

# Étude des déterminants de la performance environnementales des entreprises du STOCXX600

Projet d'Initiation à R et Dataviz

**M1 IFIM 2023/2024**

Université Sorbonne Paris Nord  
Darvin TERENCE SURENDRA

**Professeur : IMEN GHATTASSI**

## Choix de la mesure de la performance environnementale

Nous disposons d'une base de données qui propose des variables économiques, financières et de gouvernance relative aux entreprises du STOXX 600.

Elle est renommée : *Projet.xlsx*

Nous allons effectuer une étude économétrique où nous étudions les déterminants de la performance environnementale de l'entreprise en fonction de ses performances économiques et financières, et ses caractéristiques de gouvernance.

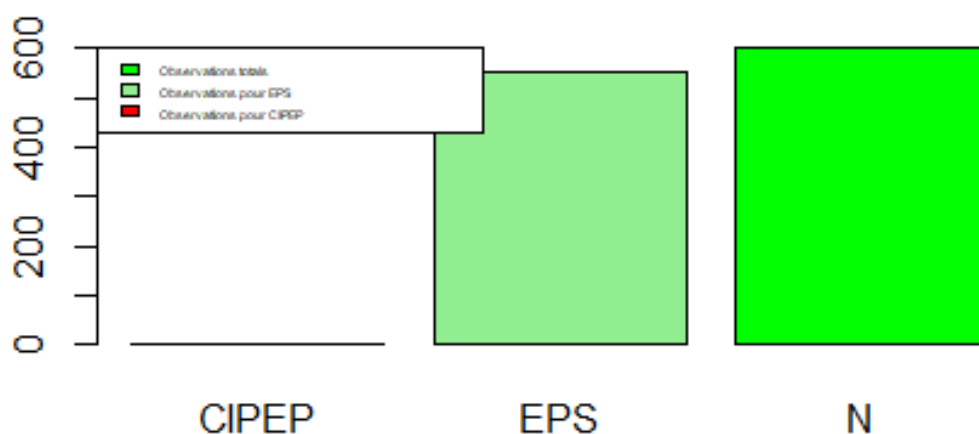
Comme mesures de la performance environnementale, nous nous sommes vus proposé deux variables :

- Environment Pillar Score (EPS)
- Carbon Intensity per Energy Produced (CIPEP)

Après une rapide analyse visuelle de la base de données, nous avons pu exclure la variable CIPEP qui est principalement composé de NA.

Pour imager notre réponse, nous avons produit un graphique du nombre de réponses collectées pour chacune de ces variables.

**Diagramme à Barres du nombre de réponses**



## Analyse en Composante Principale sur les variables financières

Comme variable financière, nous avons plusieurs mesures de la variable du prix :

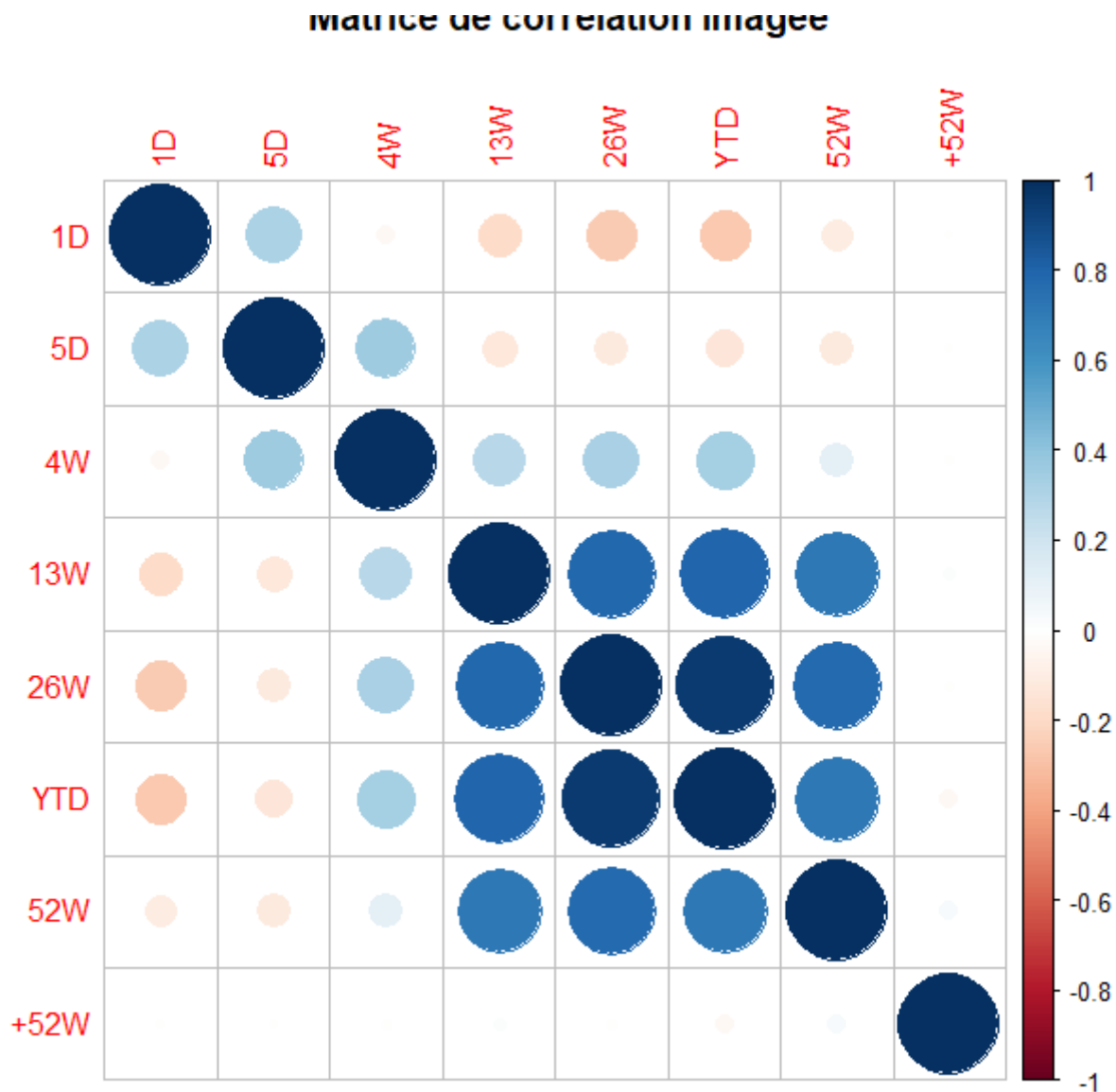
NOM DE LA VARIABLE	VARIABLE RENOMME
1-DAY PRICE PCT CHANGE	1D
5-DAY PRICE PCT CHANGE	5D
4-WEEK PRICE PCT CHANGE	4W
13-WEEK PRICE PCT CHANGE	13W
26-WEEK PRICE PCT CHANGE	26W
YTD PRICE PCT CHANGE	YTD
52-WEEK PRICE PCT CHANGE	52W
PRICE 52 WEEK HIGH	+52W

Il s'annonce être très difficile de réussir à expliquer les données de 8 variables (retranscrit donc en 8 dimensions). En usant de l'ACP, nous allons projeter les données sur un hyperplan de cet espace à 8 dimensions. Ainsi, nous allons certes perdre de l'information, mais pouvoir lire plus simplement les données. Par ailleurs, nos 8 variables ne sont pas

forcément indépendantes l'une de l'autre. Nous allons donc nous débarrasser si possible de variables corrélées entre elle.

Dans un premier temps, nous observons les corrélations entre les diverses variables. Et ceux, à travers cette matrice de corrélation obtenu à partir de la librairie *Corrplot*:

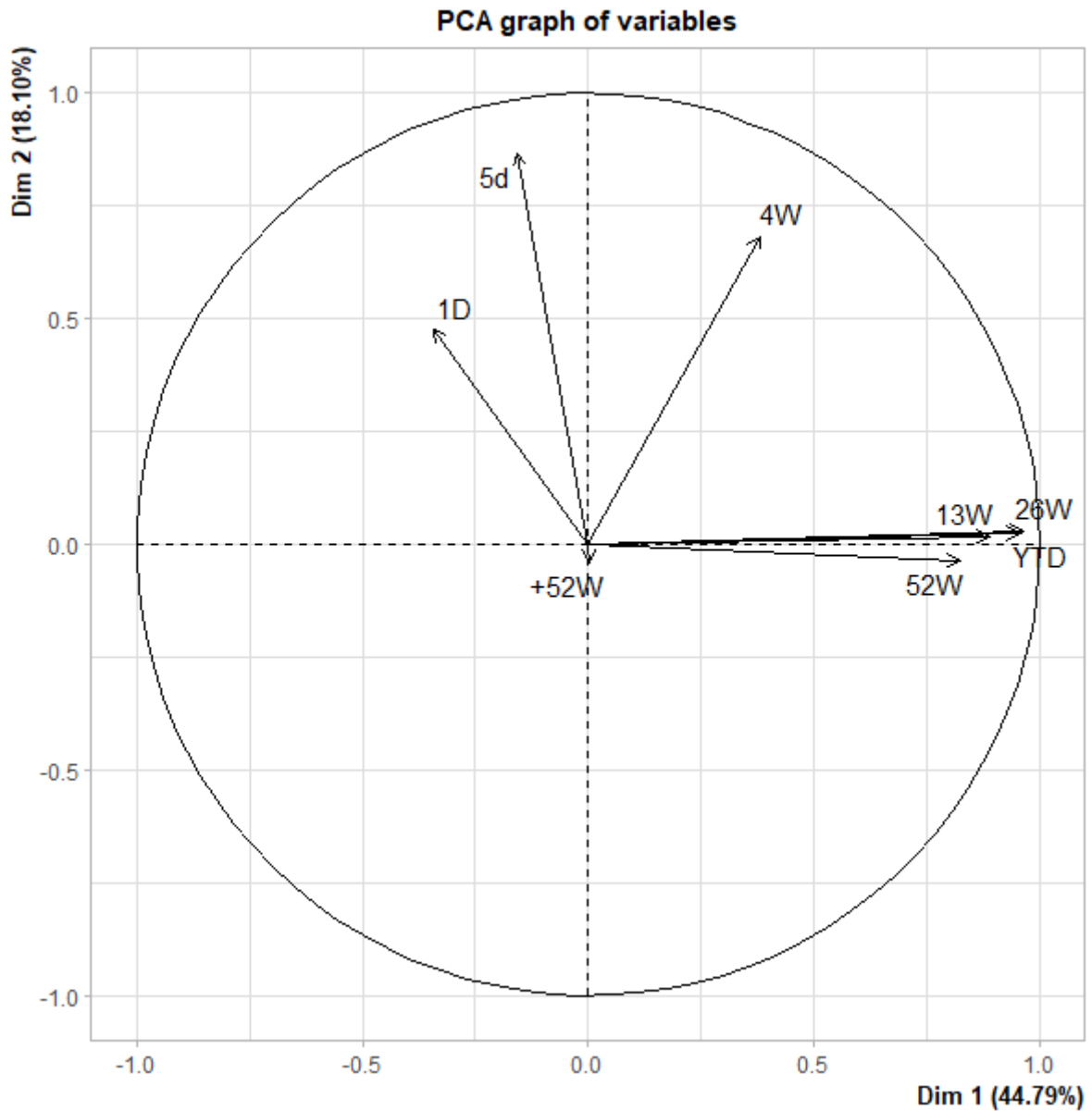
Matrice de corrélation



À partir de ce graphique, on peut conclure qu'on ne peut pas considérer certaines variables tout en gardant un niveau convenable d'information.

Passons désormais à l'Analyse en Composante Principale.

En utilisant la fonction *PCA()* de la bibliothèque *FactoMineR*, nous avons le graphique suivant :



Ce graphique nous montre à quel degré une variable est retranscrit sur la Dim 1 ou la Dim2 (qui sont orthogonaux entre eux). Ainsi, la 1<sup>ère</sup> variable (Dim 1), nous résumera 44,79% de l'information, tandis que si l'on considère aussi la 2<sup>ème</sup> variable (Dim2), nous résumerions 62,89% de l'information.

Enfin, on souligne le fait que la variable +52W semble être décorrélé de toutes les autres variables.

À partir de cette ACP, nous allons désormais considérer une nouvelle variable  $F_i$ , qui sera la projection de nos observations sur la variable Dim 1.

### *Dataviz, Statistiques descriptives et régression linéaire*

Nous allons désormais chercher à expliquer la variable Environment Pillar Score (EPS), à partir des variables explicatives suivantes :

NOM DE LA VARIABLE	VARIABLE RENOMME
COUNTRY OF DOMICIL	COD
GENERAL INDUSTRY CLASSIFICATION	Variable qui n'était pas dans la base de données
EMPLOYEES	EMP
OPERATING PROFIT MARGIN	MARGIN
NET SALES OR REVENUES	SALES
MARKET CAP	MCAP
RETURN ON INVESTED CAPITAL	RCAP
CSR SUSTAINABILITY COMMITTEE	CSR
VALUE – BOARD STRUCTURE/INDEPENDENT BOARD MEMBERS	BOARD
Fi	Fi

Voici les statistiques descriptives des variables quantitatives de cette liste :

EPS	EMP	MARGIN	SALES
Min. : 0.00	Min. : 0	Min. : -275.78	Min. : 52061
1st Qu.:49.15	1st Qu.: 3851	1st Qu.: 7.28	1st Qu.: 2385684
Median :69.01	Median : 14522	Median : 12.86	Median : 7159000
Mean :63.95	Mean : 40720	Mean : 20.04	Mean : 23724980
3rd Qu.:82.92	3rd Qu.: 43495	3rd Qu.: 23.02	3rd Qu.: 23969000
Max. :98.16	Max. :671205	Max. : 731.44	Max. :554184257
NA's :46	NA's :28	NA's :28	NA's :27

MCAP	RCAP	BOARD	Fi
Min. :1.480e+09	Min. : -50.960	Min. : 0.00	Min. : -5.5788
1st Qu.:5.548e+09	1st Qu.: 4.610	1st Qu.: 50.00	1st Qu.: -1.2093
Median :1.386e+10	Median : 8.175	Median : 65.00	Median : -0.0587
Mean :4.050e+10	Mean : 10.837	Mean : 64.15	Mean : 0.0000
3rd Qu.:3.639e+10	3rd Qu.: 13.688	3rd Qu.: 80.00	3rd Qu.: 1.0247
Max. :1.322e+12	Max. :437.740	Max. :100.00	Max. :11.2132
	NA's :30	NA's :46	

Et voici ceux des variables quantitatives :

Country of domicile

BD	65	LX	8
BG	16	MA	1
BM	1	NL	30
DK	23	NW	17
ES	22	OE	9
FA	1	PO	8
FN	17	PT	3
FR	69	SD	61
IM	1	SW	53
IR	11	UK	131
IT	27		

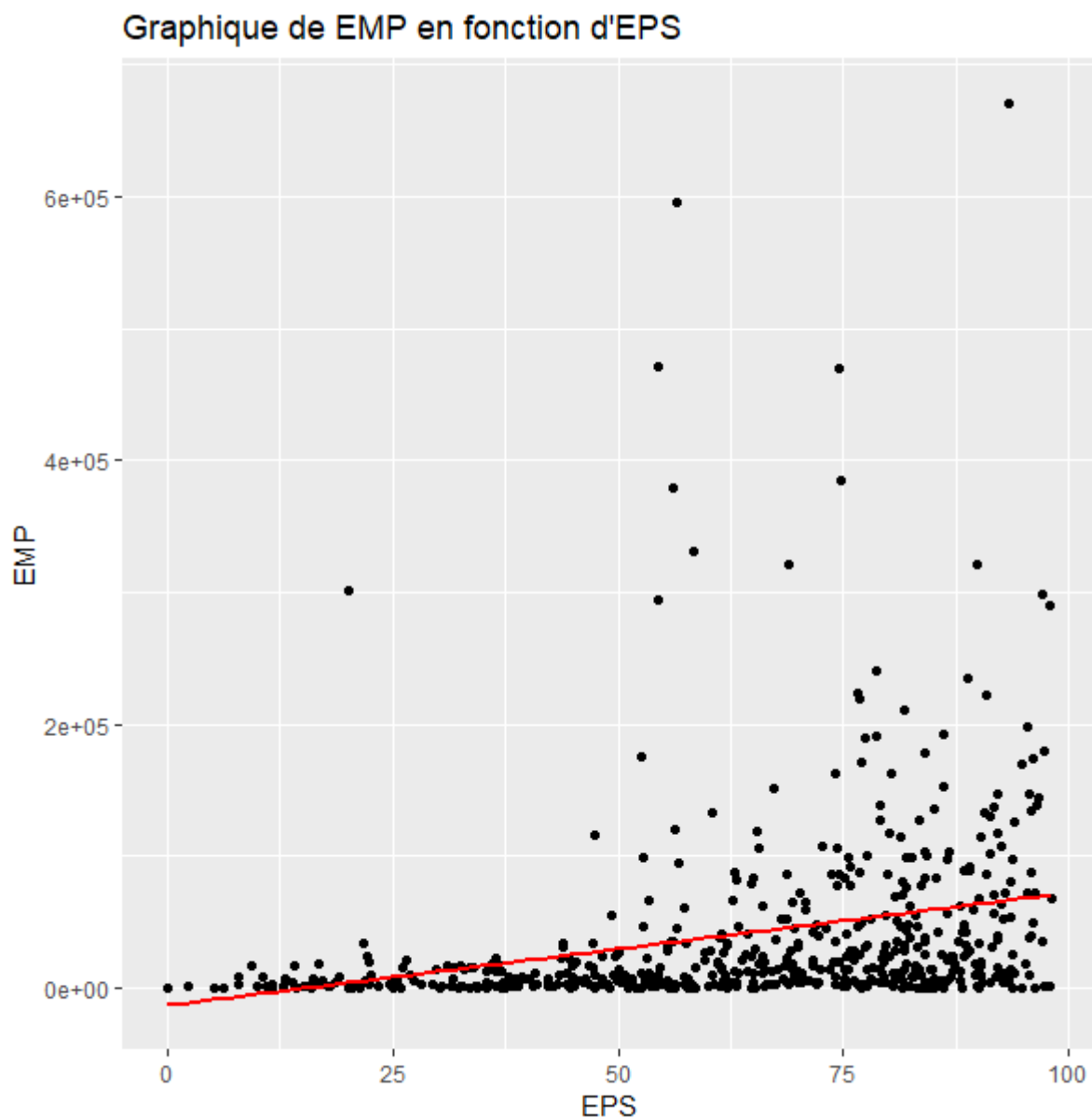
CSR Sustainability Committee

N	107
Y	447

Nous allons désormais nous intéresser à une série de nuage de point qui représente les variables explicatives en fonction de la variable à expliquer et de leurs régressions linéaires :

### *Nombre d'employé en fonction d'EPS*

En premier graphique nous avons un nuage de point du nombre d'employé selon la valeur d'EPS :



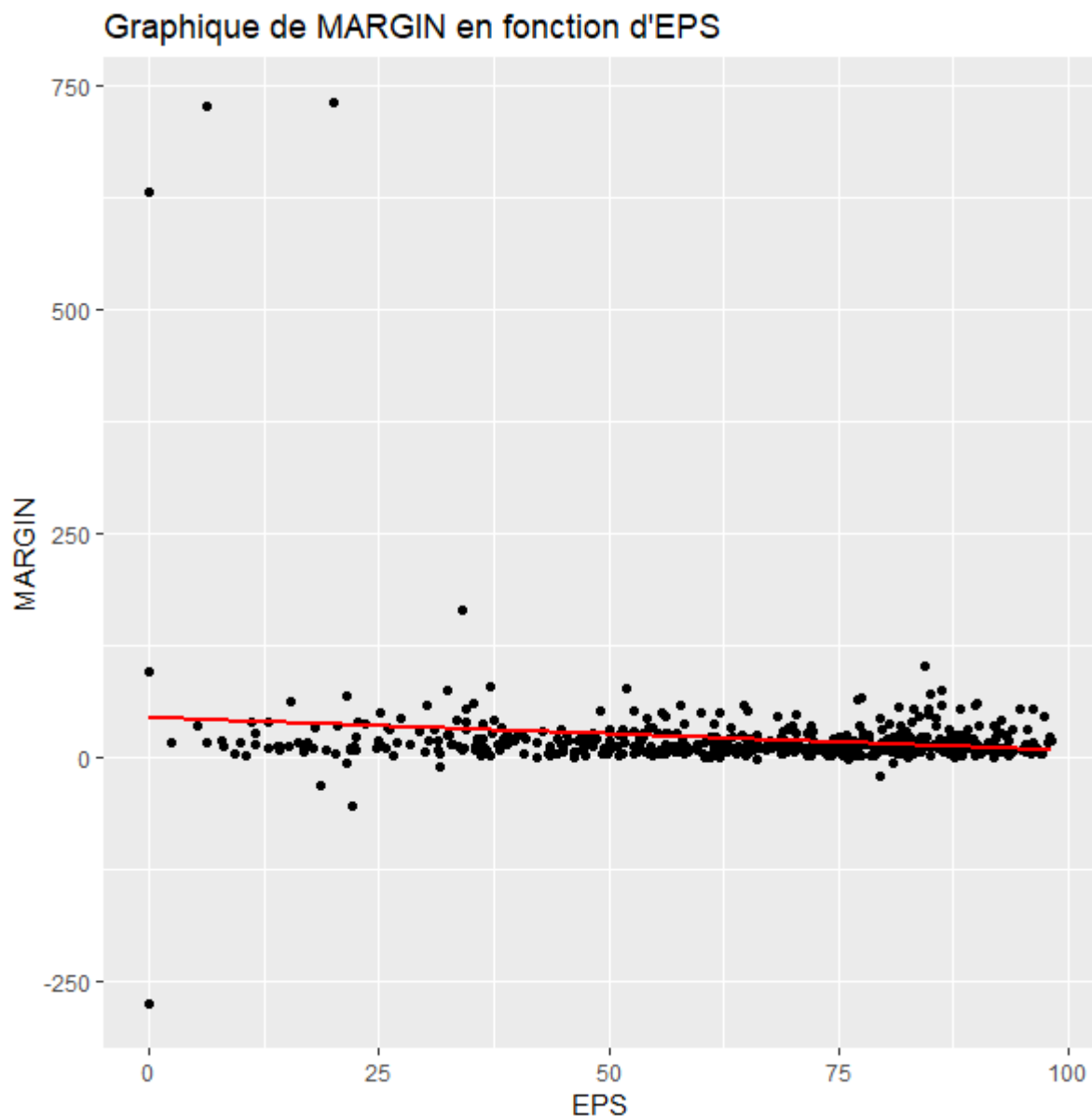
On remarque une certaine dispersion des données par rapports à la régression linéaire. Le fait que l'on ait certains cas extrême (avec une forte valeur pour EMP) fait que la pente



de la régression soit positive et se retrouve au-dessus de notre amas de point qui longe la droite  $y = 0$ .

### *Profit fonctionnel en fonction d'EPS*

Puis nous avons comme second graphique, le profit fonctionnel de l'entreprise en fonction de l'EPS.



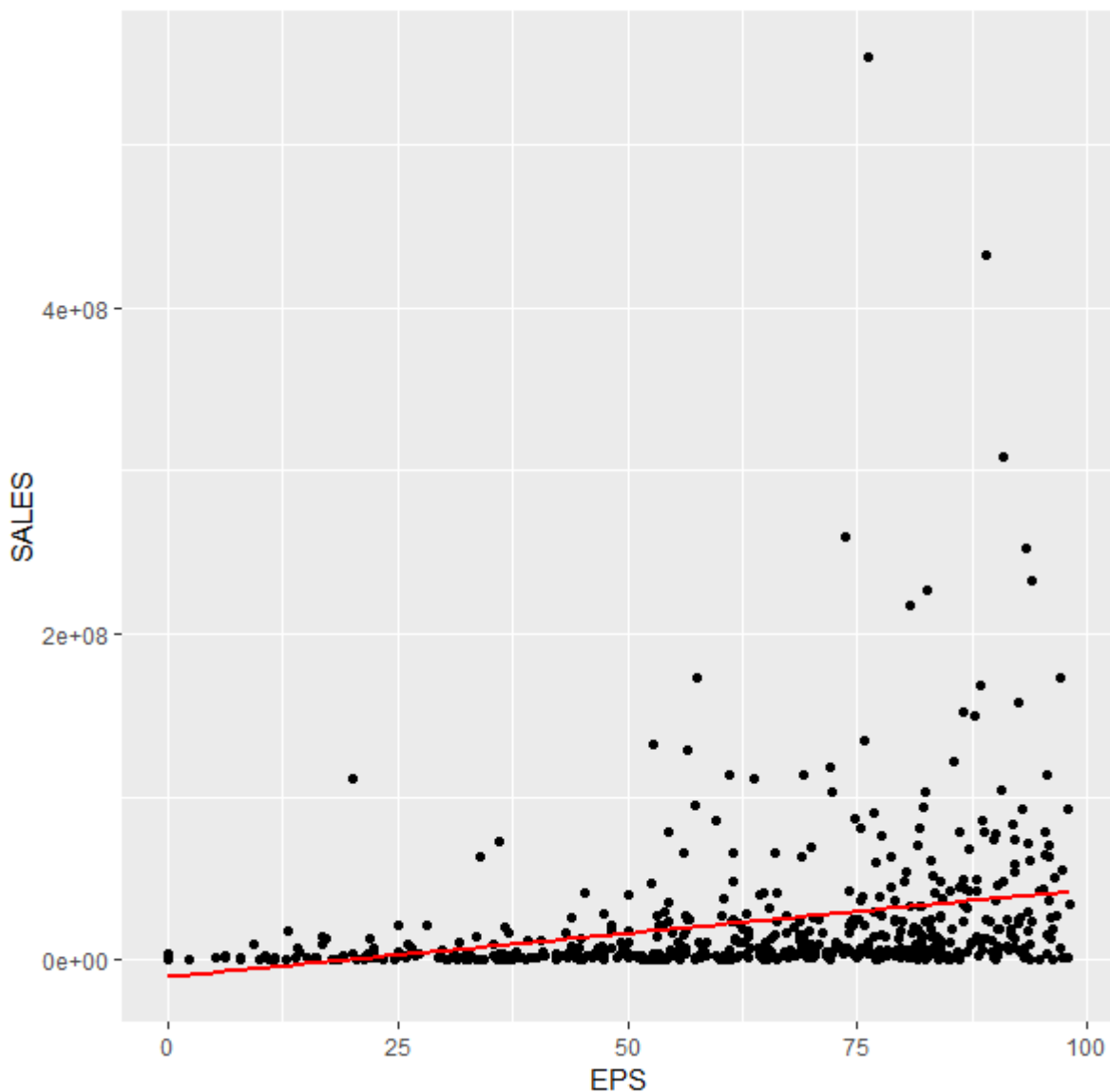
Ici, notre régression linéaire semble suivre l'amas de point. Ainsi, nous avons une pente qui décroît et donc signifie que plus une entreprise voit sa valeur EPS augmenter et plus

le profit opérationnel de l'entreprise diminue (Par analogie, si on augmente le profit opérationnel de l'entreprise la valeur EPS diminue).

### *Revenues commerciaux net en fonction d'EPS*

Nous allons désormais voir les revenus commerciaux net en fonction de l'EPS.

**Graphique de SALES en fonction d'EPS**

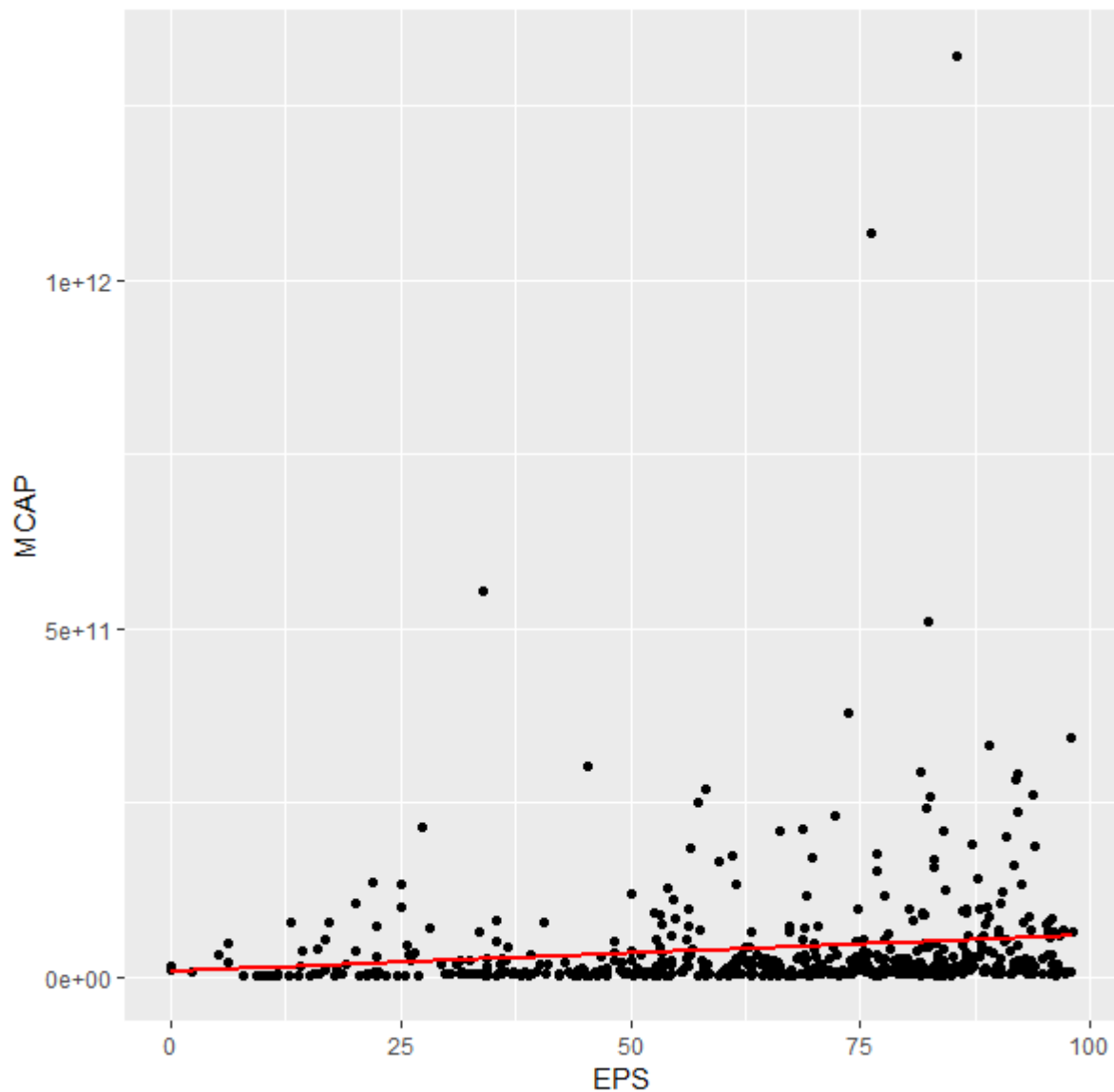


Ici, le graphique ressemble au 1<sup>er</sup> et donc nous effectuons une analyse analogue à la précédente.

## Capitalisation boursière en fonction d'EPS

Notre 4<sup>ème</sup> graphique est la capitalisation boursière de l'entreprise en fonction de l'EPS.

Graphique de MCAP en fonction d'EPS

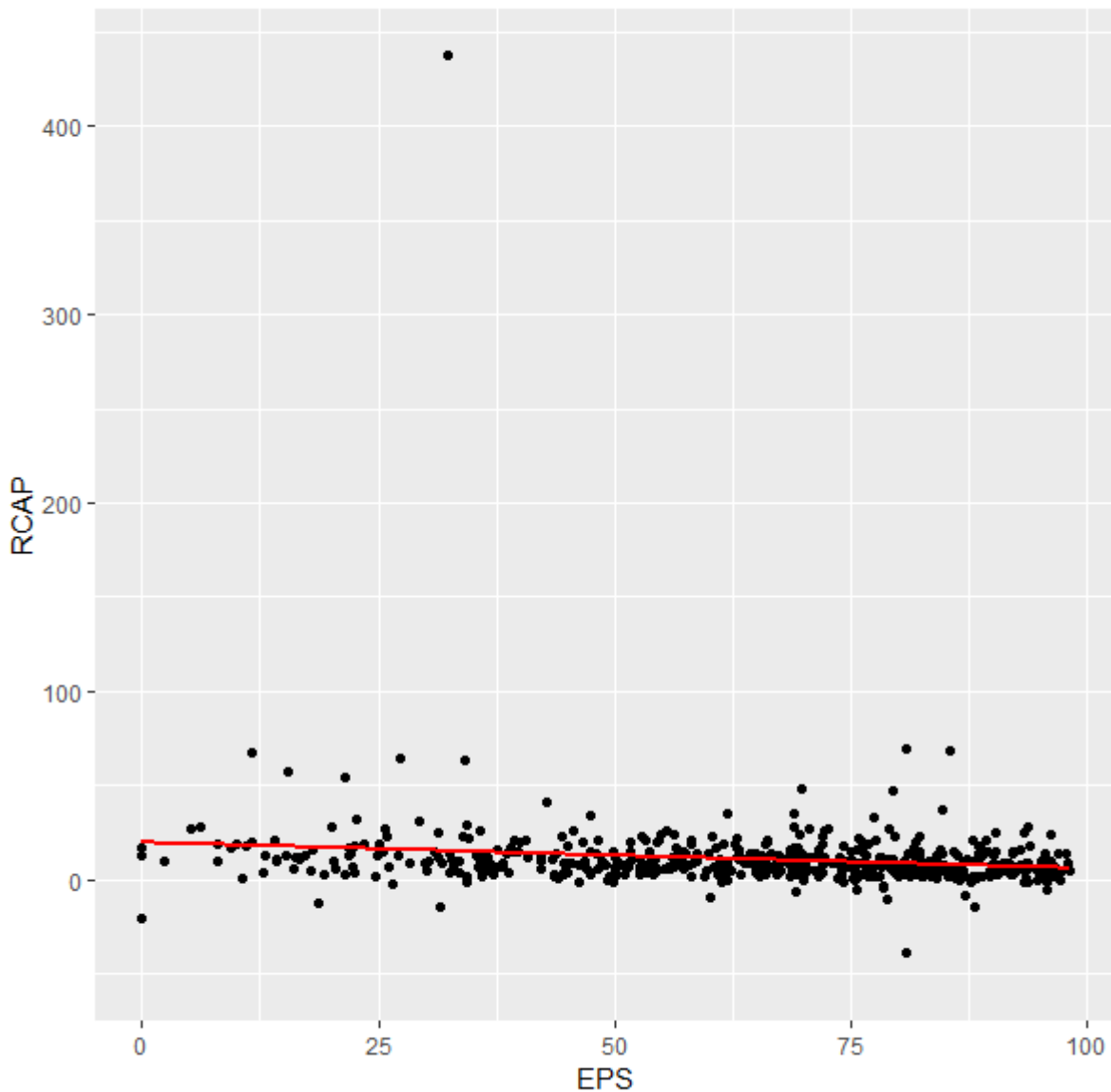


D'après la régression linéaire, on aurait tendance à dire que la capitalisation boursière impacte positivement la variable EPS, or ici on voit un important amas de point le long de la droite  $y = 0$ .

## Rendement du capital investi en fonction d'EPS

5<sup>ème</sup> graphique qui représente le rendement du capital investi selon la variable EPS.

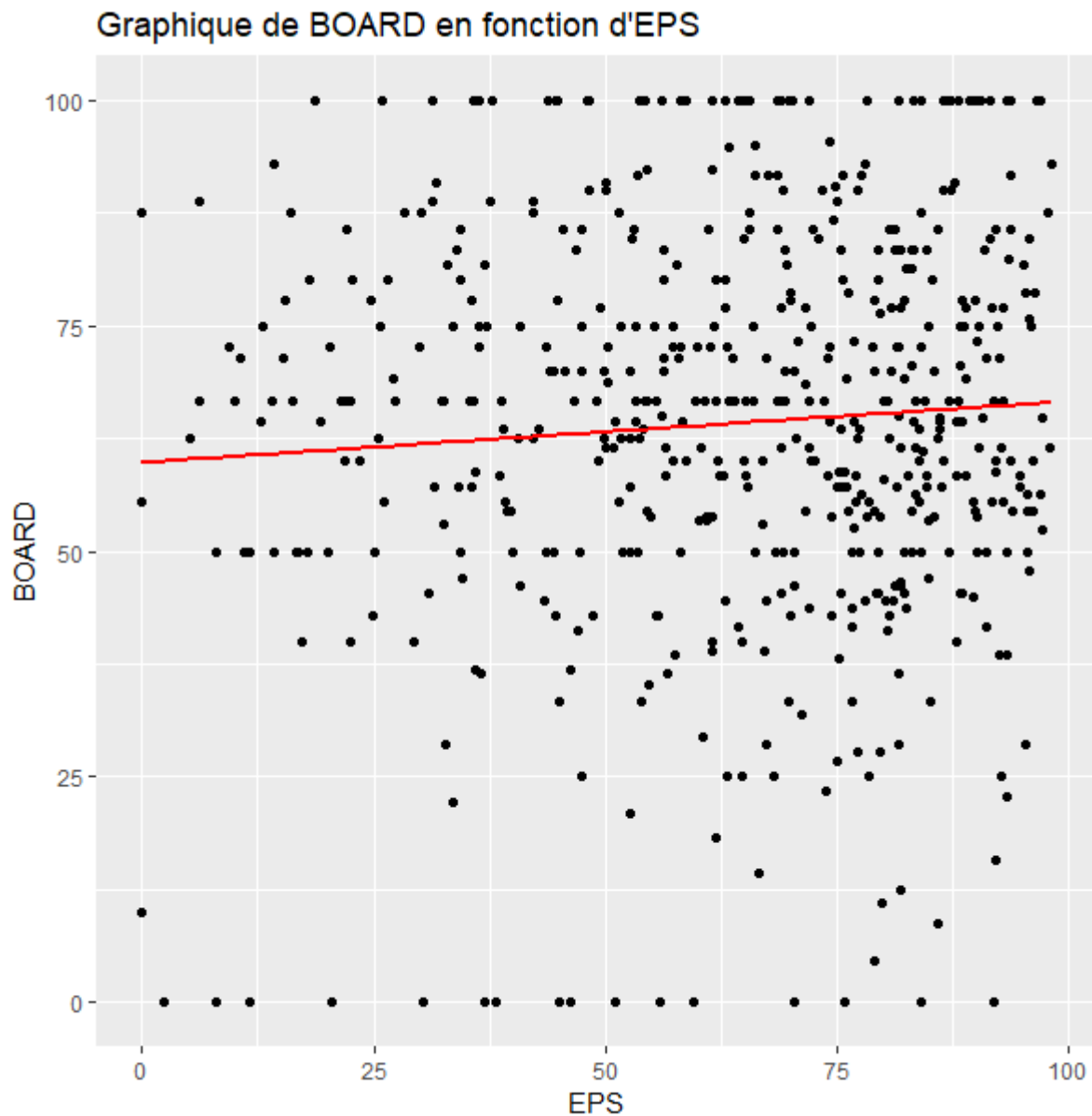
Graphique de RCAP en fonction d'EPS



Notre régression linéaire semble concorder avec notre amas de point et est de pente décroissante. Donc l'augmentation du rendement du capital investi fait décroître la variable EPS.

## Pourcentage de membre indépendant dans le board en fonction d'EPS

6<sup>ème</sup> graphique : le pourcentage de membre indépendant dans le board selon la variable EPS.

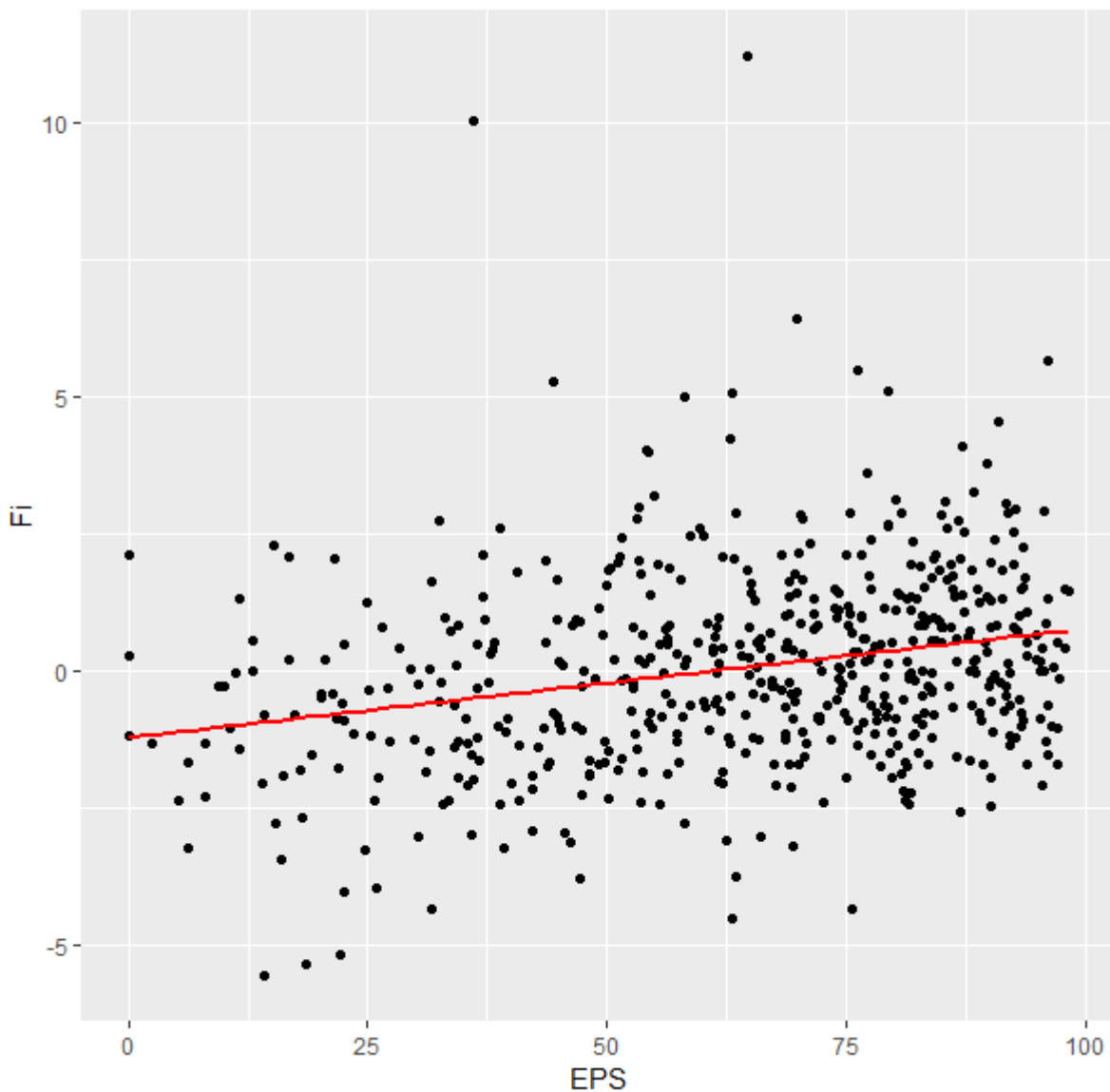


Il semble y avoir une décorrélation entre ces deux variables.

## Fi en fonction d'EPS

7<sup>ème</sup> variable : Le prix en fonction de la variable EPS.

Graphique de Fi en fonction d'EPS

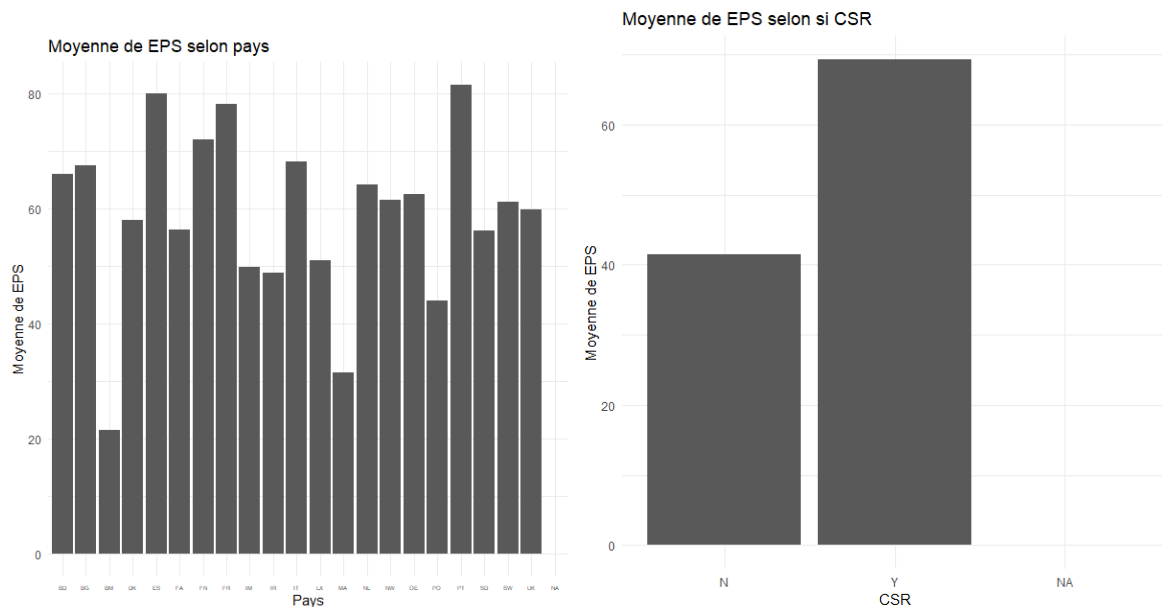


Par analogie au précédent graphique, il semble être décorrélé.

À travers ces graphiques on voit la corrélation entre variable explicative et variable à expliquer.

Et donc, par analyse graphique on peut écarter l'impact de certaines variables due à la disparité de leurs nuages de points tel que le graphique de Board ou de Fi (Non pas qu'elles n'ont aucun impact, mais plutôt qu'il est moins important que d'autre).

Enfin, pour nos variables qualitatives, nous avons opté pour des histogrammes :



Ainsi, au travers de ces divers graphiques, on voit que la présence d'un comité de soutenabilité impact positivement la variable EPS. Tandis que l'autre graphique ne nous donne pas d'informations flagrantes.

*Estimer  $EPS_i$*

On cherche à estimer chaque  $EPS_i$ , dont la formule théorique est la suivante :

$$EPS_i = \alpha + \sum_k b_k X_{k_i} + \epsilon_i$$

Pour le déterminer nous allons utiliser la régression linéaire.

Voici les coefficients déterminés pour les variables quantitatives :

	Estimé
$\alpha$	39,6756451
EMP	2,3479E-05
MARGIN	-0,0299437
SALES	6,6975E-08
MCAP	1,9368E-11
RCAP	-0,0468674
BOARD	0,09378313
Fi	1,57530427

En fonction du pays de domiciliation nous sommes ces valeurs :

CODBG	7,77202349
CODBM	-28,058679
CODDK	-9,4046732
CODES	14,0585799
CODFA	-11,581313
CODFN	6,23399048
CODFR	10,2001279
CODIM	-15,242112
CODIR	-12,302502
CODIT	1,92807162
CODLX	-4,9434033
CODMA	-36,277617
CODNL	-0,4159234
CODNW	-7,774569
CODOE	1,86316349
CODPO	-7,0186274
CODPT	13,5207359
CODSD	-4,1769134
CODSW	-1,1614981
CODUK	-3,2970453

Et enfin, s'il y a un comité de soutenabilité, nous sommes si c'est oui la valeur suivante :

CSRY	19,8923317
------	------------

À partir de ce dernier travail, nous avons pu quantifier l'impact de chacune de ces variables sur la valeur d'EPS. Et donc, pu déterminer en fonction des performances économiques, financières, et des caractéristiques de gouvernance, la performance environnementale des entreprises du STOXX600.