
dim(countryData)
```

```
## [1] 173   5
```

```
# structure
str(countryData)
```

```
## 'data.frame':    173 obs. of  5 variables:
##  $ Country : Factor w/ 173 levels "Afghanistan",..: 1 2 3 4 5 6 7 8 9 10 ...
##  $ HDI.Rank: int  172 70 96 148 45 86 2 19 91 53 ...
##  $ HDI     : num  0.398 0.739 0.698 0.486 0.797 0.716 0.929 0.885 0.7 0.771 ...
##  $ CPI     : num  1.5 3.1 2.9 2 3 2.6 8.8 7.8 2.4 7.3 ...
##  $ Region  : Factor w/ 6 levels "Americas","Asia Pacific",..: 2 4 5 6 1 4 2 3 4 1 ...
```

```
#summary of data
summary(countryData)
```

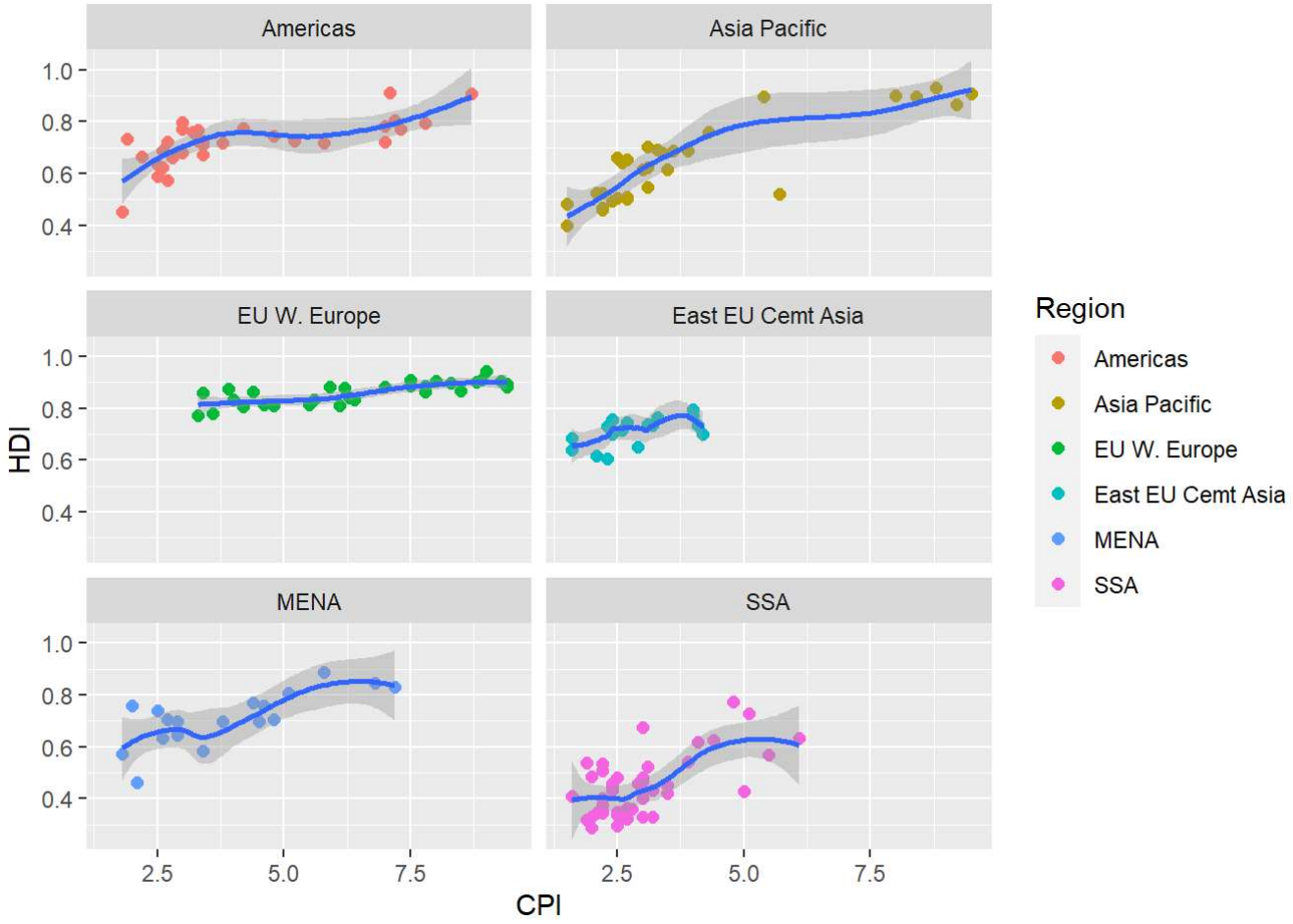
```
##         Country       HDI.Rank           HDI              CPI       
##  Afghanistan:  1   Min.   :  1.00   Min.   :0.2860   Min.   :1.500  
##  Albania    :  1   1st Qu.: 47.00   1st Qu.:0.5090   1st Qu.:2.500  
##  Algeria    :  1   Median : 96.00   Median :0.6980   Median :3.200  
##  Angola     :  1   Mean   : 95.28   Mean   :0.6581   Mean   :4.052  
##  Argentina  :  1   3rd Qu.:143.00   3rd Qu.:0.7930   3rd Qu.:5.100  
##  Armenia    :  1   Max.   :187.00   Max.   :0.9430   Max.   :9.500  
##  (Other)    :167                                                    
##                Region    

##  Americas         :31  
##  Asia Pacific     :30  
##  EU W. Europe     :30  
##  East EU Cemt Asia:18  
##  MENA             :18  
##  SSA              :46  
##
```

```
countryData %>%
  group_by(HDI) %>%
  arrange(desc(HDI)) %>%
  top_n(n = 20 , wt = HDI)
```

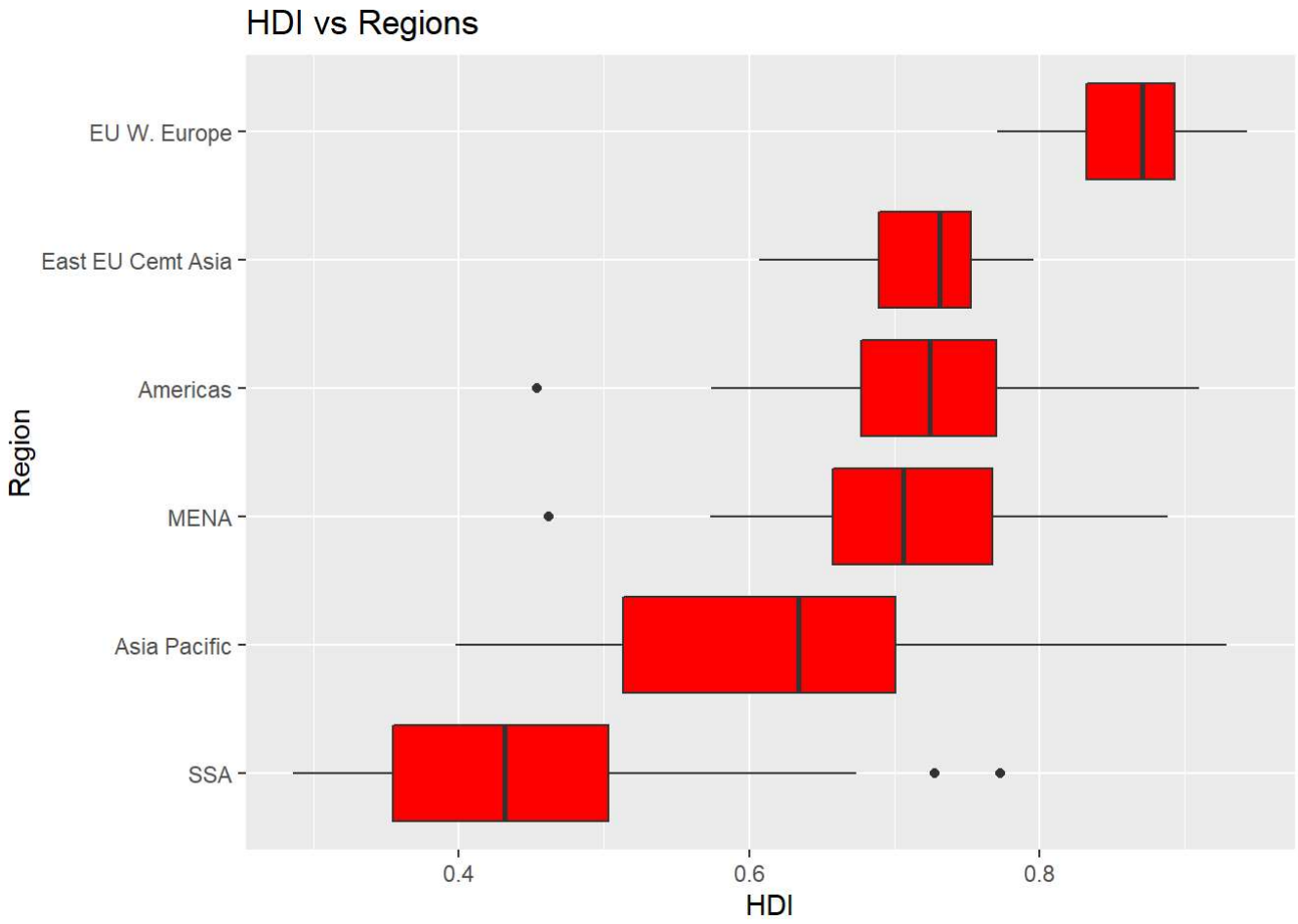
```
## # A tibble: 173 x 5
## # Groups:   HDI [148]
##    Country       HDI.Rank   HDI   CPI Region       
##    <fct>            <int> <dbl> <dbl> <fct>        
##  1 Norway               1 0.943   9   EU W. Europe 
##  2 Australia            2 0.929   8.8 Asia Pacific 
##  3 Netherlands          3 0.91    8.9 EU W. Europe 
##  4 United States        4 0.91    7.1 Americas     
##  5 Canada               6 0.908   8.7 Americas     
##  6 Ireland              7 0.908   7.5 EU W. Europe 
##  7 New Zealand          5 0.908   9.5 Asia Pacific 
##  8 Germany              9 0.905   8   EU W. Europe 
##  9 Sweden              10 0.904   9.3 EU W. Europe 
## 10 Switzerland         11 0.903   8.8 EU W. Europe 
## # ... with 163 more rows
```

```
ggplot(countryData , aes(x = CPI , y = HDI)) + geom_point(aes(color = Region) , size = 2) + g
eom_smooth() + facet_wrap(~ Region , nrow = 3)
```

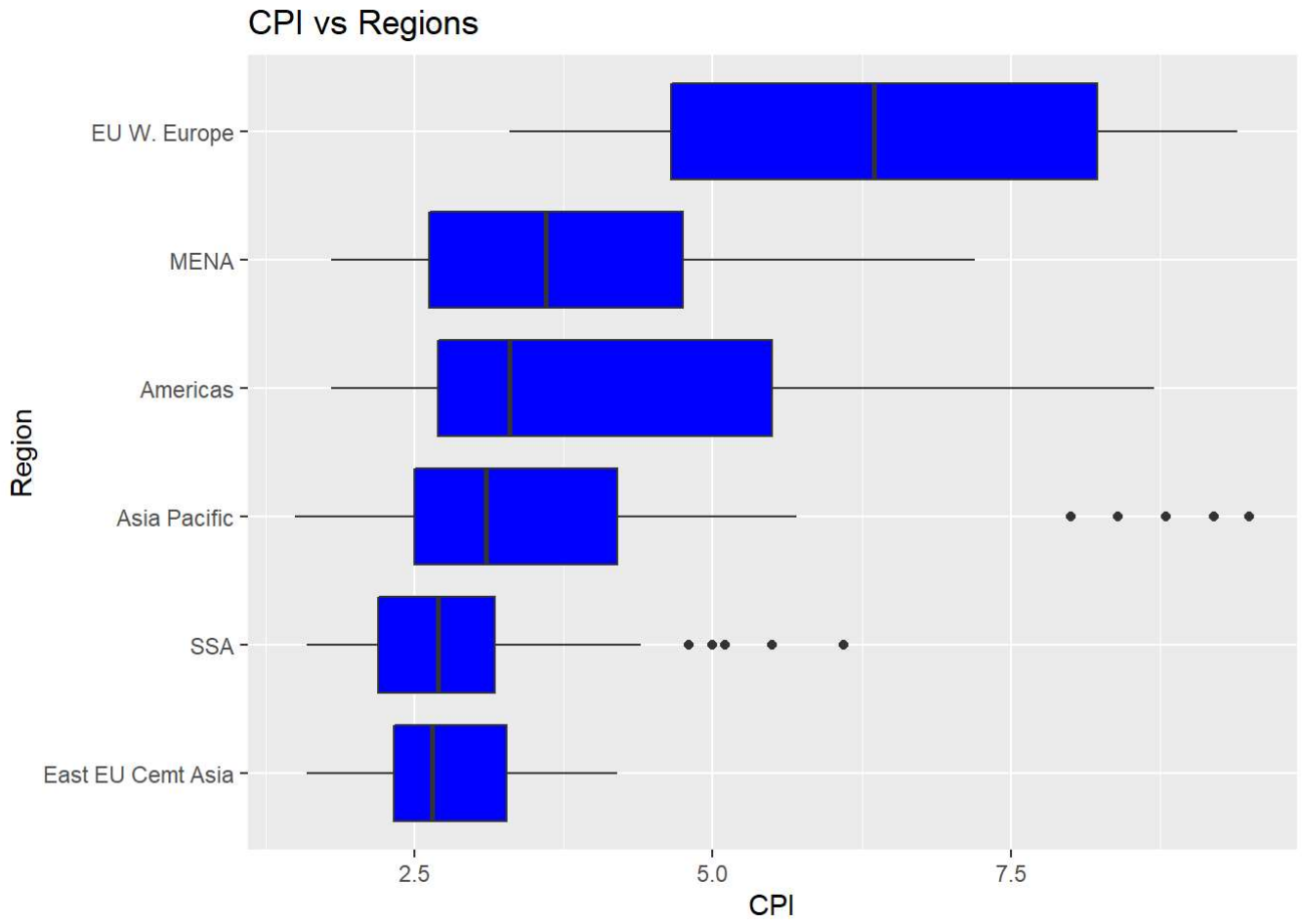
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
ggplot(countryData) + geom_boxplot(
  mapping = aes(x = reorder(Region , HDI , FUN = median ) , y = HDI) , fill = "red"
) + labs(x = "Region" , y = "HDI" , title = "HDI vs Regions") + coord_flip()
```



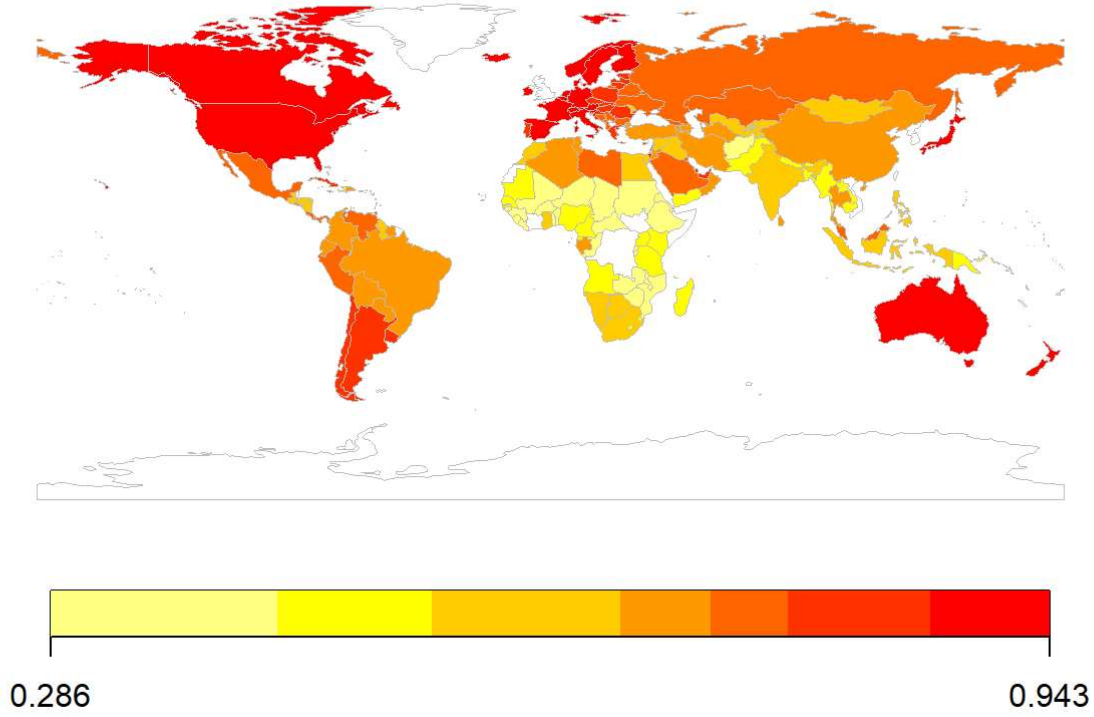

```
ggplot(countryData) + geom_boxplot(mapping = aes(
  x = reorder(Region , CPI , FUN = median) ,
  y = CPI
) , fill = "blue") + labs(x = "Region" , y = "CPI" , title = "CPI vs Regions") + coord_flip()
```




```
n <- joinCountryData2Map(countryData , joinCode = "NAME" , nameJoinColumn = "Country")
```

```
## 169 codes from your data successfully matched countries in the map
## 4 codes from your data failed to match with a country code in the map
## 74 codes from the map weren't represented in your data
```

```
mapCountryData(n , nameColumnToPlot = "HDI" , mapTitle = "Human Development Index")
```


Human Development Index
0.286
0.943