

R project: Data analysis and visualization of a publicly available dataset

Julie Pevernagie

January 2019

Human Freedom Index

About the dataset

The Human Freedom Index presents the state of human freedom in the world based on a broad measure that encompasses personal, civil, and economic freedom. Human freedom is a social concept that recognizes the dignity of individuals and is defined here as negative liberty or the absence of coercive constraint. Because freedom is inherently valuable and plays a role in human progress, it is worth measuring carefully. The Human Freedom Index is a resource that can help to more objectively observe relationships between freedom and other social and economic phenomena, as well as the ways in which the various dimensions of freedom interact with one another.

This dataset uses 79 distinct indicators of personal and economic freedom in the following areas:

Personal Freedom

- Rule of Law
- Security and Safety
- Movement
- Religion
- Association, Assembly, and Civil Society
- Expression and Information
- Identity and Relationships

Economical Freedom

- Size of Government
- Legal System and Property Rights
- Access to Sound Money
- Freedom to Trade Internationally
- Regulation of Credit, Labor, and Business

The HFI is the most comprehensive freedom index so far created for a globally meaningful set of countries. The HFI covers 162 countries for 2016, the most recent year for which sufficient data are available and it includes three countries - Belarus, Iraq, and Sudan - that were added this year. The index ranks countries beginning in 2008, the earliest year for which a robust enough index could be produced.

On a scale of 0 to 10, where 10 represents more freedom, the average human freedom rating for 162 countries in 2016 was 6.89. Among countries included in this year's and last year's report, the level of freedom decreased slightly (-0.01) compared with 2015, with 63 countries increasing their ratings and 87 decreasing. Since 2008, the level of global freedom has also decreased slightly (-0.06), with 56 countries in the index increasing their ratings and 81 decreasing.

Loading the dataset and required packages:

```
# check installed pkg
pkg <- installed.packages()[, "Package"]
loadpkg <- c("ggplot2", "rworldmap", "reshape2", "RColorBrewer", "ggplus")

library(devtools)
devtools::install_github("guiastronnec/ggplus")

# install pkg if necessary
for ( i in 1:length(loadpkg) ){
  if(!(loadpkg[i] %in% pkg)) {
    install.packages(loadpkg[i])
  }
}

# load the required packages
lapply(loadpkg, library, character.only = TRUE)

# load data for global human freedom index 2008-2016
url <- "https://github.com/Juliepvr/hello-world/raw/master/hfi_cc_2018.csv"
hfi <- read.csv(url)
```

Modifying the data for easy processing:

```
hfi$year<- as.factor(hfi$year)

# data for Belgium:
hfiBEL <- hfi[hfi$ISO_code=="BEL",]

# data for 2016:
### global:
hfi2016 <- hfi[hfi$year==2016,]
### for Belgium:
hfiBEL2016<- hfiBEL[hfiBEL$year==2016,]

# load separte text files with titles
titles <- scan(
  "https://raw.githubusercontent.com/Juliepvr/hello-world/master/hfi-titles",
  what="", sep="\n")
subt <- scan(
  "https://raw.githubusercontent.com/Juliepvr/hello-world/master/hfi-list-column-names",
  what="", sep="\n")

# execute R commands in textfile line by line:
subtitles<- c()
for(i in 1:length(subt)){
  tmp<-list(eval(parse(text=subt[i])))
  subtitles<-append(subtitles, tmp, after=length(subtitles))
}

# Filter out columns to be used from BEL 2016
```

```

skip_this <- grep( "(\\_([a-zA-Z]{1,}){3,})") , colnames(hfiBEL2016),
                  value=TRUE)
hfiBEL2016 <- hfiBEL2016[,!(colnames(hfiBEL2016) %in% skip_this)]

# totals and subcategories separate
seperate_this<-grep( "(\\_([a-zA-Z]{1,}){2,})") , colnames(hfiBEL2016),
                    value=TRUE)
totalsBEL<-hfiBEL2016[,!(colnames(hfiBEL2016) %in% seperate_this)]
totalsBEL<-totalsBEL[,c(5:11,14:18)] # cut out pf and ef totals and rank
cols<-colnames(totalsBEL)
detailsBEL<-hfiBEL2016[,seperate_this]

for(i in 1:length(cols)){
  assign(cols[i], detailsBEL[,grep(paste(cols[i], "_*", sep = ""),
                                   colnames(detailsBEL))])
}

# picking a color palette for uniform plotting, custom colors for better
# representation of the data.

clr <- c("black",brewer.pal(11,"RdGy")[c(11,10,1)], brewer.pal(6,"RdYlGn"))

```

Worldwide Freedom 2016

```
# Worldmap with totals for 2016

# create SpatialPolygonsDataFrame: match countries between data and package
sPDF <- joinCountryData2Map( hfi2016
                             , joinCode = "ISO3"
                             , nameJoinColumn = "ISO_code")

## 161 codes from your data successfully matched countries in the map
## 1 codes from your data failed to match with a country code in the map
## 82 codes from the map weren't represented in your data

# draw map
mapParams <- mapCountryData(sPDF, nameColumnToPlot="hf_score",
                             mapTitle="Human Freedom Index 2016",
                             catMethod=0:10,
                             colourPalette=clr,
                             addLegend=FALSE )

# draw legend
do.call( addMapLegend,
         c(mapParams,
           legendLabels="all",
           legendWidth=0.5,
           legendIntervals="page",
           legendMar = 4 ))
```

A world map showing the distribution of the world's population by country. The map is color-coded to represent population density, with colors ranging from light yellow (low density) to dark red (high density). High-density areas (dark red) are concentrated in Europe, India, and East Asia. Moderate density (orange/red) is seen in North America, China, and parts of Europe and Africa. Low density (light yellow) is prevalent in Russia, Canada, and Australia. The map also shows the outlines of the continents and major bodies of water.



Distribution of Freedom by Region

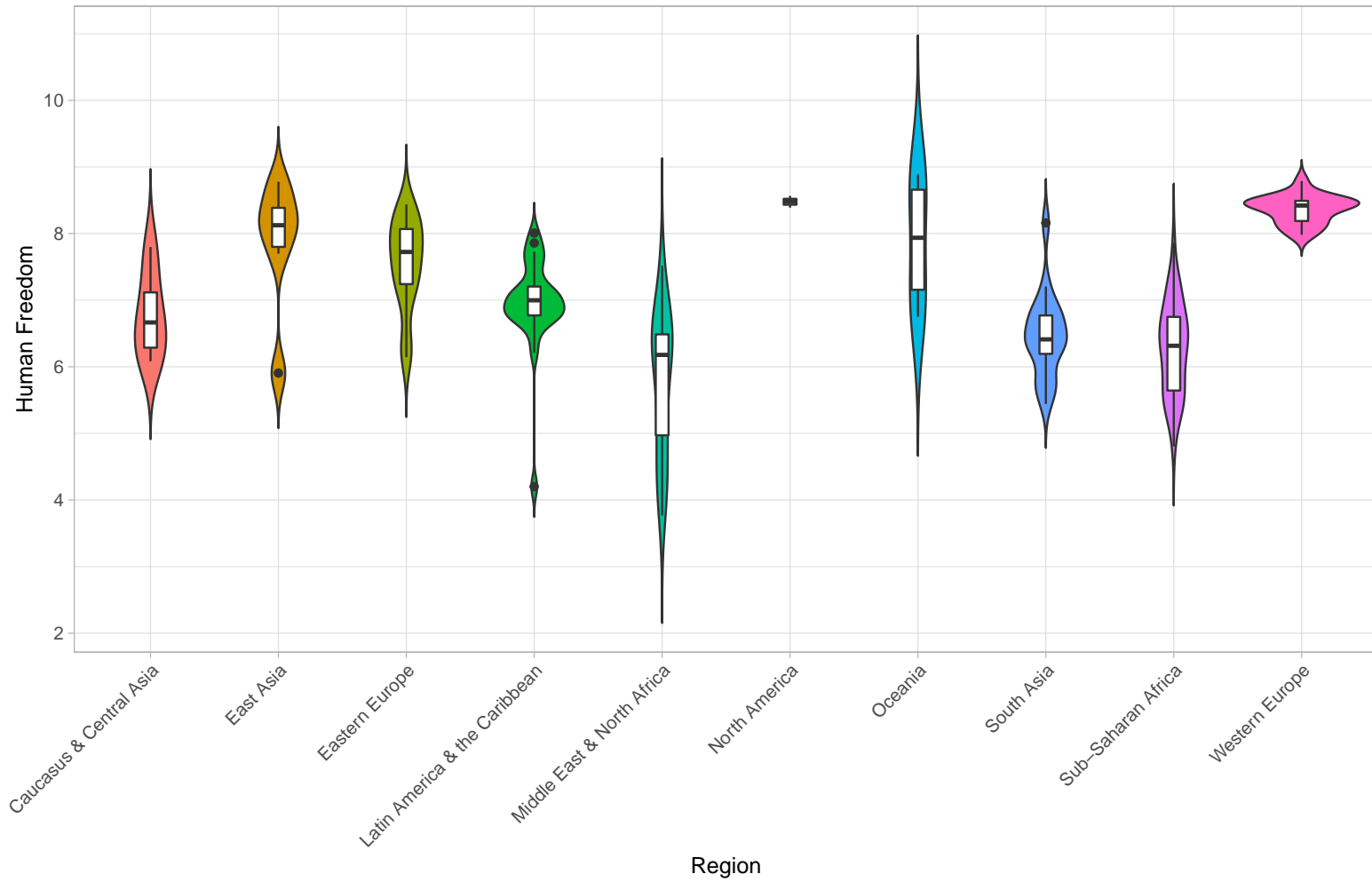
The violin plots show that western countries enjoy more freedom and the scores are more uniform amongst the countries, as opposed to the Middle East with pronounced outliers. Since North America consist of only 2 countries (Canada and USA), there is not enough data to generate a violinplot.

```
# Violin plots + boxplots per region

viobox <- ggplot(hfi2016, aes(x=hfi2016$region, y=hfi2016$hf_score,
                             fill=hfi2016$region)) +
  geom_violin(trim=FALSE, show.legend = FALSE) +
  geom_boxplot(width=0.1, fill="white") +
  labs(title="Human Freedom by Region", x="Region", y="Human Freedom")

viobox + theme_light() + theme(axis.text.x=element_text(angle=45, hjust=1),
                               plot.margin=margin(1, 1, 1, 2, "cm"),
                               plot.title = element_text(hjust = 0.5))
```

Human Freedom by Region



Belgium

The following graph illustrates the evolution of personal freedom, economical freedom and the resulting human freedom in Belgium since 2008.

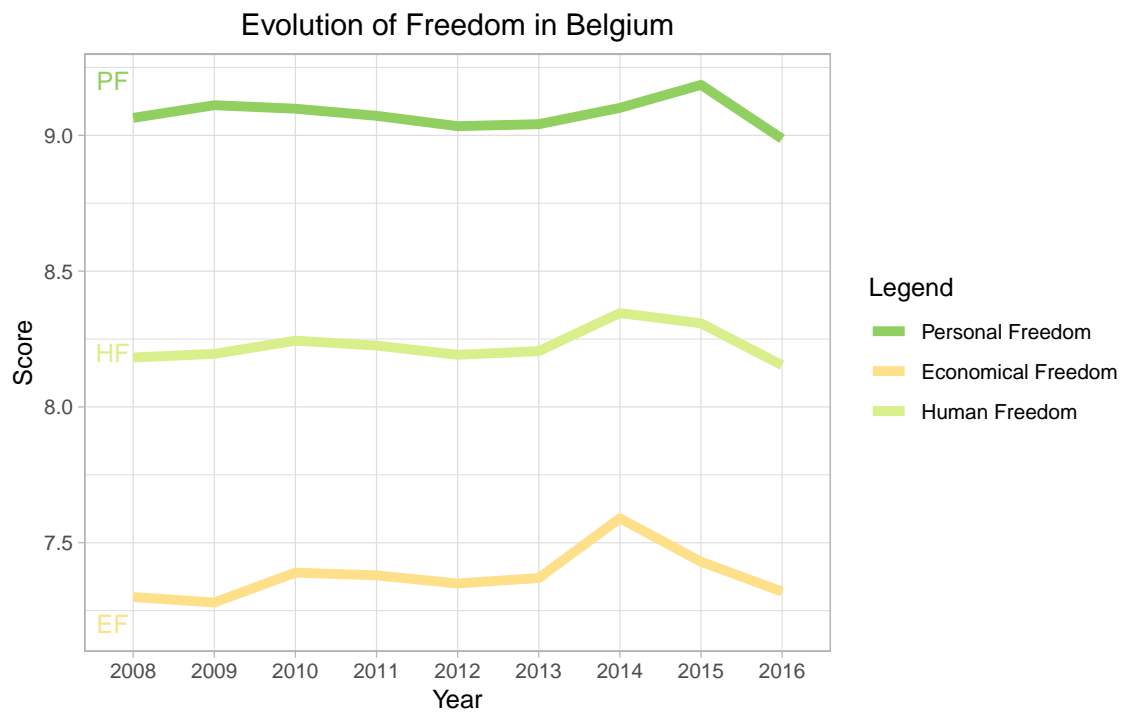
```
# Overview Belgium 2008-2016

summBEL <- data.frame(hfiBEL$year, hfiBEL$pf_score, hfiBEL$ef_score, hfiBEL$hf_score)
evolBEL <- data.frame(hfiBEL$year, hfiBEL$hf_score)
summBEL <- melt(summBEL)
evolBEL <- melt(evolBEL)
levels(summBEL$variable) <- c("Personal Freedom", "Economical Freedom",
                             "Human Freedom")

a <- round(mean(hfiBEL$pf_score, digit=0))
b <- round(mean(hfiBEL$ef_score, digit=0))
c <- round(mean(hfiBEL$hf_score, digit=0))

p1 <- ggplot(summBEL, aes(summBEL$hfiBEL.year, summBEL$value,
                          fill = summBEL$variable))

# graph evolution areas of freedom
p1 + geom_line(aes(color = summBEL$variable), group = summBEL$variable, size=2) +
  labs(title="Evolution of Freedom in Belgium", x="Year", y="Score") +
  guides(color=guide_legend("Legend")) +
  scale_color_manual(values=clr[c(a,b,c)]) +
  theme_light() +
  theme(plot.margin=margin(1, 1, 1, 1, "cm"),
        plot.title = element_text(hjust = 0.5)) +
  annotate("text", x=.75, y=c(a,b,c)+.2, label = c("PF", "EF", "HF"),
          color=clr[c(a,b,c)])
```

Details

```
# Details Belgium
BELdata <- list()

for(i in 1:length(cols)) {
  BELdata[[i]] <- eval(parse(text=cols[i]))
}

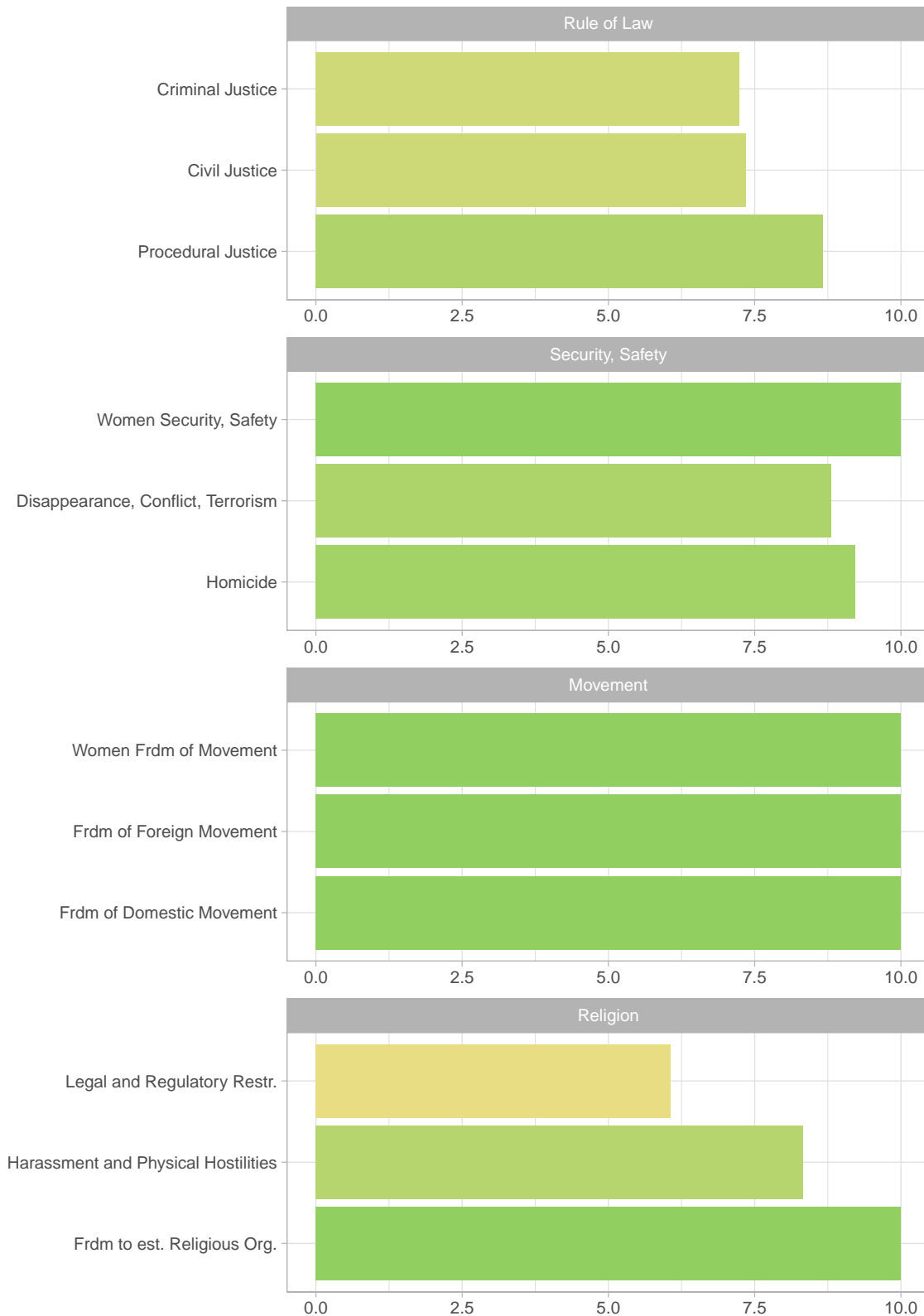
BELdata<-melt(BELdata)
BELdata$variable<- factor(BELdata$variable, levels = BELdata$variable, labels =
  unlist(subtitles))
BELdata$L1<- factor(BELdata$L1, levels = 1:12, labels = titles)

p2 <-ggplot(BELdata, value.name=titles,aes(x=variable,y=value, fill=value))+
  geom_col(position="dodge")+
  ylim(0,10) +
  coord_flip()+
  scale_fill_gradient2( midpoint=5, low=clr[5], mid=clr[7],high=clr[9]) +
  labs(title="Areas of Freedom for Belgium in 2016",x=element_blank(),
    y=element_blank()) +
  theme_light() +
  theme(plot.title = element_text(hjust = 0.5),legend.position="none")

# Plot on multiple pages (ggplus package from github)
facet_multiple(plot = p2,
  facets = "L1",
  scales = "free",
  ncol = 1,
  nrow = 4)
```

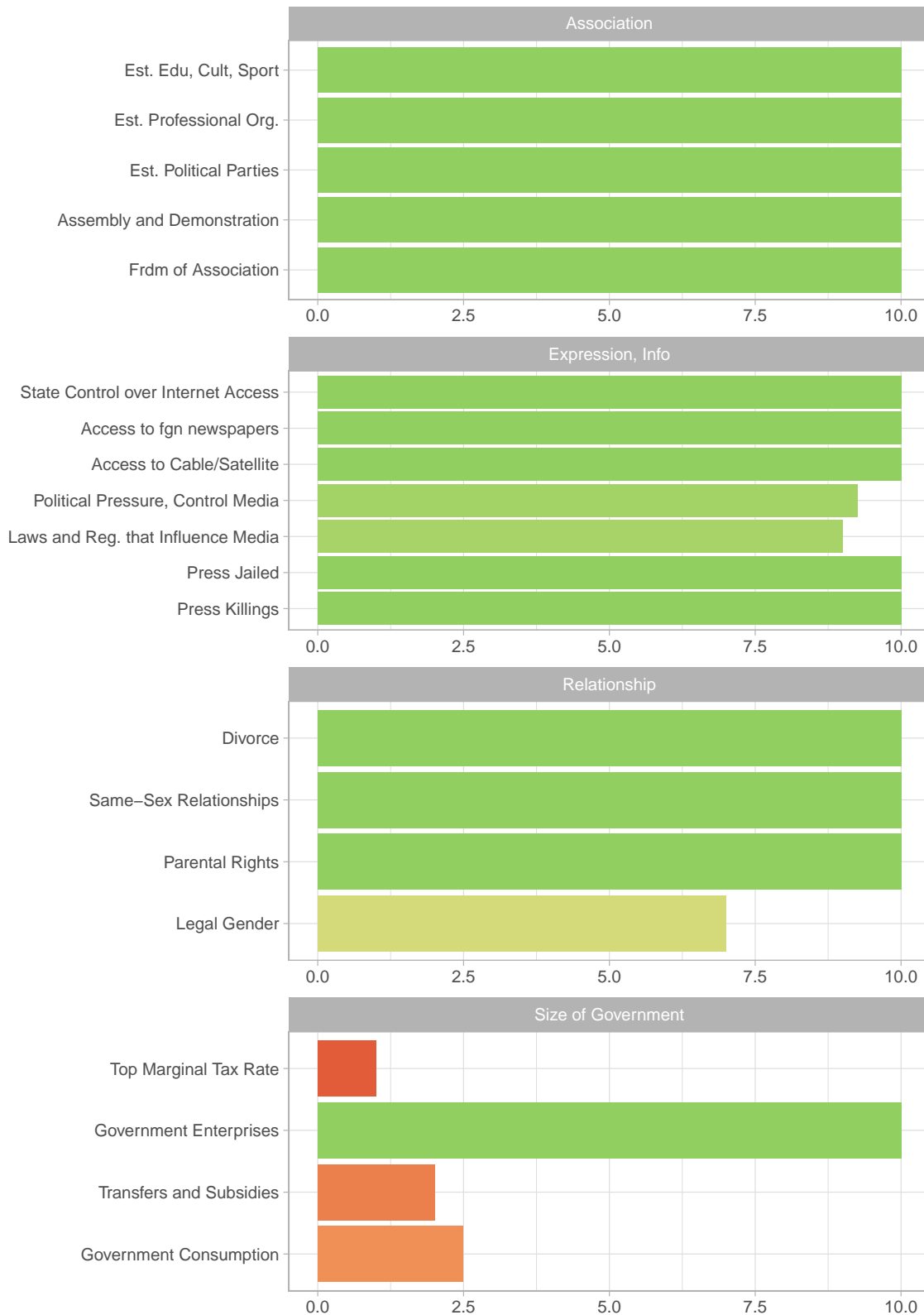
Areas of Freedom for Belgium in 2016

Page 1 of 3



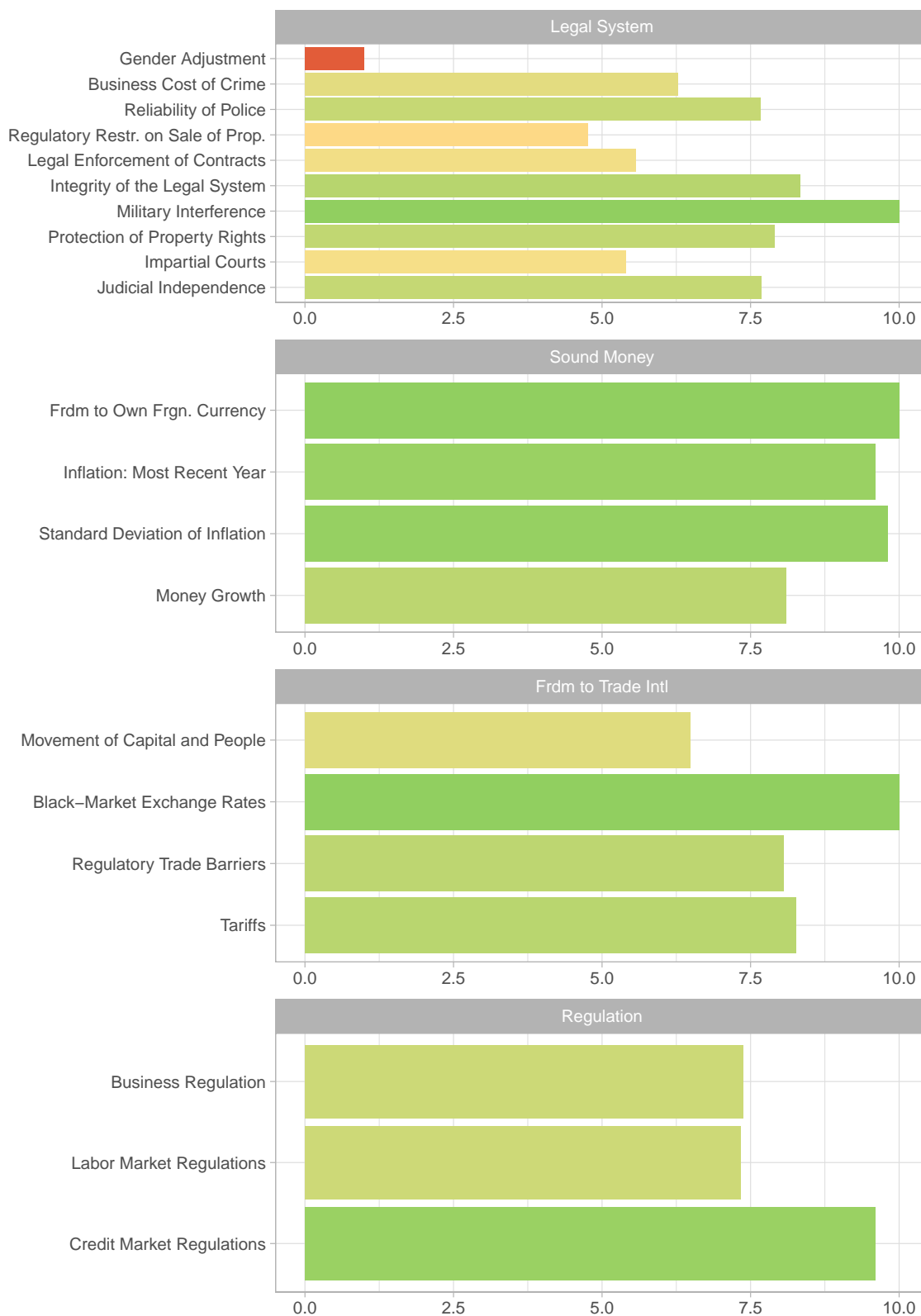
Areas of Freedom for Belgium in 2016

Page 2 of 3



Areas of Freedom for Belgium in 2016

Page 3 of 3



Women's freedom of movement

One of the components of the human freedom index measures the freedom of women to move outside the home. The following elements were considered: ability to choose their places of residence, to visit their families and friends, or to apply for a passport. A rating of 10 was assigned to countries with no restrictions on women's movement outside the home. A rating of 5 was assigned to countries where (some) women can leave home sometimes but with restrictions. A rating of 0 was assigned to countries where women can never leave home without restrictions (e.g., they need a male companion).

```
# map women

hfi2016$pf_movement_women <- as.factor(hfi2016$pf_movement_women)
levels(hfi2016$pf_movement_women) <- c("none", "limited", "all")

# create SpatialPolygonsDataFrame: match countries between data and package
sPDF <- joinCountryData2Map( hfi2016
                             , joinCode = "ISO3"
                             , nameJoinColumn = "ISO_code")

## 161 codes from your data successfully matched countries in the map
## 1 codes from your data failed to match with a country code in the map
## 82 codes from the map weren't represented in your data

# draw map
mapParams <- mapCountryData(sPDF, nameColumnToPlot="pf_movement_women",
                             mapTitle="Women's freedom by law to move outside the
                             home 2016",
                             catMethod="categorical",
                             colourPalette=clr[c(5,7,9)])
```

A world map showing the distribution of the 'category' variable across countries. The legend indicates three categories: 'none' (red), 'limited' (orange), and 'all' (green). The map shows that the 'none' category is concentrated in the Middle East and parts of Africa. The 'limited' category is found in North Africa, the Middle East, and parts of Asia and South America. The 'all' category is the most widespread, covering North America, South America, Europe, Russia, China, India, and Australia.