

# Projet ADD : Rashid Rafi, Arakchou Ayour

M1 IM

## Présentation

Le Human Freedom Index (HFI) représente une échelle de 0 à 10 visant à mesurer la liberté humaine parmi les pays du monde entier. Ces mesures prennent en compte plusieurs aspects de la liberté : liberté personnelle, civile, économique. Cet index représente une mesure de la liberté humaine, dans le sens d'absence de contrainte coercitive. Le HFI 2019 est la 5ème édition (le premier étant de 2015), et comporte 76 indicateurs distincts de la liberté personnelle et économique dans les domaines suivants :

- Rule of Law
- Security and Safety
- Movement
- Religion
- Association, Assembly, and Civil Society
- Expression and Information
- Identity and Relationships
- Size of Government
- Legal System and Property Rights
- Access to Sound Money
- Freedom to Trade Internationally
- Regulation of Credit, Labor, and Business

## But, Objectifs

Un des objectifs du HFI est de visualiser la liberté humaine de manière générale dans le monde, suivant les différents pays. Un autre serait de comprendre ce qu'est la liberté humaine, mieux la définir, et regarder sa relation avec les autres phénomènes économiques et sociaux. Finalement, est-ce le HFI est un bon index ? On pourrait au final faire un résumé avec un classement, puis essayer d'expliquer ce classement.

## Import du dataset et selection de nos variables

```
data <- read_csv("hfi_cc_2018.csv")
head(data)
```

```
## # A tibble: 6 x 123
##   year ISO_code countries region pf_rol_procedur... pf_rol_civil pf_rol_criminal
##   <dbl> <chr>    <chr>    <chr>          <dbl>          <dbl>          <dbl>
## 1  2016 ALB      Albania  Easte...      6.66           4.55           4.67
## 2  2016 DZA      Algeria  Middl...      NA             NA             NA
## 3  2016 AGO      Angola   Sub-S...      NA             NA             NA
## 4  2016 ARG      Argentina Latin...  7.10           5.79           4.34
## 5  2016 ARM      Armenia  Cauca...      NA             NA             NA
## 6  2016 AUS      Australia Ocean...  8.44           7.53           7.36
## # ... with 116 more variables: pf_rol <dbl>, pf_ss_homicide <dbl>,
## #   pf_ss_disappearances_disap <dbl>, pf_ss_disappearances_violent <dbl>,
## #   pf_ss_disappearances_organized <dbl>,
## #   pf_ss_disappearances_fatalities <dbl>, pf_ss_disappearances_injuries <dbl>,
## #   pf_ss_disappearances <dbl>, pf_ss_women_fgm <dbl>,
## #   pf_ss_women_missing <dbl>, pf_ss_women_inheritance_widows <dbl>,
## #   pf_ss_women_inheritance_daughters <dbl>, pf_ss_women_inheritance <dbl>,
## #   pf_ss_women <dbl>, pf_ss <dbl>, pf_movement_domestic <dbl>,
## #   pf_movement_foreign <dbl>, pf_movement_women <dbl>, pf_movement <dbl>,
## #   pf_religion_estop_establish <dbl>, pf_religion_estop_operate <dbl>,
## #   pf_religion_estop <dbl>, pf_religion_harassment <dbl>,
## #   pf_religion_restrictions <dbl>, pf_religion <dbl>,
## #   pf_association_association <dbl>, pf_association_assembly <dbl>,
## #   pf_association_political_establish <dbl>,
## #   pf_association_political_operate <dbl>, pf_association_political <dbl>,
## #   pf_association_prof_establish <dbl>, pf_association_prof_operate <dbl>,
## #   pf_association_prof <dbl>, pf_association_sport_establish <dbl>,
## #   pf_association_sport_operate <dbl>, pf_association_sport <dbl>,
## #   pf_association <dbl>, pf_expression_killed <dbl>,
## #   pf_expression_jailed <dbl>, pf_expression_influence <dbl>,
## #   pf_expression_control <dbl>, pf_expression_cable <dbl>,
## #   pf_expression_newspapers <dbl>, pf_expression_internet <dbl>,
## #   pf_expression <dbl>, pf_identity_legal <dbl>,
## #   pf_identity_parental_marriage <dbl>, pf_identity_parental_divorce <dbl>,
## #   pf_identity_parental <dbl>, pf_identity_sex_male <dbl>,
## #   pf_identity_sex_female <dbl>, pf_identity_sex <dbl>,
## #   pf_identity_divorce <dbl>, pf_identity <dbl>, pf_score <dbl>,
## #   pf_rank <dbl>, ef_government_consumption <dbl>,
## #   ef_government_transfers <dbl>, ef_government_enterprises <dbl>,
## #   ef_government_tax_income <dbl>, ef_government_tax_payroll <dbl>,
## #   ef_government_tax <dbl>, ef_government <dbl>, ef_legal_judicial <dbl>,
## #   ef_legal_courts <dbl>, ef_legal_protection <dbl>, ef_legal_military <dbl>,
## #   ef_legal_integrity <dbl>, ef_legal_enforcement <dbl>,
## #   ef_legal_restrictions <dbl>, ef_legal_police <dbl>, ef_legal_crime <dbl>,
## #   ef_legal_gender <dbl>, ef_legal <dbl>, ef_money_growth <dbl>,
## #   ef_money_sd <dbl>, ef_money_inflation <dbl>, ef_money_currency <dbl>,
## #   ef_money <dbl>, ef_trade_tariffs_revenue <dbl>,
## #   ef_trade_tariffs_mean <dbl>, ef_trade_tariffs_sd <dbl>,
## #   ef_trade_tariffs <dbl>, ef_trade_regulatory_nontariff <dbl>,
## #   ef_trade_regulatory_compliance <dbl>, ef_trade_regulatory <dbl>,
## #   ef_trade_black <dbl>, ef_trade_movement_foreign <dbl>,
## #   ef_trade_movement_capital <dbl>, ef_trade_movement_visit <dbl>,
## #   ef_trade_movement <dbl>, ef_trade <dbl>,
## #   ef_regulation_credit_ownership <dbl>, ef_regulation_credit_private <dbl>,
## #   ef_regulation_credit_interest <dbl>, ef_regulation_credit <dbl>.
```

```
## #   ef_regulation_labor_minwage <dbl>, ef_regulation_labor_firing <dbl>,
## #   ef_regulation_labor_bargain <dbl>, ef_regulation_labor_hours <dbl>, ...
```

Notre jeu de donnée n'est pas très bien présenté : nous avons tous les pays en 2016, puis ces mêmes pays en 2015, etc. Puisque l'unité expérimentale ici est un pays sur une année précise, on veut réarranger notre jeu de façon à avoir tous les pays en ordre alphabétique, pour une meilleur visibilité.

```
data2 <- data[order(data$countries),]
head(data2)
```

```
## # A tibble: 6 x 123
##   year ISO_code countries region pf_rol_procedur... pf_rol_civil pf_rol_criminal
##   <dbl> <chr>    <chr>    <chr>          <dbl>          <dbl>          <dbl>
## 1  2016 ALB      Albania  Easte...        6.66           4.55           4.67
## 2  2015 ALB      Albania  Easte...        6.11           4.80           4.58
## 3  2014 ALB      Albania  Easte...        5.57           5.04           4.30
## 4  2013 ALB      Albania  Easte...         5             4.9            3.6
## 5  2012 ALB      Albania  Easte...         5             4.9            3.6
## 6  2011 ALB      Albania  Easte...        5.67           5.07           4.10
## # ... with 116 more variables: pf_rol <dbl>, pf_ss_homicide <dbl>,
## #   pf_ss_disappearances_disap <dbl>, pf_ss_disappearances_violent <dbl>,
## #   pf_ss_disappearances_organized <dbl>,
## #   pf_ss_disappearances_fatalities <dbl>, pf_ss_disappearances_injuries <dbl>,
## #   pf_ss_disappearances <dbl>, pf_ss_women_fgm <dbl>,
## #   pf_ss_women_missing <dbl>, pf_ss_women_inheritance_widows <dbl>,
## #   pf_ss_women_inheritance_daughters <dbl>, pf_ss_women_inheritance <dbl>,
## #   pf_ss_women <dbl>, pf_ss <dbl>, pf_movement_domestic <dbl>,
## #   pf_movement_foreign <dbl>, pf_movement_women <dbl>, pf_movement <dbl>,
## #   pf_religion_estop_establish <dbl>, pf_religion_estop_operate <dbl>,
## #   pf_religion_estop <dbl>, pf_religion_harassment <dbl>,
## #   pf_religion_restrictions <dbl>, pf_religion <dbl>,
## #   pf_association_association <dbl>, pf_association_assembly <dbl>,
## #   pf_association_political_establish <dbl>,
## #   pf_association_political_operate <dbl>, pf_association_political <dbl>,
## #   pf_association_prof_establish <dbl>, pf_association_prof_operate <dbl>,
## #   pf_association_prof <dbl>, pf_association_sport_establish <dbl>,
## #   pf_association_sport_operate <dbl>, pf_association_sport <dbl>,
## #   pf_association <dbl>, pf_expression_killed <dbl>,
## #   pf_expression_jailed <dbl>, pf_expression_influence <dbl>,
## #   pf_expression_control <dbl>, pf_expression_cable <dbl>,
## #   pf_expression_newspapers <dbl>, pf_expression_internet <dbl>,
## #   pf_expression <dbl>, pf_identity_legal <dbl>,
## #   pf_identity_parental_marriage <dbl>, pf_identity_parental_divorce <dbl>,
## #   pf_identity_parental <dbl>, pf_identity_sex_male <dbl>,
## #   pf_identity_sex_female <dbl>, pf_identity_sex <dbl>,
## #   pf_identity_divorce <dbl>, pf_identity <dbl>, pf_score <dbl>,
## #   pf_rank <dbl>, ef_government_consumption <dbl>,
## #   ef_government_transfers <dbl>, ef_government_enterprises <dbl>,
## #   ef_government_tax_income <dbl>, ef_government_tax_payroll <dbl>,
## #   ef_government_tax <dbl>, ef_government <dbl>, ef_legal_judicial <dbl>,
## #   ef_legal_courts <dbl>, ef_legal_protection <dbl>, ef_legal_military <dbl>,
## #   ef_legal_integrity <dbl>, ef_legal_enforcement <dbl>,
## #   ef_legal_restrictions <dbl>, ef_legal_police <dbl>, ef_legal_crime <dbl>,
```

```
## # ef_legal_gender <dbl>, ef_legal <dbl>, ef_money_growth <dbl>,
## # ef_money_sd <dbl>, ef_money_inflation <dbl>, ef_money_currency <dbl>,
## # ef_money <dbl>, ef_trade_tariffs_revenue <dbl>,
## # ef_trade_tariffs_mean <dbl>, ef_trade_tariffs_sd <dbl>,
## # ef_trade_tariffs <dbl>, ef_trade_regulatory_nontariff <dbl>,
## # ef_trade_regulatory_compliance <dbl>, ef_trade_regulatory <dbl>,
## # ef_trade_black <dbl>, ef_trade_movement_foreign <dbl>,
## # ef_trade_movement_capital <dbl>, ef_trade_movement_visit <dbl>,
## # ef_trade_movement <dbl>, ef_trade <dbl>,
## # ef_regulation_credit_ownership <dbl>, ef_regulation_credit_private <dbl>,
## # ef_regulation_credit_interest <dbl>, ef_regulation_credit <dbl>,
## # ef_regulation_labor_minwage <dbl>, ef_regulation_labor_firing <dbl>,
## # ef_regulation_labor_bargain <dbl>, ef_regulation_labor_hours <dbl>, ...
```

Précisons la signification de nos variables : les préfixes **pf** pour Personal Freedom, **ef** pour Economic Freedom, et **hf** pour Fuman Freedom.

Nous avons beaucoup de variables (123 variables), nous avons pensé à faire un pivot\_longer sur nos variables pf\_, ef\_, et hf\_ mais nous remarquons quelque chose : les variables pf\_() sont des moyennes des variables pf\_()\_(). Regardons cela de plus près :

```
new <- data2 %>%
  mutate(test = (pf_rol_procedural + pf_rol_civil + pf_rol_criminal)/3)
head(new$test)
```

```
## [1] 5.291752 5.164786 4.971520 4.500000 4.500000 4.947322
```

```
head(data2$pf_rol)
```

```
## [1] 5.291752 5.164786 4.971520 4.493651 4.493651 4.900000
```

Nous avons bien (à quelques arrondis près) la même colonne pf\_rol si on moyenne 3 sous-catégories de Rule Of Law. Nous décidons alors de travailler uniquement sur ces moyennes, donc sur ces variables qui décrivent les domaines cités plus haut. Réarrangeons notre dataset :

```
data3 <- data2 %>%
  select(countries,
         region,
         year,
         pf_rol,
         pf_ss,
         pf_movement,
         pf_religion,
         pf_association,
         pf_expression,
         pf_identity,
         pf_score,
         pf_rank,
         ef_government,
         ef_legal,
         ef_money,
         ef_trade,
```

```

ef_regulation,
ef_score,
ef_rank,
hf_score,
hf_rank
)

```

Nous avons ici enlevé l'ISOCode puisqu'il n'est pas utile à garder, mais nous gardons la variable **région** pour ensuite faire une étude qualitative en fonction des régions du monde.

## Personal Freedom vs Economic Freedom

La première chose que l'on constate est que le jeu de donnée est regroupé en 2 aspect de la liberté : la liberté personnelle et économique. Comparons alors ces 2 libertés.

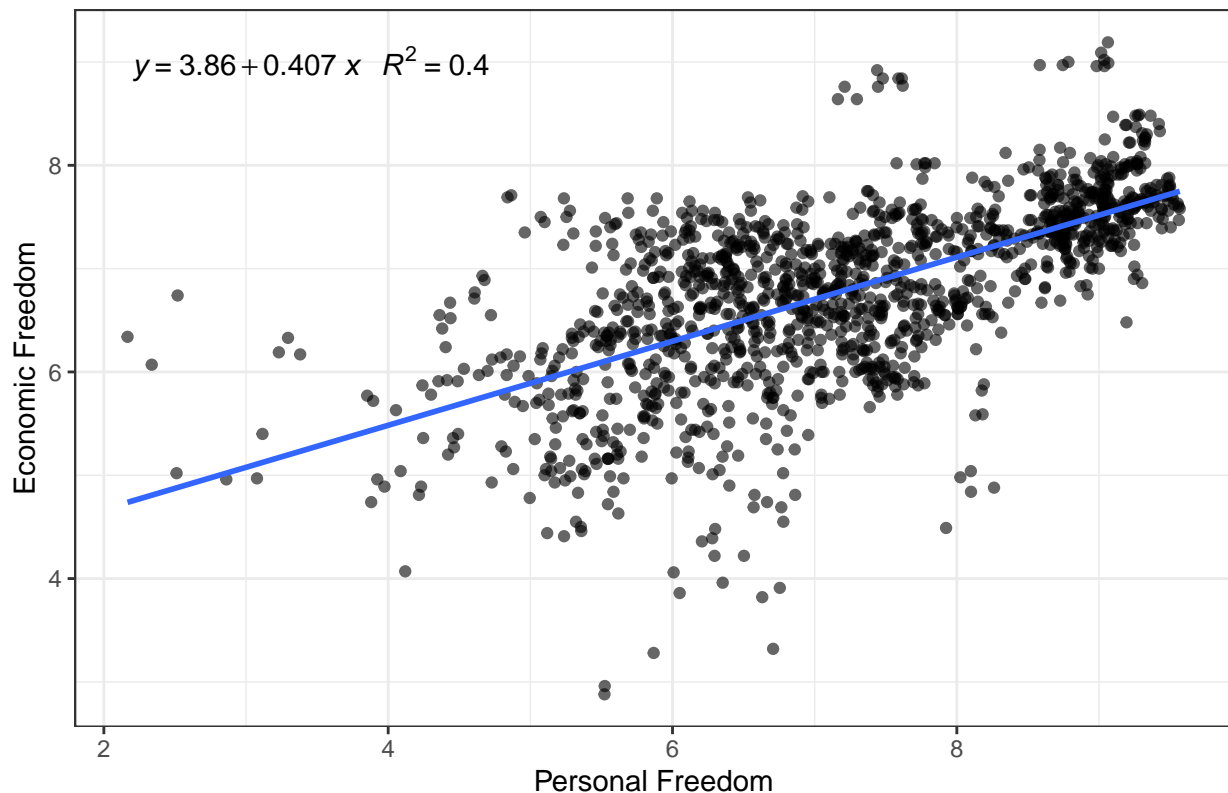
```

library(ggpmisc)

ggplot(data = data3, mapping = aes(x = pf_score, y = ef_score)) +
  geom_point(alpha = 0.6) +
  geom_smooth(method = "lm", formula = y ~ x, se = FALSE) +
  stat_poly_eq(formula = y ~ x,
               aes(label = paste(..eq.label.., ..rr.label.., sep = "~~~")),
               parse = TRUE) +
  labs(x = "Personal Freedom",
       y = "Economic Freedom",
       title = "Personal Freedom en fonction de Economic Freedom") +
  theme_bw() #theme black and white

```

## Personal Freedom en fonction de Economic Freedom



On observe que la grande majorité des pays sont concentrés à droite du nuage de points, ce qui signifie qu'il s'agit de pays pour lesquels l'indice de liberté humaine atteint un score élevé. De plus, on constate qu'il y a peu de pays avec un indice humain très bas, malgré tout de même la présence de quelques observations.

Nous avons également appliqué un modèle de régression linéaire grâce à l'argument **method** de la fonction **geom\_smooth**. De cela, on a extrait l'équation de la droite et le R-squared grâce à la fonction **stat\_poly\_eq** et la librairie **ggpmisc**, qui est une mesure statistique qui indique si le modèle est proche des données et le décrit bien. Après documentation, nous comprenons que le R-squared est un pourcentage de la variabilité décrit par le modèle linéaire.

$R\text{-squared} = \text{Variance expliquée} / \text{Variance totale}$

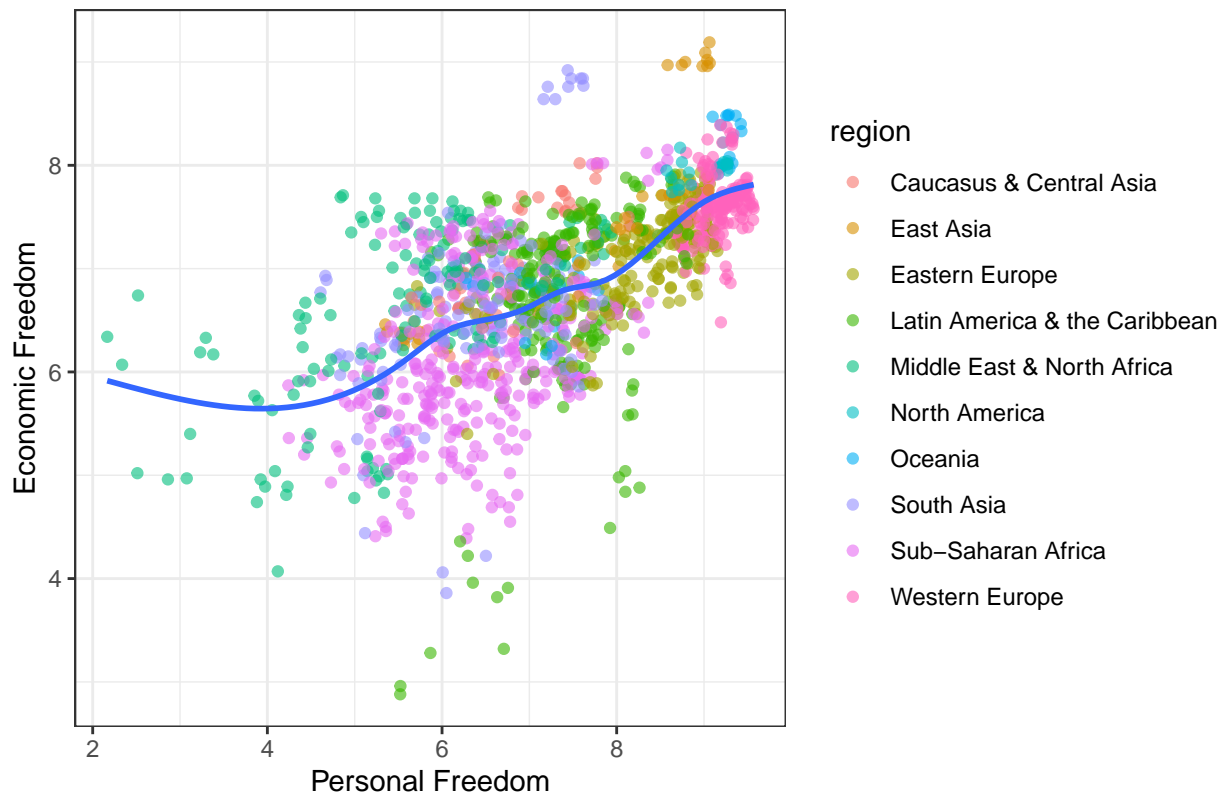
0% indique que le modèle explique 0% de la variabilité des données autour de sa moyenne. 100% indique que le modèle explique toute la variabilité des données autour de sa moyenne.

En général, plus le R-squared est grand, mieux le modèle décrit les données. Ici, nous avons un R-squared de 40%, donc 40% de la variance totale est expliqué par un modèle de régression linéaire, ce qui n'est pas très élevé. Un modèle linéaire ne convient pas.

Essayons maintenant de regarder cela par rapport aux différentes régions du monde.

```
ggplot(data = data3, mapping = aes(x = pf_score, y = ef_score)) +  
  geom_point(mapping = aes(color = region), alpha = 0.6) +  
  geom_smooth(se = FALSE) +  
  labs(x = "Personal Freedom",  
       y = "Economic Freedom",  
       title = "Personal Freedom en fonction de Economic Freedom") +  
  theme_bw()
```

## Personal Freedom en fonction de Economic Freedom

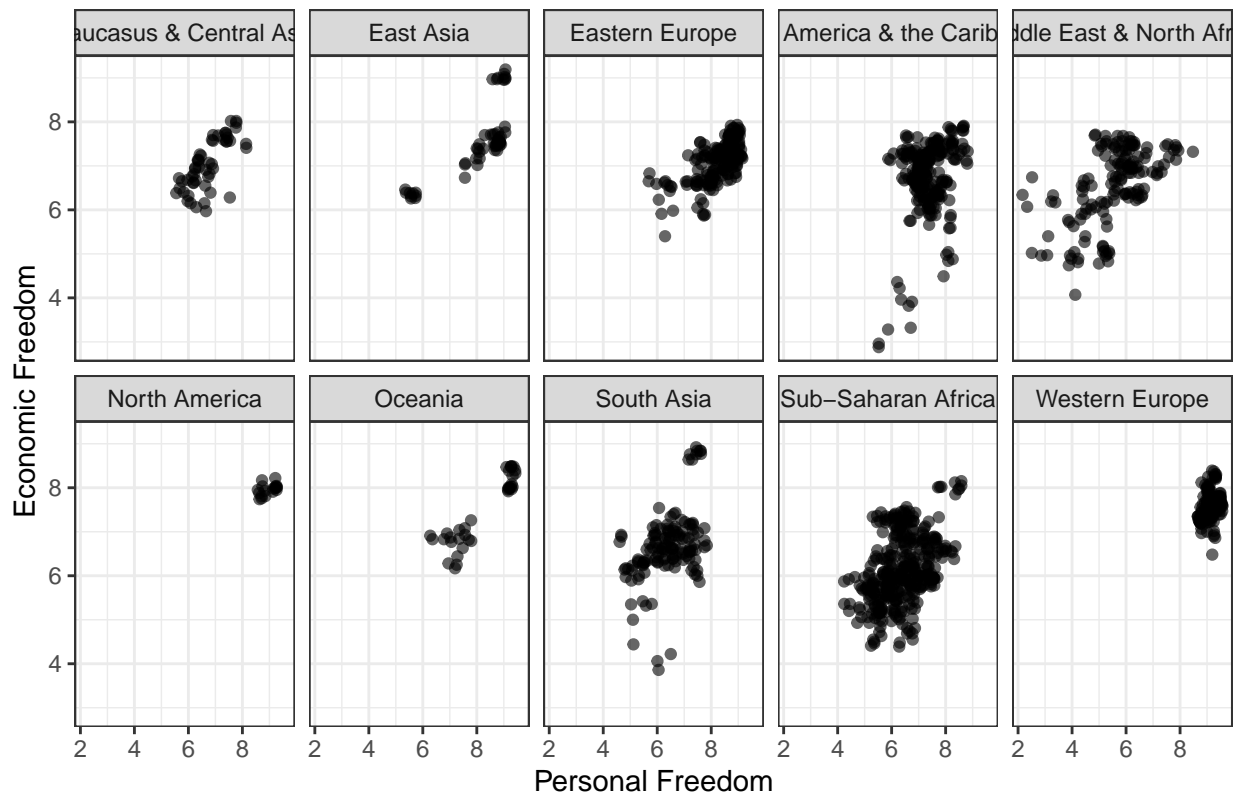


Les régions **Sub-Saharan Africa**, **Latin America & the Caribbean** et **Middle East & North Africa** sont celles qui comptent les pays avec les plus faibles scores de liberté personnelles et économiques. À l'inverse, ce sont les régions **Western Europe** et **Eastern Europe**, ainsi que **North America** qui compte les pays avec le plus haut score. On peut mettre ici en évidence un lien avec le niveau de développement des régions : les pays les plus développés ont un score très élevé (les pays du Nord) tandis que les pays sous-développés ont un score très bas (pays en développement, pays sous-développé, PMA).

Essayons d'utiliser un `facet_wrap` pour une autre visualisation.

```
ggplot(data = data3, mapping = aes(x = pf_score, y = ef_score)) +
  geom_point(alpha = 0.6) +
  facet_wrap(~ region, nrow = 2) +
  labs(x = "Personal Freedom",
       y = "Economic Freedom",
       title = "Personal Freedom en fonction de Economic Freedom") +
  theme_bw()
```

## Personal Freedom en fonction de Economic Freedom



Ce graphique nous permet d'avoir une vision plus globale de l'indice humain en fonction de chaque région. Il peut être intéressant d'étudier les régions **Western Europe** et **Sub-Saharan Africa** et de dresser un classement des pays selon l'indice de liberté humaine. Nous utilisons des requêtes SQL.

```
library(sqldf)
# Western Europe
we = sqldf( "SELECT countries, hf_score
            FROM data3
            WHERE region = 'Western Europe' AND year = 2016" );
we2 = as_tibble(we)
we_croissant <- arrange(we2, hf_score); we_croissant
```

```
## # A tibble: 18 x 2
##   countries      hf_score
##   <chr>         <dbl>
## 1 Italy          7.98
## 2 France         8.01
## 3 Iceland        8.15
## 4 Belgium        8.15
## 5 Spain          8.16
## 6 Portugal       8.28
## 7 Malta          8.35
## 8 Sweden         8.39
## 9 Austria        8.41
## 10 Luxembourg    8.43
## 11 Germany       8.46
```



```
## 12 Norway      8.47
## 13 Finland     8.47
## 14 United Kingdom 8.50
## 15 Ireland     8.50
## 16 Denmark     8.55
## 17 Netherlands 8.55
## 18 Switzerland 8.79
```

```
we_decroissant <- arrange(we2, desc(hf_score)); we_decroissant
```

```
## # A tibble: 18 x 2
##   countries      hf_score
##   <chr>          <dbl>
## 1 Switzerland    8.79
## 2 Netherlands    8.55
## 3 Denmark        8.55
## 4 Ireland        8.50
## 5 United Kingdom 8.50
## 6 Finland        8.47
## 7 Norway         8.47
## 8 Germany        8.46
## 9 Luxembourg     8.43
## 10 Austria        8.41
## 11 Sweden         8.39
## 12 Malta          8.35
## 13 Portugal       8.28
## 14 Spain          8.16
## 15 Belgium        8.15
## 16 Iceland        8.15
## 17 France         8.01
## 18 Italy          7.98
```

```
# Sub-Saharan Africa
af = sqldf( "SELECT countries, hf_score
            FROM data3
            WHERE region = 'Sub-Saharan Africa' AND year = 2016 " ) ;
af2 = as_tibble(af)
af_croissant <- arrange(af2, hf_score); af_croissant
```

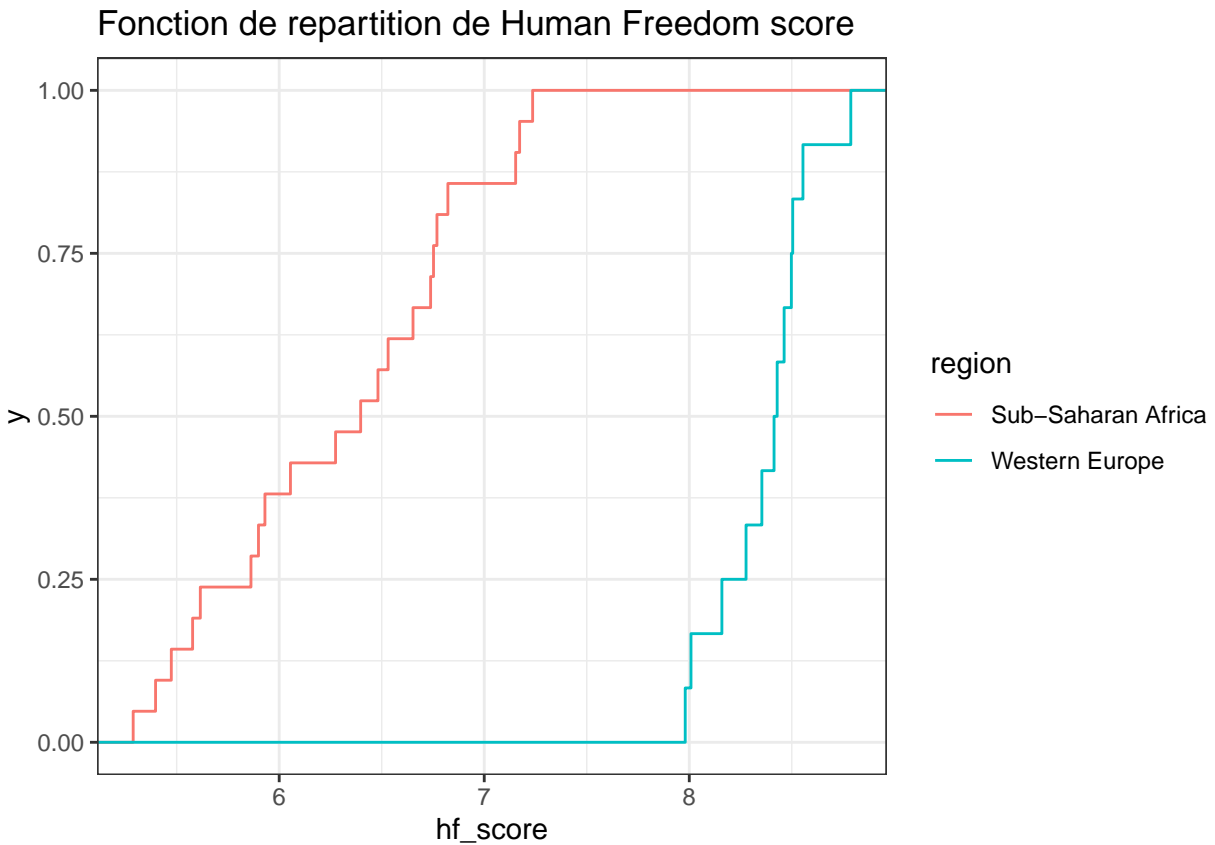
```
## # A tibble: 42 x 2
##   countries      hf_score
##   <chr>          <dbl>
## 1 Sudan          4.80
## 2 Burundi        5.17
## 3 Central Afr. Rep. 5.29
## 4 Congo, Dem. R.   5.31
## 5 Ethiopia        5.40
## 6 Mauritania      5.47
## 7 Chad           5.47
## 8 Cameroon        5.58
## 9 Gabon           5.58
## 10 Zimbabwe       5.62
## # ... with 32 more rows
```

```
af_decroissant <- arrange(af2, desc(hf_score)); af_decroissant
```

```
## # A tibble: 42 x 2
##   countries    hf_score
##   <chr>        <dbl>
## 1 Mauritius    7.86
## 2 Cape Verde   7.33
## 3 Seychelles   7.30
## 4 Ghana        7.24
## 5 South Africa 7.17
## 6 Botswana     7.15
## 7 Rwanda       6.97
## 8 Namibia      6.90
## 9 Kenya       6.82
## 10 Uganda      6.77
## # ... with 32 more rows
```

On construit un classement croissant et décroissant pour les deux régions afin d'observer quels sont les pays avec l'indice humain de liberté le plus élevé et le plus faible. De plus nous pouvons également construire une fonction de répartition qui nous permet de constater l'écart important entre les deux régions.

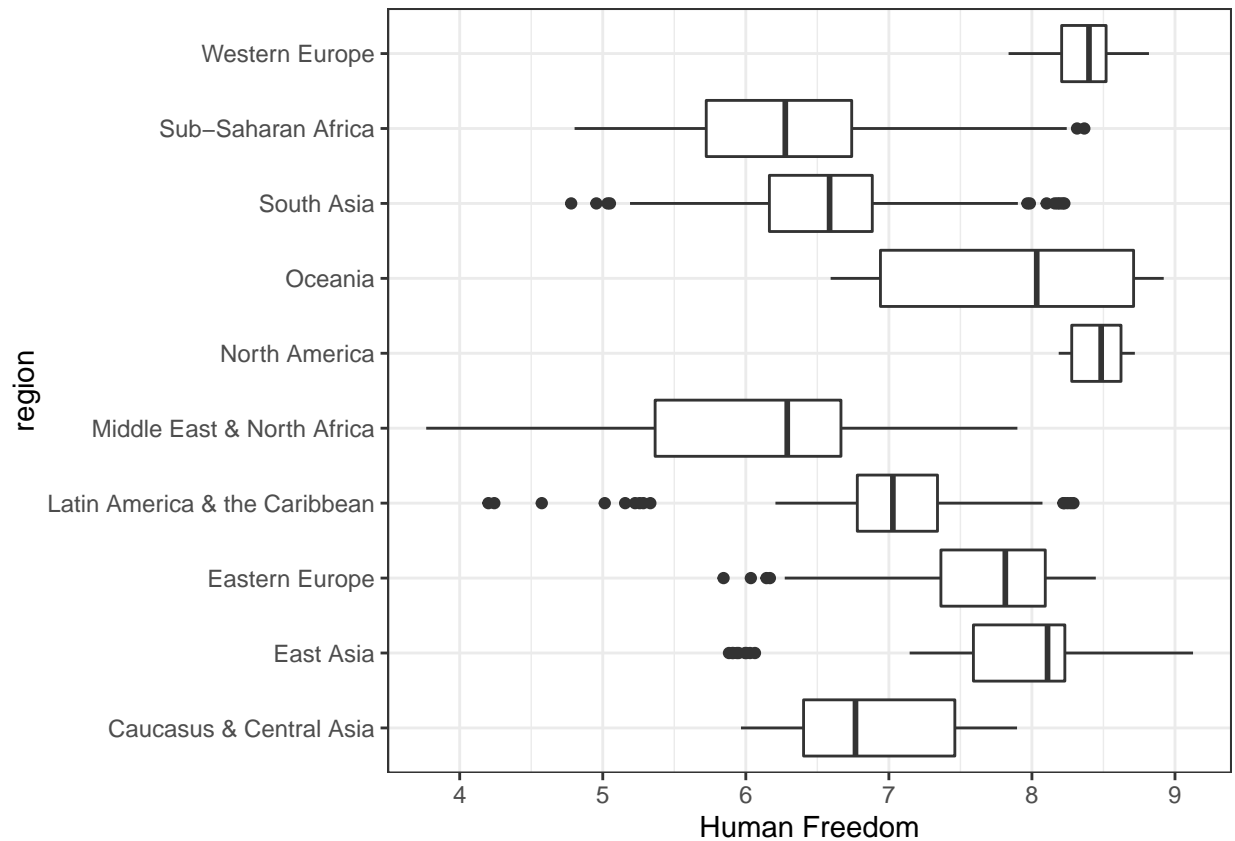
```
data_we_af <- filter(data3, region == c('Western Europe', 'Sub-Saharan Africa'), year == 2016)
ggplot(data_we_af, aes(hf_score, colour = region)) +
  stat_ecdf() +
  labs(title = 'Fonction de repartition de Human Freedom score') +
  theme_bw()
```



Pour la region africaine, près de 50% des pays ont un score humain inferieur à 6.5 alors qu'à l'inverse, dans la region Western Europe, 50% des pays ont un score inferieur à 8.5 approximativement.

Un bon outil pour comparer les differentes regions est le boxplot.

```
ggplot(data = data3, mapping = aes(x = region, y = hf_score)) +  
  geom_boxplot() +  
  labs(y = "Human Freedom") +  
  coord_flip() +  
  theme_bw()
```



`coord_flip()` est utilisé pour avoir une meilleure visibilité sur les noms des régions. Nous avons également la possibilité de recoder les différentes modalités de la variable `region` de la manière suivante :

```
library(forcats) # afin de pouvoir utiliser la fonction fct_recode
```

```
class(data3$region) # il faut vérifier qu'il s'agit bien d'une variable qualitative
```

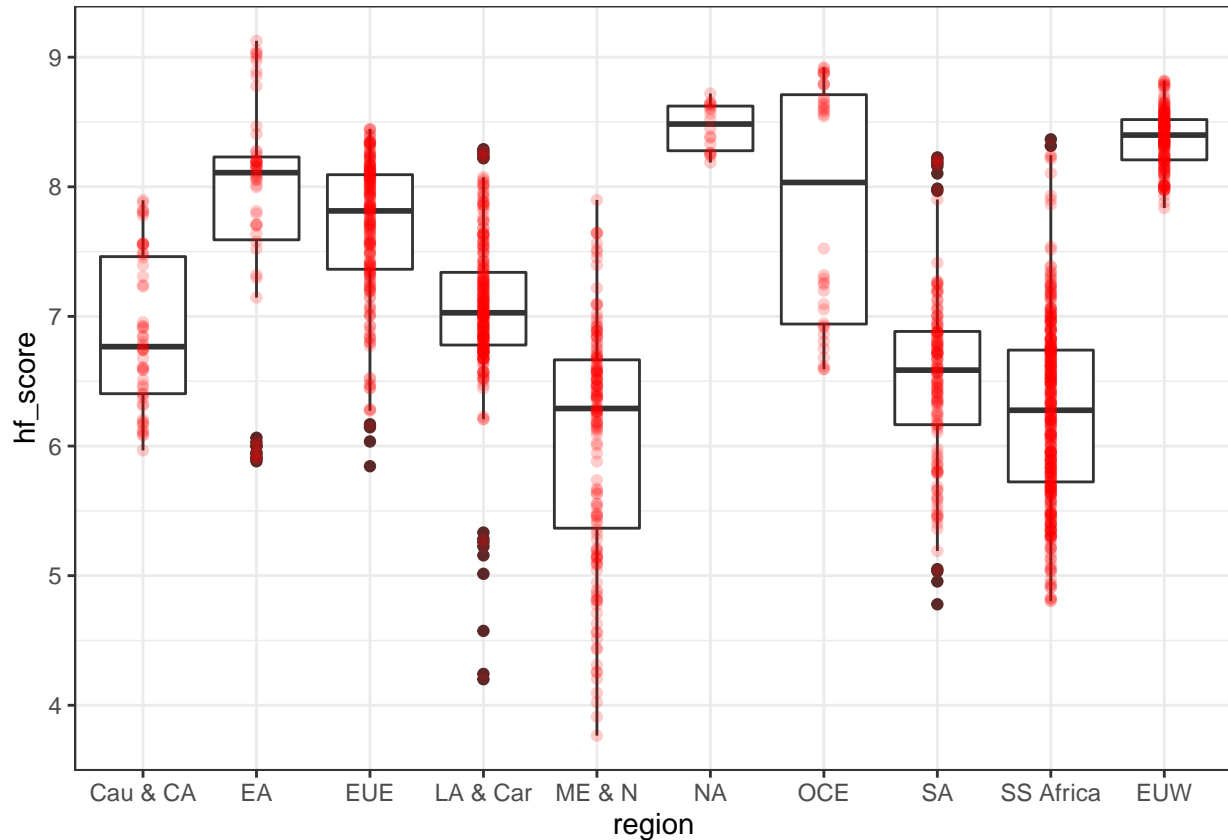
```
## [1] "character"
```

```
data_recode <- data3 # On crée une copie de notre jeu de données initiale afin de garder une sauvegarde
```

```
data_recode$region <- fct_recode(data_recode$region,
  "Cau & CA" = "Caucasus & Central Asia",
  "EA" = "East Asia",
  "EUE" = "Eastern Europe" ,
  "LA & Car" = "Latin America & the Caribbean" ,
  "ME & N" = "Middle East & North Africa" ,
  "NA" = "North America" ,
  "OCE" = "Oceania" ,
  "SA" = "South Asia" ,
  "SS Africa" = "Sub-Saharan Africa" ,
  "EUW" = "Western Europe" )
```

Comme dit précédemment, le boxplot est un bon moyen de comparer les régions.

```
ggplot(data_recode) +
  geom_boxplot(aes(x = region, y = hf_score)) +
  geom_point(aes(x = region, y = hf_score), col = "red", alpha = 0.2) +
  theme_bw()
```



**North America** et **Western Europe** sont les deux regions ayant les meilleurs resultats en terme de score de liberte humaine avec une forte concentration autour de la mediane.

## Corrélation

Etudions la corrélation de nos variables. Nous ne voulons pas utiliser une heatmap classique car nous ne trouvons pas ça très beau à voir (et parfois difficile à comprendre) donc nous avons opté pour un corrélogramme, en gardant les coefficients de corrélations.

```
#install.packages("corrplot")
library(corrplot)

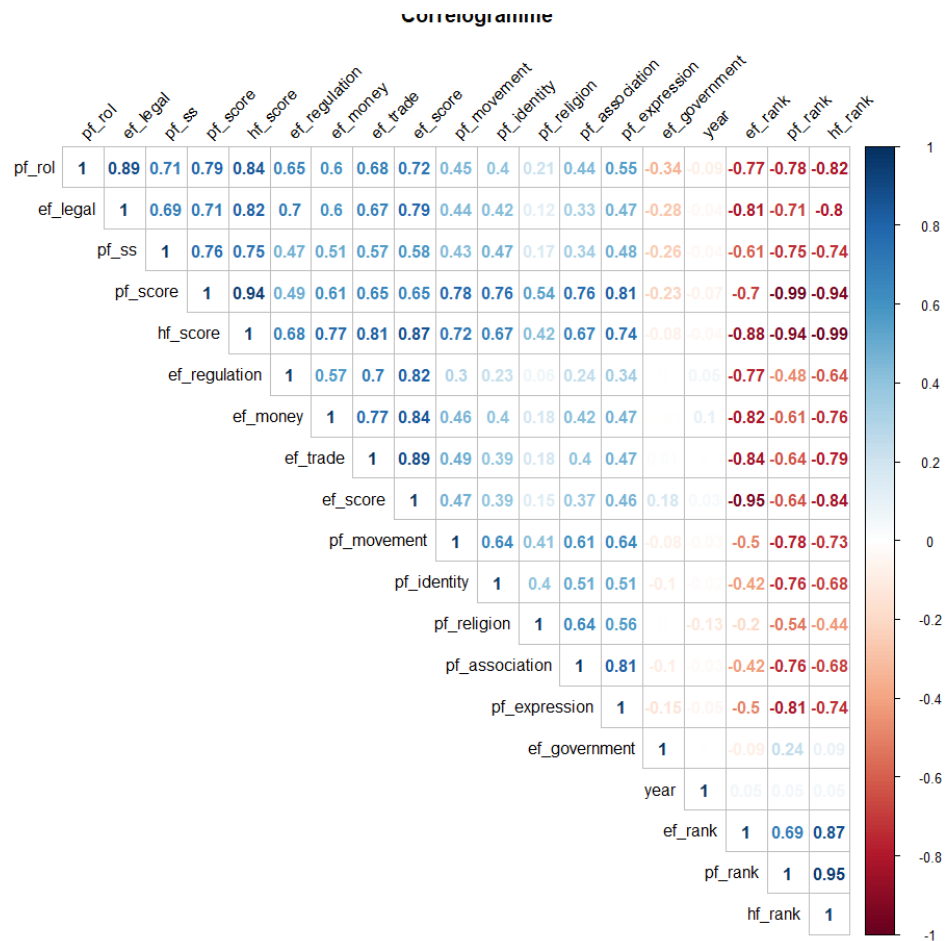
donnee <- select(data3, -c(countries, region)) #on enlève les variables quali
b <- na.omit(donnee) #on enlève les NA
p <- round(cor(b), 2) #on arrondi les corrélations à 2 chiffres après la virgule

corrplot(p,
  method = "number",
  type = "upper",
  order = "hclust",
```

```
tl.col = "black",
tl.srt = 45,
title = "Corrélogramme")
```

Nous avons un petit soucis de rendu, l'image est compressée et les coefficients ne sont pas lisibles. Lorsque nous plottons dans R, il faut zoomer pour mieux voir le corrélogramme (avec le bouton "zoom" sur R) mais cet image zoomé n'est pas accessible directement. Nous avons essayé de changer les attributs fig.width ou encore out.width mais cela ne résout pas le problème. La seule solution est donc d'enregistrer cette image zoomée puis de l'importer via R.

```
knitr::include_graphics("test.png")
```



On peut changer l'argument "method" pour visualiser en cercle, en carré, etc.

Les corrélations positives sont affichées en bleu et les corrélations négatives en rouge. L'intensité de la couleur est proportionnelle aux coefficients de corrélation.

On observe plusieurs choses :

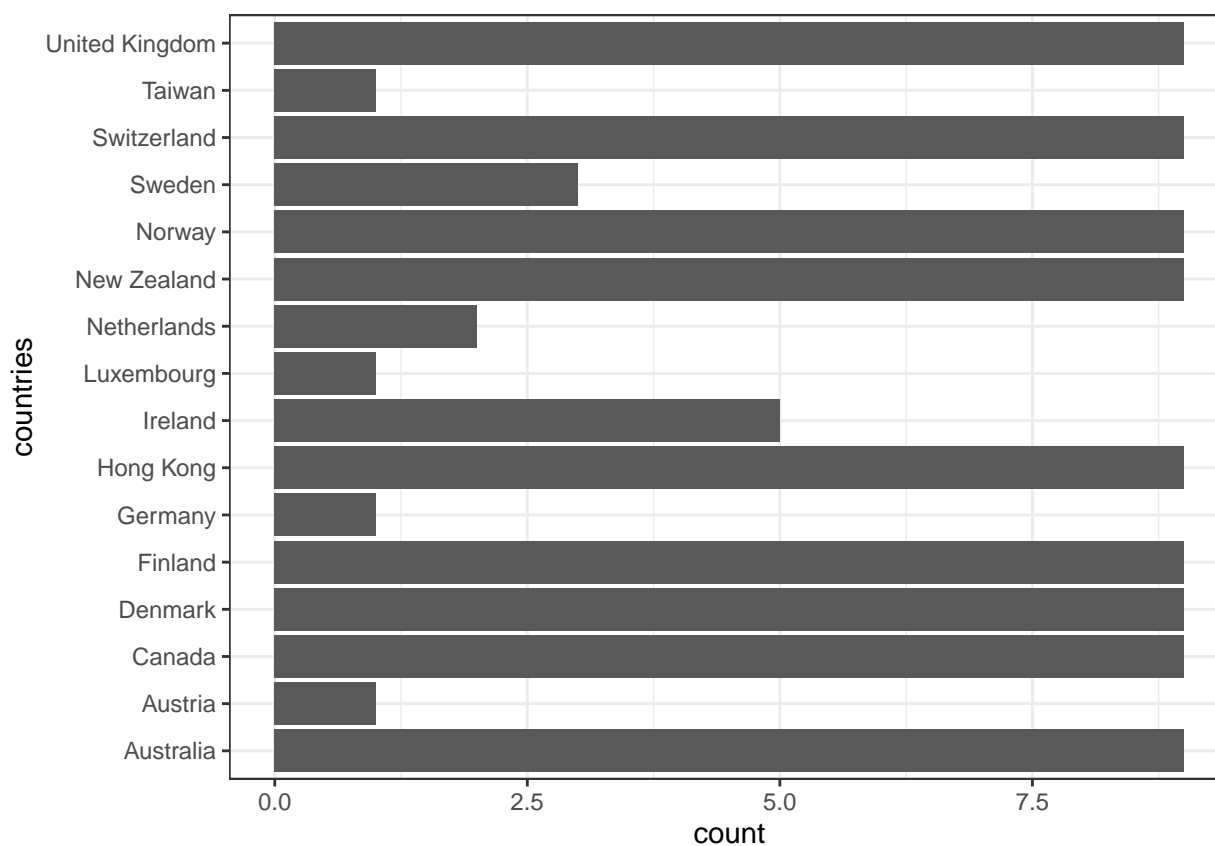
- La taille du gouvernement (ef\_government) n'impacte pas la liberté (corrélation très faible),
- La religion (pf\_religion) n'affecte pas la liberté, contrairement à l'opinion publique.
- Rule of Law a beaucoup d'impact sur la liberté économique et la liberté humaine.
- Corrélation plus élevée de ef\_trade (international trade) que ef\_money (access to money) sur ef\_score (liberté économique)

## Classement

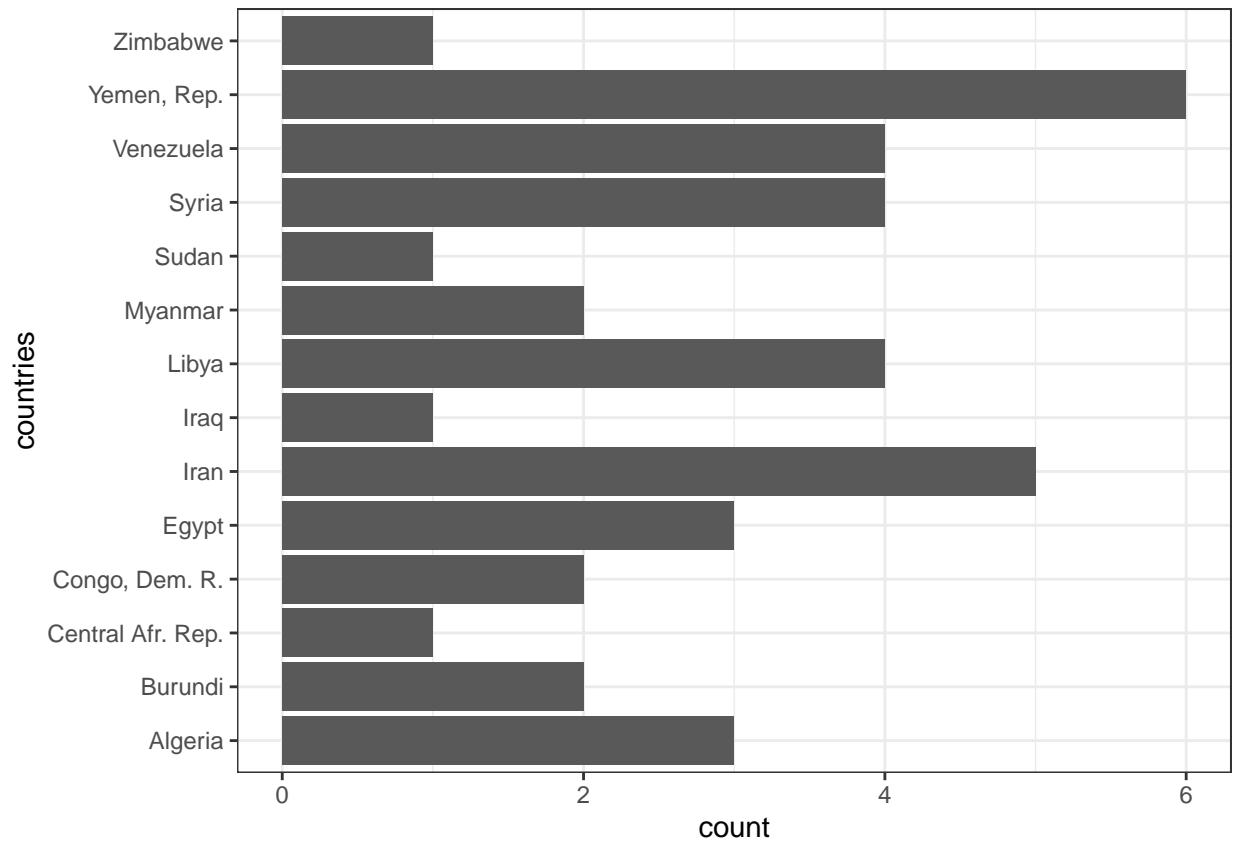
Essayons de regarder le classement des 10 pays les “plus libres” et les “moins libres”, et regardons le nombre d’occurrence de ces derniers suivant les années. On décide de faire des barplot.

```
top10 <- filter(data3, hf_rank <= 10)
last10 <- filter(data3, hf_rank >=152) #il y a 162 pays

#top10
ggplot(data = top10, mapping = aes(x = countries)) +
  geom_bar() +
  coord_flip() +
  theme_bw()
```



```
#last10
ggplot(data = last10, mapping = aes(x = countries)) +
  geom_bar() +
  coord_flip() +
  theme_bw()
```



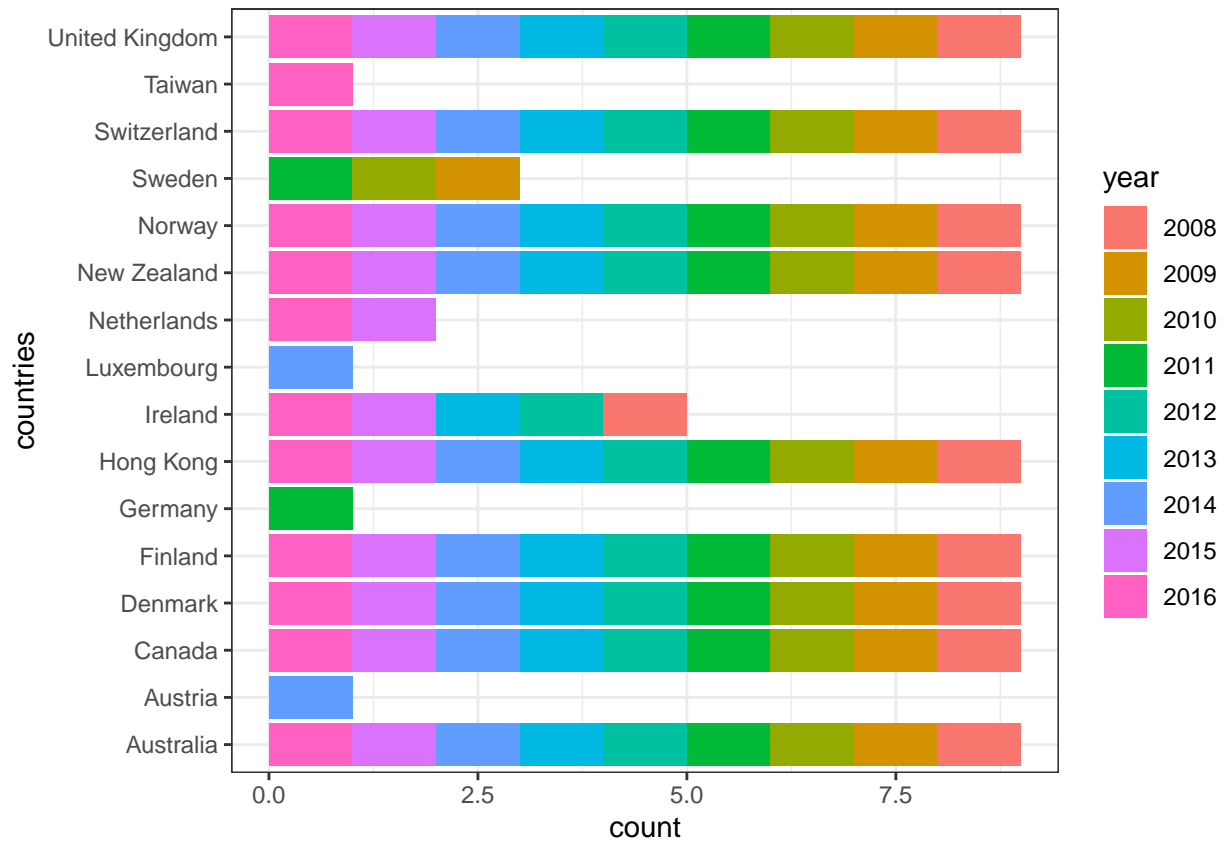
Pour faire un barplot et remplir par année, nous devons d'abord transformer la variable year en string, pour catégoriser cette variable.

```
z <- data3
z$year <- as.character(z$year)

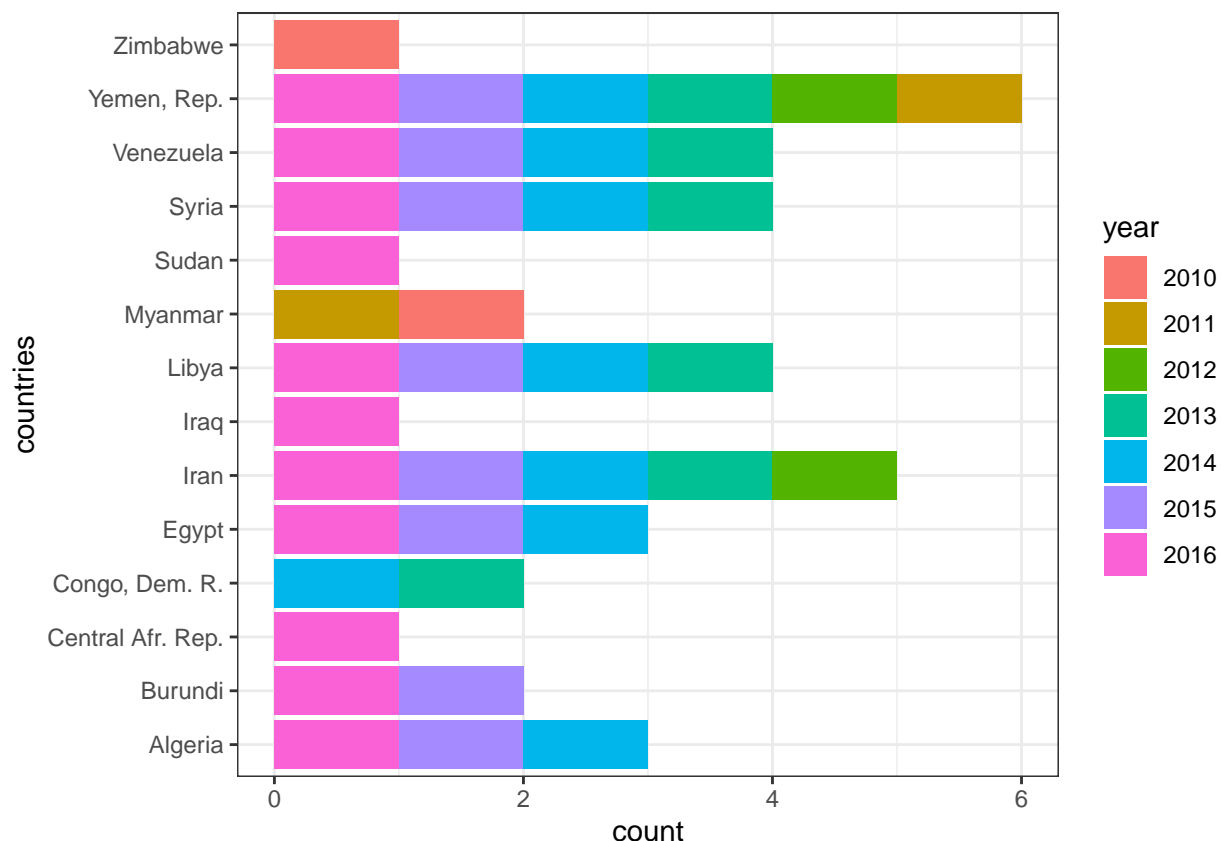
newtop10 <- filter(z, hf_rank <= 10)
newlast10 <- filter(z, hf_rank >= 152)

ggplot(data = newtop10, mapping = aes(x = countries, fill = year)) +
  geom_bar() +
  coord_flip() +
  theme_bw()
```





```
ggplot(data = newlast10, mapping = aes(x = countries, fill = year)) +
  geom_bar() +
  coord_flip() +
  theme_bw()
```



Ce graphique nous permet d'observer combien de fois les pays membres du last 10 sont apparus dans ce classement et également en quelle année ils sont apparus pour la première fois. Le premier graphique nous permet de voir que le Royaume-Uni, la Suisse, Norvège, Nouvelle-Zélande, l'Irlande, Hong Kong, Finlande, Danemark, Canada, et Australie sont les pays qui étaient présents dans tous les top 10 annuels depuis 2008.

Ce classement va dans le sens des pays développés ayant un grand score, étant constamment dans le top 10 contrairement aux pays sous-développés. Certains pays comme la Syrie ou le Yémen qui ont une situation économique et politique compliquée se retrouvent constamment dans le last 10.

## Conclusion

Nous pouvons affirmer par notre étude que le développement du pays est fortement lié à la liberté humaine. La situation politique, économique et sociale d'un pays influe sur la valeur de cet index faisant de lui un indicateur pertinent. En effet, il prend en compte les différents aspects de la liberté et on retrouve bien les pays que l'on attend en haut du classement. Un individu vivant dans un pays très développé s'attend à être plus libre (économiquement et personnellement) que dans un pays sous-développé, pauvre, en guerre, etc.