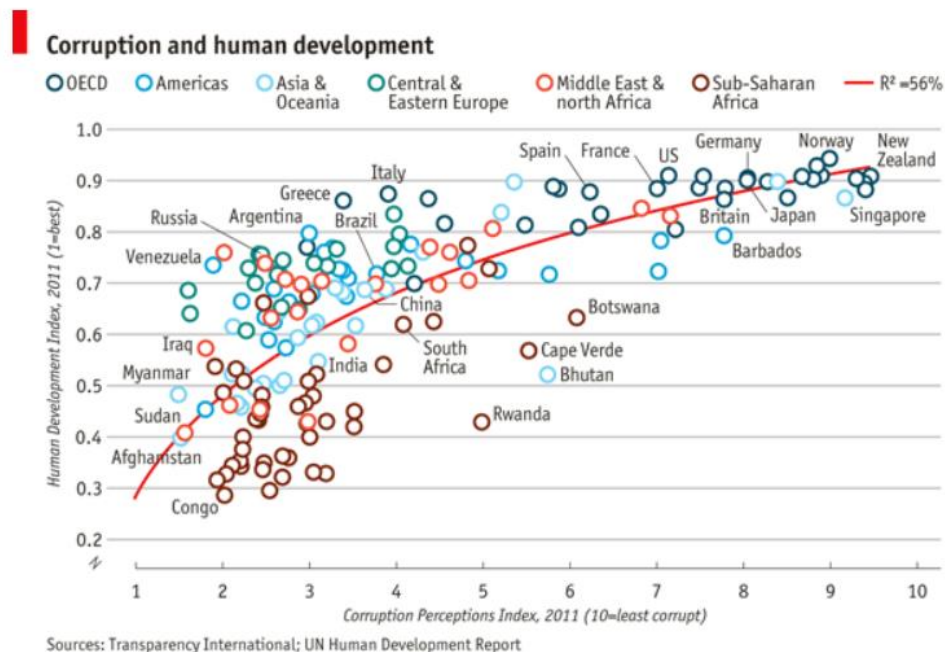


Hiver 2019

Professeur : Amor Amami

## Laboratoire : visualisation avec ggplot2

**Objectif :** nous recréerons ce graphique de [The Economist](#) :



1. Importez les bibliothèques ggplot2, data.table et utilisez fread pour charger le fichier csv 'Economist\_Data.csv' dans un cadre de données appelé df (Conseil: utilisez drop = 1 pour ignorer la première colonne).

```
library(ggplot2)
library(data.table)
```

```
df <- fread('Economist_Data.csv',drop=1)
```

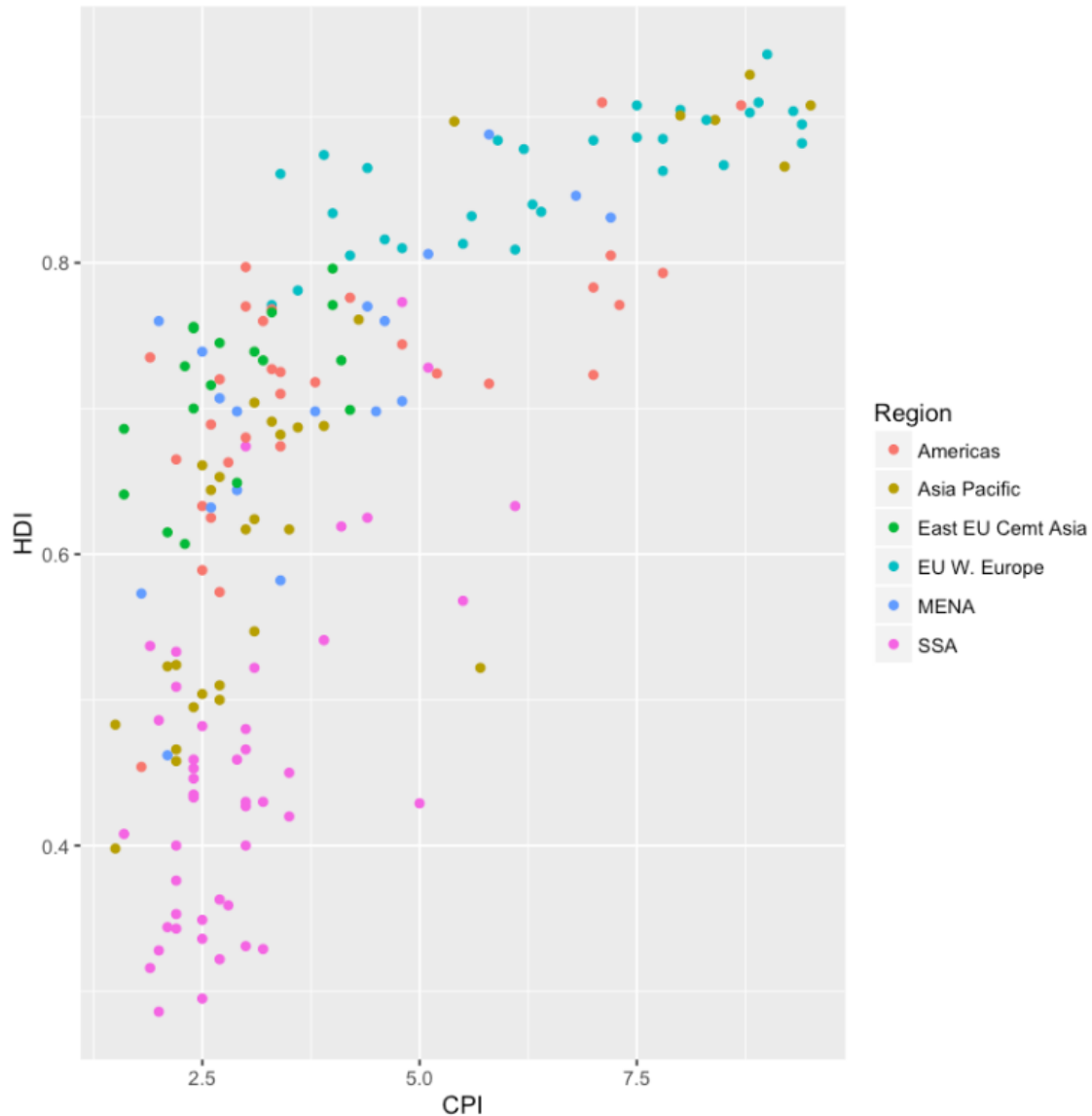
2. Afficher les 6 premiers éléments :

```
head(df)
```

	Country	HDI.Rank	HDI	CPI	Region
1	Afghanistan	172	0.398	1.5	Asia Pacific
2	Albania	70	0.739	3.1	East EU Cemt Asia
3	Algeria	96	0.698	2.9	MENA
4	Angola	148	0.486	2	SSA
5	Argentina	45	0.797	3	Americas
6	Armenia	86	0.716	2.6	East EU Cemt Asia

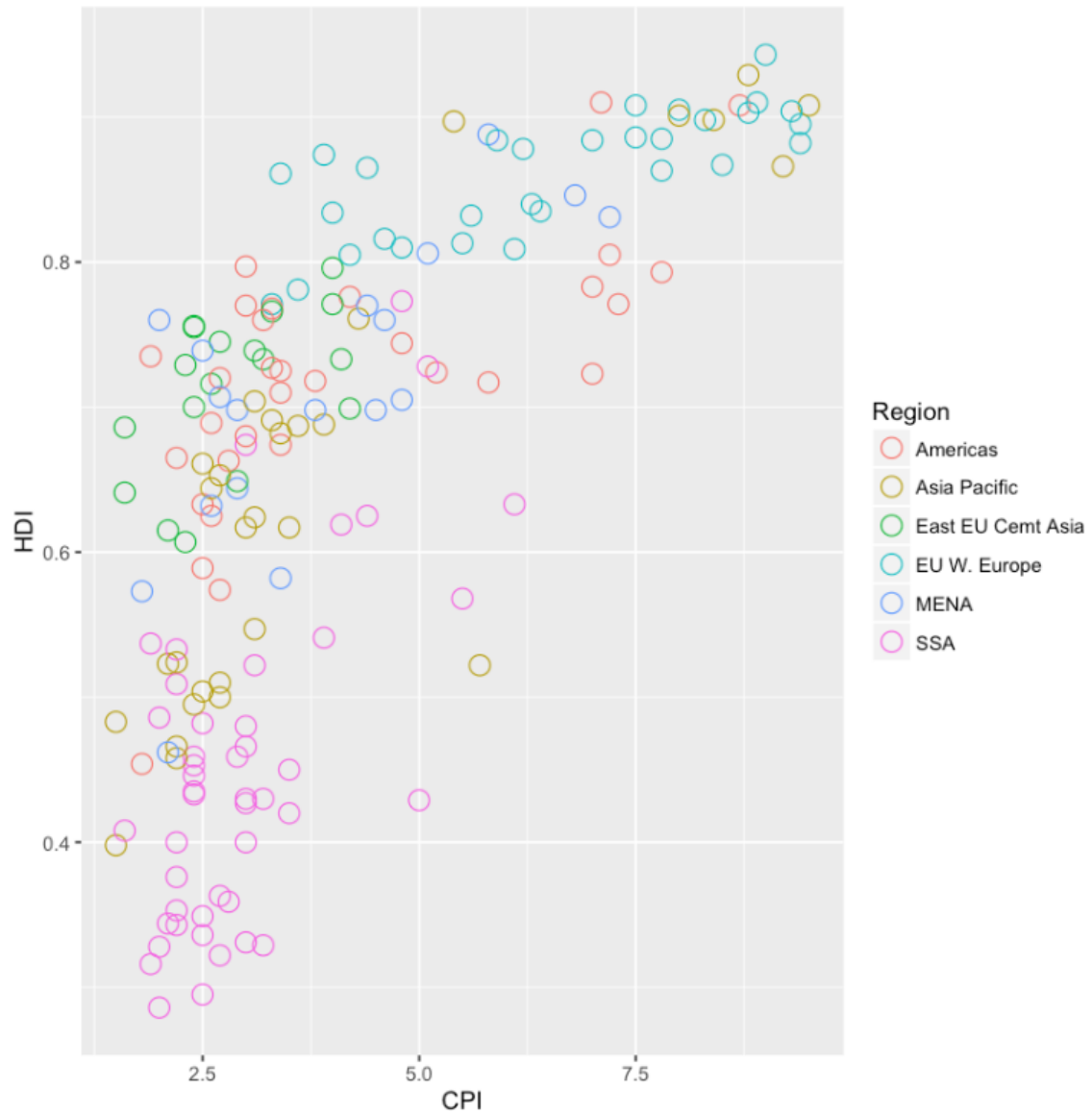
3. Utilisez ggplot () + geom\_point () pour créer un objet de diagramme de dispersion appelé **pl** . Vous devrez spécifier x = CPI et y = HDI et color = Région comme esthétique

```
pl <- ggplot(df,aes(x=CPI,y=HDI,color=Region)) + geom_point()  
pl
```



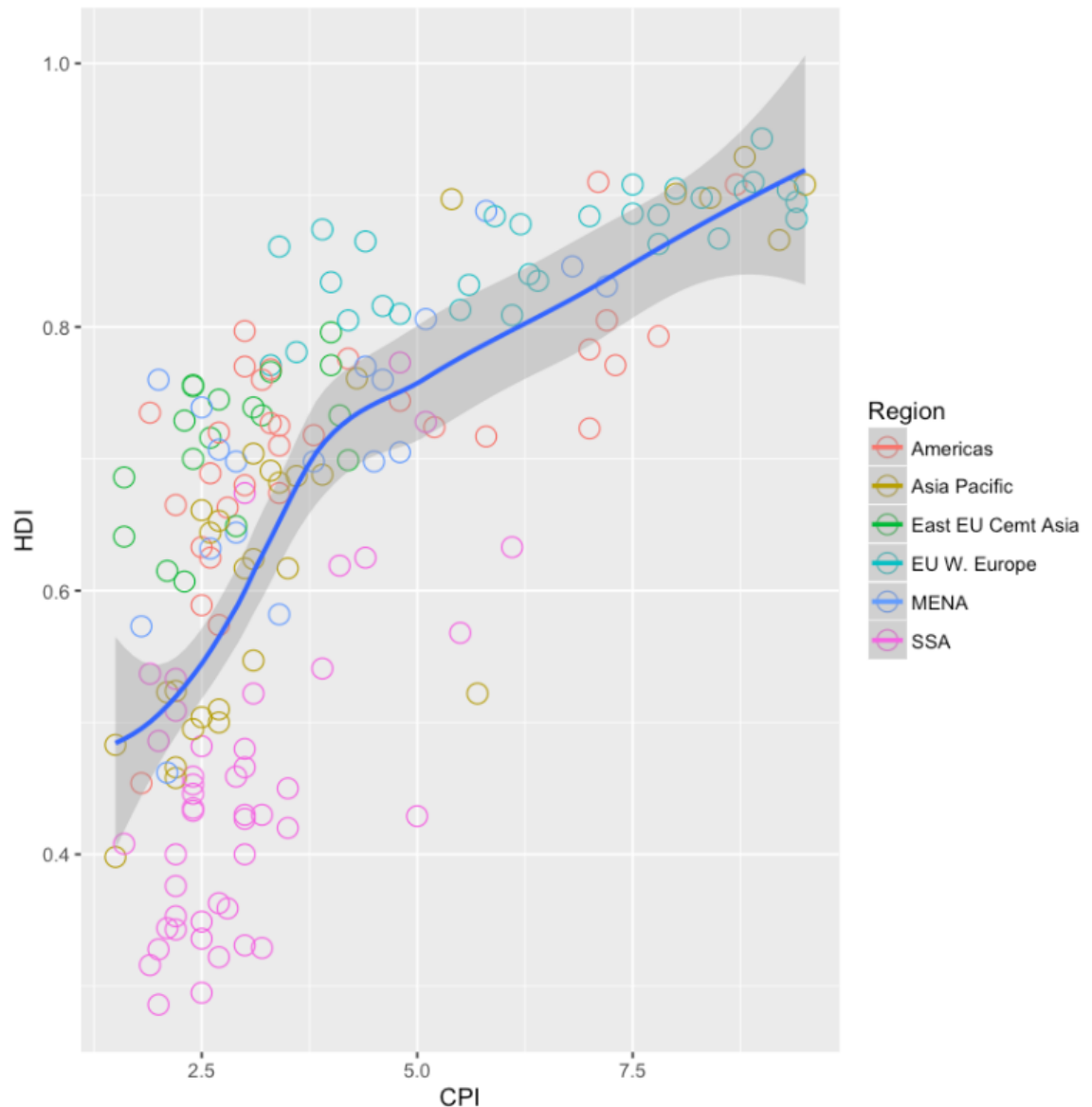
4. Changer les points pour être plus grands cercles vides. (Vous devrez revenir en arrière et ajouter des arguments à `geom_point()` et le réaffecter à `pl`.) Vous devrez déterminer quelle **shape=** et **size=**

```
pl <- ggplot(df,aes(x=CPI,y=HDI,color=Region)) + geom_point(size=4,shape=1)
pl
```



5. Ajoutez `geom_smooth(aes(groupe = 1))` pour ajouter une ligne de tendance

```
pl + geom_smooth(aes(group=1))
```



6. Nous voulons éditer davantage cette ligne de tendance. Ajoutez les arguments suivants à `geom_smooth` (en dehors de `aes`):

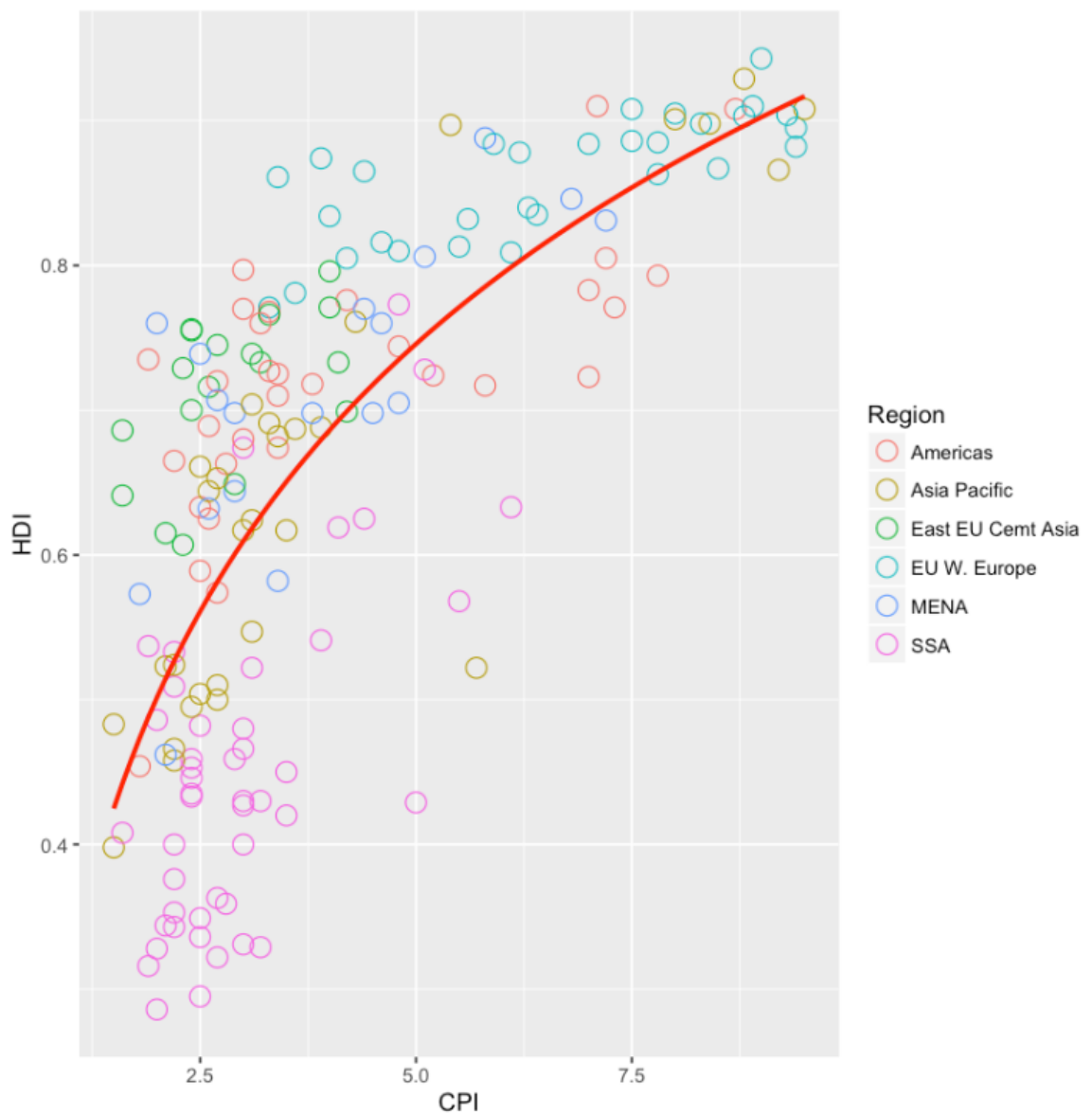
- `méthode = 'lm'`
- `formule = y ~ log (x)`
- `se = FALSE`
- `couleur = 'rouge'`

Pour plus d'informations sur ces arguments, consultez la [documentation](#) sous la liste *Arguments* pour plus de détails.

Attribuer tout cela à pl2

```
pl2 <- pl + geom_smooth(aes(group=1),method = 'lm',formula = y~log(x),se=
FALSE,color='red')
```

pl2



7. Ça commence vraiment à ressembler! Mais nous devons encore ajouter des étiquettes, nous pouvons

utiliser `geom_text` ! Ajoutez `geom_text(aes(label = Country))` à `pl2` et voyez ce qui se passe. (Indice: il devrait y avoir beaucoup trop d'étiquettes)

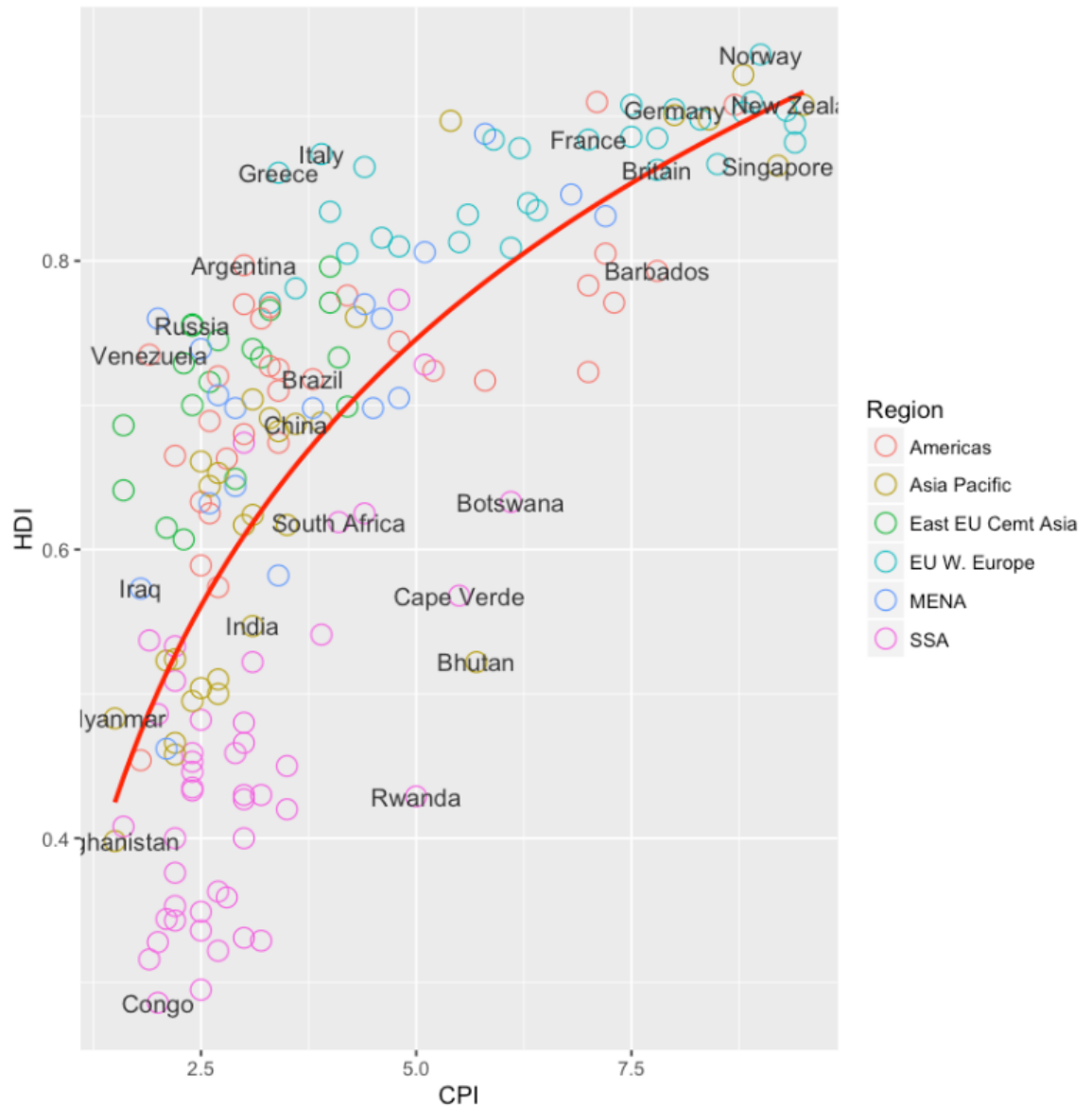
```
pl2 + geom_text(aes(label=Country))
```



8. Étiqueter un sous-ensemble est en fait assez délicat! Nous allons donc simplement vous donner la réponse car il faudrait sélectionner manuellement le sous-ensemble de pays que nous voulons étiqueter!

```
pointsToLabel <- c("Russia", "Venezuela", "Iraq", "Myanmar", "Sudan",  
  "Afghanistan", "Congo", "Greece", "Argentina", "Brazil",  
  "India", "Italy", "China", "South Africa", "Spain",  
  "Botswana", "Cape Verde", "Bhutan", "Rwanda", "France",  
  "United States", "Germany", "Britain", "Barbados", "Norway", "Ja  
pan",  
  "New Zealand", "Singapore")  
  
pl3 <- pl2 + geom_text(aes(label = Country), color = "gray20",  
  data = subset(df, Country %in% pointsToLabel), check_overlap = TR  
UE)  
pl3
```





9. Ajoutez `theme_bw()` à votre graphique et enregistrez-le dans `pl4`

```
pl4 <- pl3 + theme_bw()
pl4
```



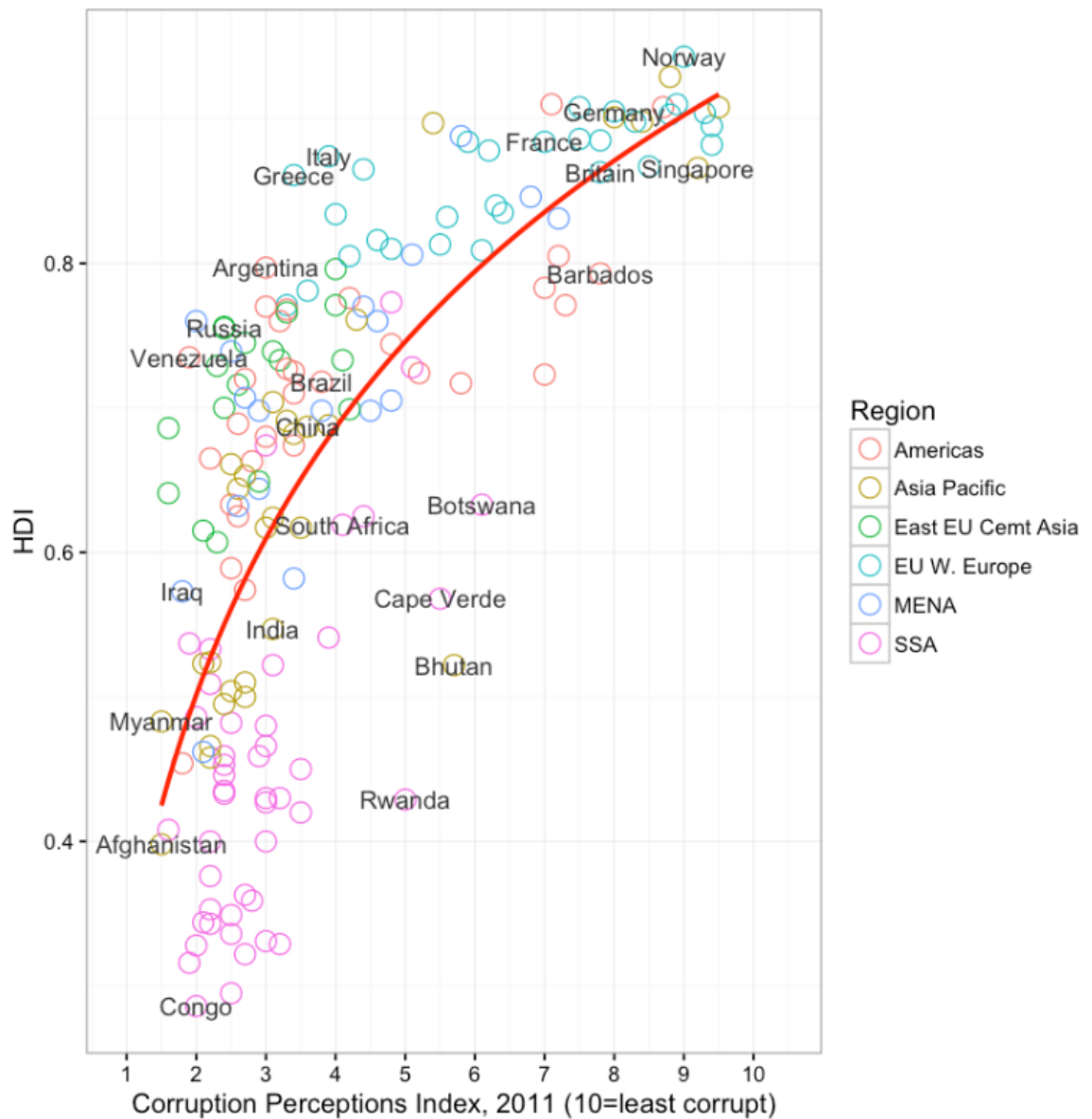
10. Ajoutez `scale_x_continuous()` et définissez les arguments suivants:

- `nom` = même axe x que le tracé de l'économiste
- `limits` = Passer un vecteur de x limites appropriées
- `pauses` = 1:10

```
pl5 <- pl4 + scale_x_continuous(name = "Corruption Perceptions Index, 2011 (10=least corrupt)",
```

```
limits = c(.9, 10.5),breaks=1:10)
```

p15



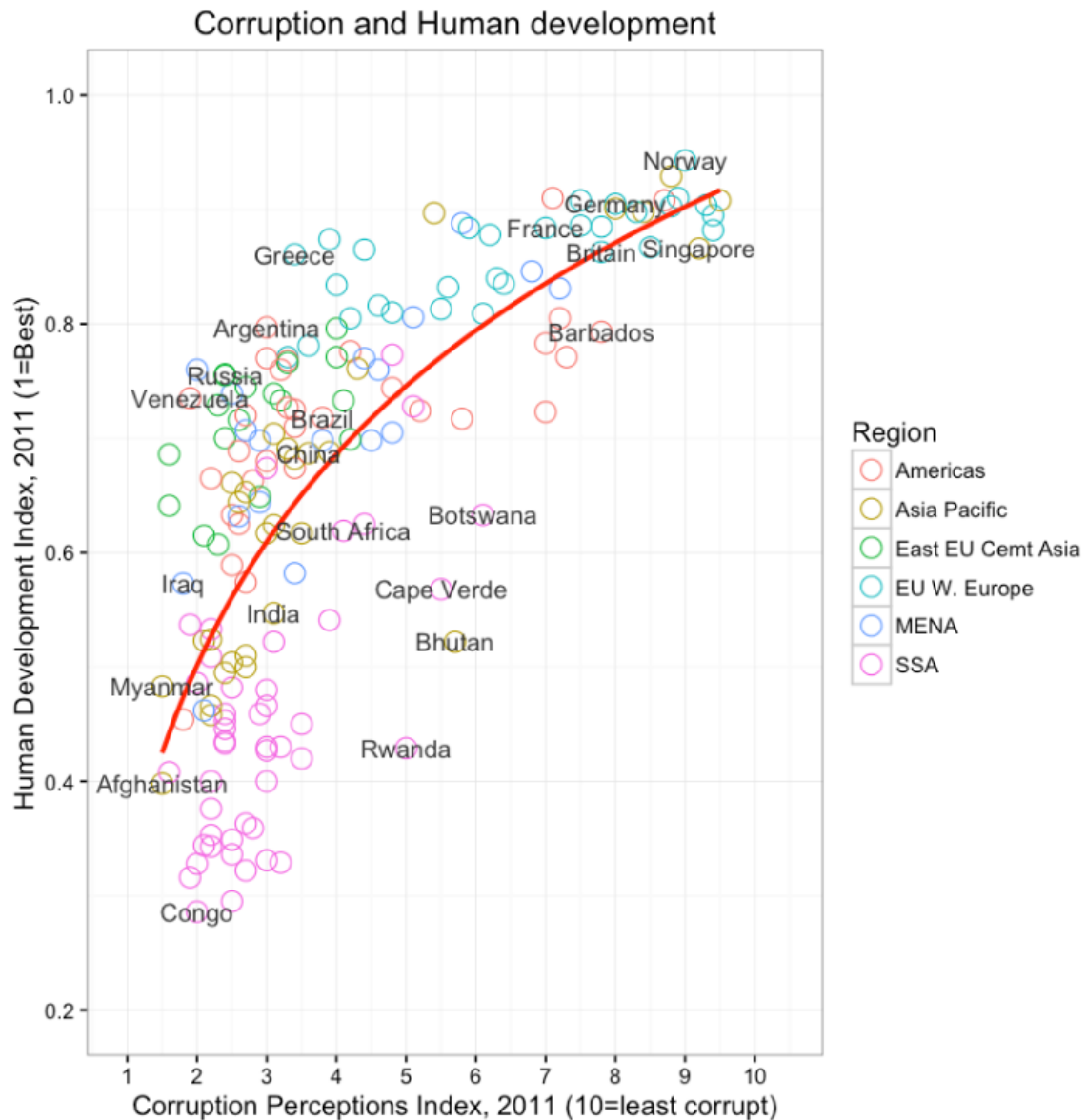
11. Utilisez `scale_y_continuous` pour effectuer des opérations similaires sur l'axe y!

```
pl6 <- pl5 + scale_y_continuous(name = "Human Development Index, 2011  
(1=Best)",  
                                limits = c(0.2, 1.0))  
pl6
```



12. Utilisez ggtitle () pour ajouter un titre.

```
pl6 + ggtitle("Corruption et développement humain")
```



### 13. Utiliser ggthemes:

```
library(ggthemes)
p16 + theme_economist_white()
```

