

# Projet Statistiques des assurances

Isabelle Ajtay (41010932)      Smail Chabane (38012939)      Yuxuan Zhang (38019811)

19 février 2023

## Table des matières

<b>Plan d'analyse</b>	<b>1</b>
<b>Les données</b>	<b>1</b>
Import des données . . . . .	1
Analyses univariées . . . . .	1
2.3.2 Analyses bivariées . . . . .	3
<b>Modélisation de Sinistre0</b>	<b>10</b>
Tester les problèmes d'endogénéité de certaines variables . . . . .	10
IV regressions . . . . .	16

## Plan d'analyse

L'approche que nous allons suivre pour cette étude :

- préparation des données pour traitement : importer, vérifier et découper les données
- analyse des corrélations des variables, deux à deux, et analyse bivariée
- analyse en composantes principales
- classification hiérarchique ascendante; interprétations dans le contexte historique.

## Les données

### Import des données

La procédure pour lire ligne par ligne ces données est longue. Donc nous les avons exportées dans un fichier .txt pour aller plus vite.

```
data = read.table("data.txt", sep = " ", header=T, encoding = "UTF-8")
#data = read.table("assurance_complete_corrige.R") #, sep = "", header=T)
```

### Analyses univariées

On a utilisé str pour afficher les informations simples concernant les variables, et summary pour afficher les données statistiques pour chaque variable.

On a représenté les boxplots des variables RUC et Sinistre0. Pour écraser les grandes valeurs, on utilise la fonction log.

```

par(mfrow=c(3,3))
boxplot(log(data[,2]),main="Boxplot de variable RUC")
hist(data[,2],main="Histogramme de variable RUC")
# erreur ici chez Eva : stat_density requires an x or y aesthetic ggplot(data = data.frame(data[,2]))
boxplot(data[,27],main="Boxplot de variable Sinistre0")
hist(data[,27],main="Histogramme de variable Sinistre0")
density(data[,27])

```

```

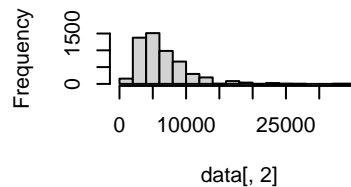
##
## Call:
## density.default(x = data[, 27])
##
## Data: data[, 27] (5352 obs.); Bandwidth 'bw' = 0.6943
##
##      x              y
## Min.   :-1.118   Min.   :1.270e-06
## 1st Qu.: 6.509   1st Qu.:8.995e-04
## Median :14.136   Median :2.309e-02
## Mean   :14.136   Mean    :3.275e-02
## 3rd Qu.:21.763   3rd Qu.:4.731e-02
## Max.   :29.390   Max.    :1.226e-01

```

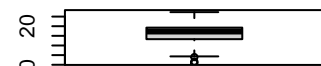
**Boxplot de variable RUC**



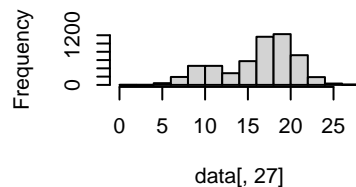
**Histogramme de variable RUC**



**Boxplot de variable Sinistre0**



**Histogramme de variable Sinisti**

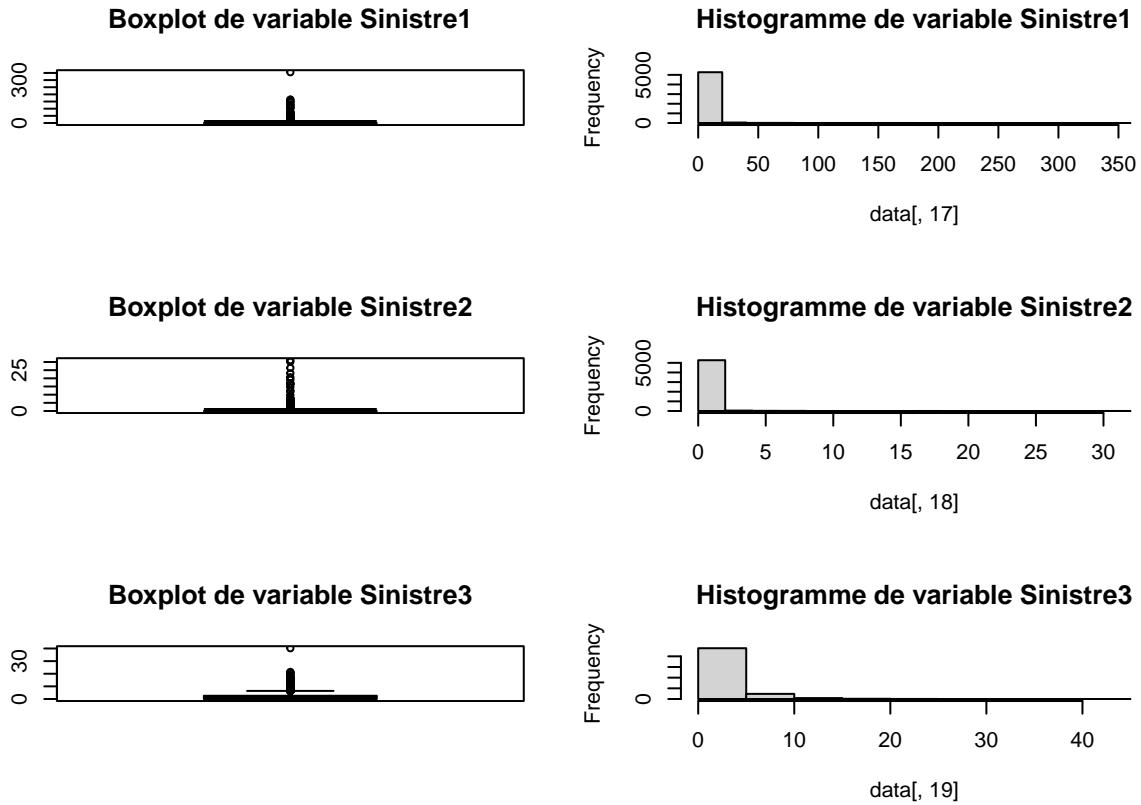


```

par(mfrow=c(3,2))
boxplot(data[,17],main="Boxplot de variable Sinistre1")
hist(data[,17],main="Histogramme de variable Sinistre1")
boxplot(data[,18],main="Boxplot de variable Sinistre2")

```

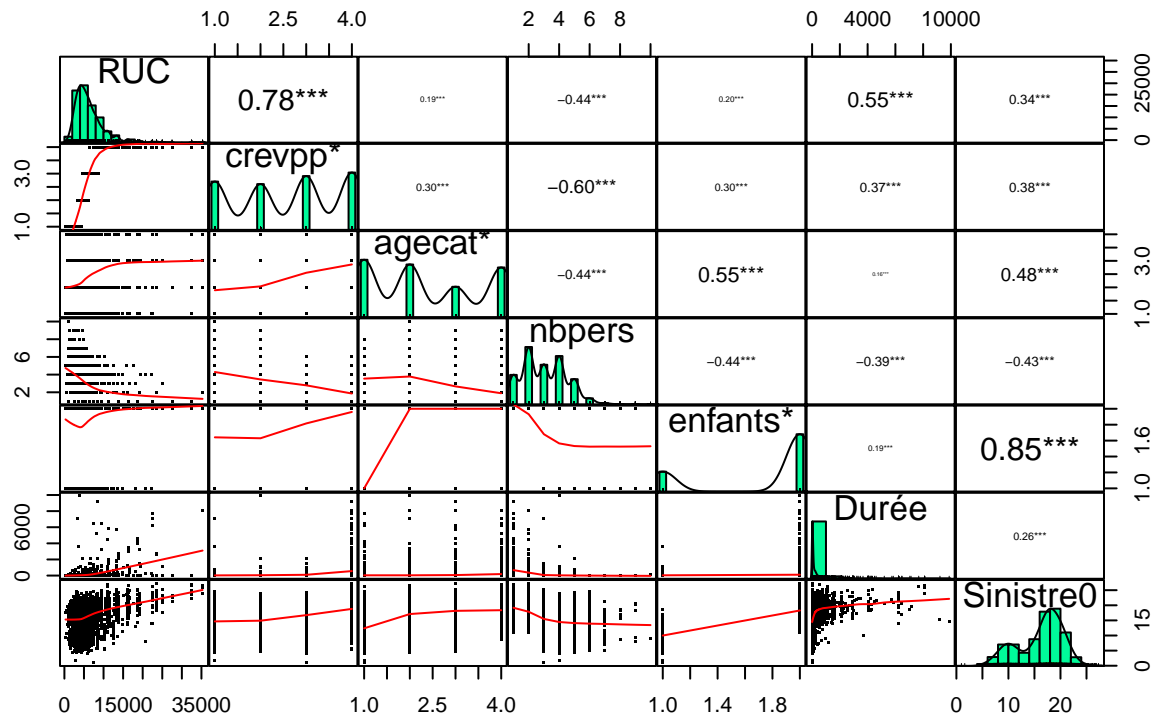
```
hist(data[,18],main="Histogramme de variable Sinistre2")
boxplot(data[,19],main="Boxplot de variable Sinistre3")
hist(data[,19],main="Histogramme de variable Sinistre3")
```



### 2.3.2 Analyses bivariées

A travers un graphique des variables numériques deux à deux, nous regardons comment évoluent les variables ensemble, et s'il y a des "tendances" reconnaissables. Par exemple, sur le graphique ci-dessous, qui contient les analyses bivariées complètes, on voit une tendance linéaire croissante (et corrélation positive significative) entre *pib* et *recc*. (fonction trouvée à ref. 7: analyse bi + corrélations).

## Correlations entre variables deux par deux



On peut constater que les variabilités des Sinistres des 3 types sont toutes grandes.

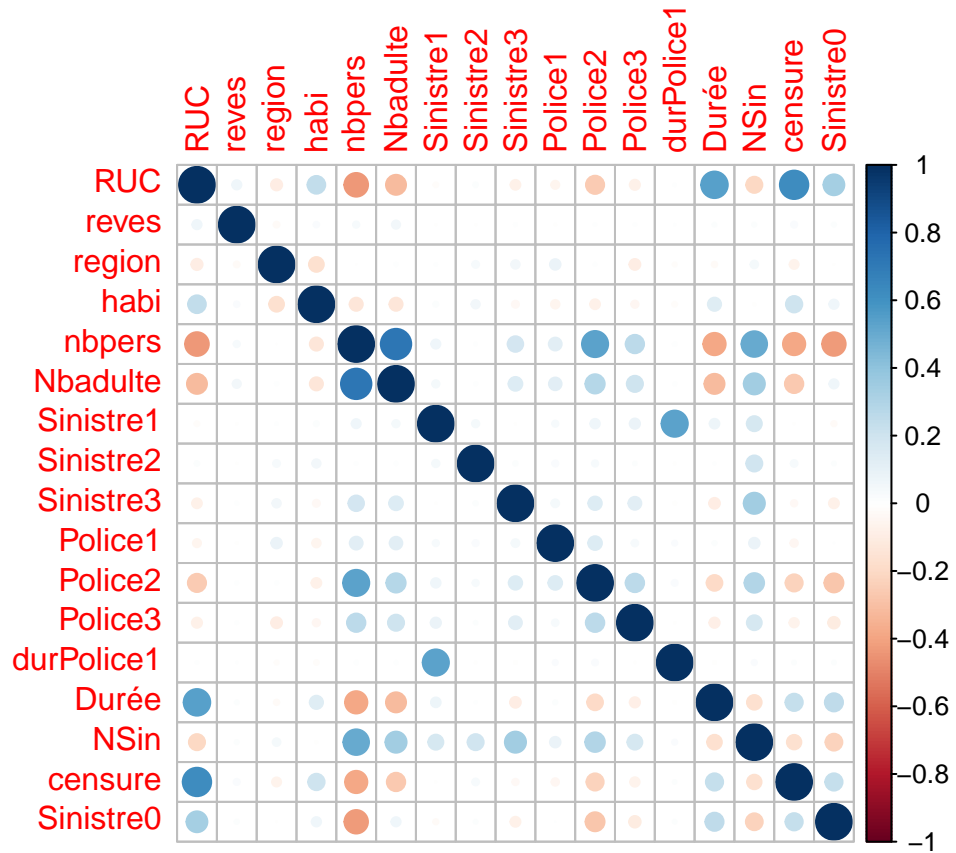
```
aggregate(data[,c(17,18,19,27)],list(data[,1]),mean)
```

```
##           Group.1 Sinistre1 Sinistre2 Sinistre3
## 1 Agr. exploitants 0.4860776 0.009051724 2.983017
## 2 Artisans, comm., chefs d'ent. 0.3559116 0.097458564 1.836381
## 3 Autres pers. sans activite prof. 1.4143169 0.107540984 1.761721
## 4 Cadres et prof. intellectuelles sup. 1.6937083 0.184058333 2.249010
## 5 Employes 1.2392184 0.159409429 1.771824
## 6 Ouvriers 1.6113979 0.128146194 1.841745
## 7 Professions intermediaires 1.7689008 0.207553551 2.214859
## 8 Retraites 0.5480867 0.185876093 1.395999
## Sinistre0
## 1 15.01849
## 2 16.12462
## 3 17.41922
## 4 16.83139
## 5 15.57469
## 6 14.30565
## 7 15.66231
## 8 18.36901
```

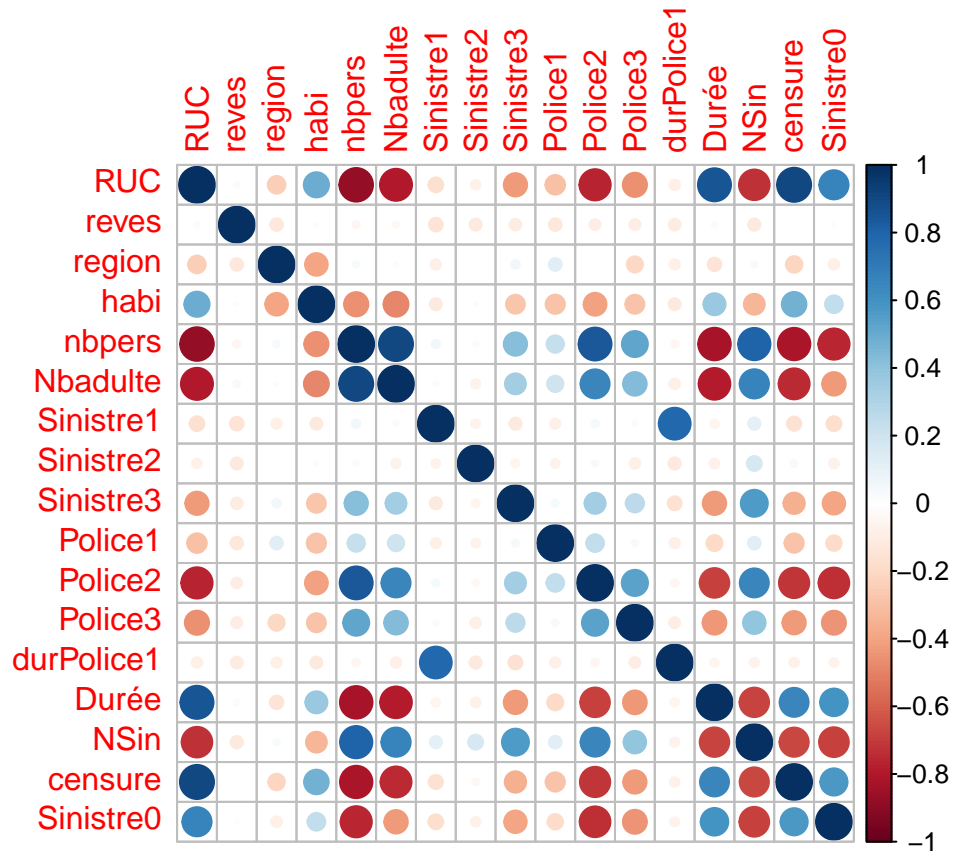
```
variables_quantitatives = data %>% select_if(is.numeric) %>% cor()
kable(variables_quantitatives, digits=3)
```

	RUC	reves	region	habi	nbpers	Nbadulte	Sinistre1	Sinistre2	Sinistre3	Police1	Police2	Police3	durPolice	Durée	NSin	censure	Sinistre0
RUC	1.000	0.066	-	0.244	-	-	-	0.017	-	-	-	-	0.005	0.549	-	0.628	0.338
		0.097		0.436	0.311	0.018		0.075	0.051	0.256	0.080			0.201			
reves	0.066	1.000	-	0.024	0.031	0.053	-	-	0.007	-	0.006	0.005	0.001	0.012	0.013	0.029	0.014
		0.027				0.001	0.001		0.002								
region	-	-	1.000	-	0.000	0.008	0.005	0.037	0.056	0.087	0.007	-	-	-	0.046	-	0.000
	0.097	0.027		0.170								0.092	0.013	0.023		0.063	
habi	0.244	0.024	-	1.000	-	-	0.019	0.054	-	-	-	-	-	0.136	-	0.209	0.068
		0.170		0.133	0.133			0.039	0.057	0.075	0.041	0.018		0.003			
nbpers	-	0.031	0.000	-	1.000	0.720	0.064	0.006	0.188	0.129	0.531	0.262	0.009	-	0.508	-	-
	0.436		0.133											0.387		0.385	0.429
Nbadulte	-	0.053	0.008	-	0.720	1.000	0.042	0.008	0.148	0.124	0.288	0.202	0.008	-	0.346	-	0.066
	0.311		0.133											0.312		0.262	
Sinistre1	-	-	0.005	0.019	0.064	0.042	1.000	0.042	-	0.030	0.070	0.082	0.534	0.073	0.175	-	-
	0.018	0.001						0.001								0.001	0.021
Sinistre2	0.017	-	0.037	0.054	0.006	0.008	0.042	1.000	0.015	0.029	0.031	0.010	-	0.004	0.200	0.037	0.012
	0.001												0.002				
Sinistre3	-	0.007	0.056	-	0.188	0.148	-	0.015	1.000	0.048	0.140	0.130	-	-	0.346	-	-
	0.075		0.039		0.001								0.006	0.090		0.032	0.077
Police1	-	-	0.087	-	0.129	0.124	0.030	0.029	0.048	1.000	0.146	0.035	0.020	0.013	0.087	-	0.000
	0.051	0.002	0.057													0.044	
Police2	-	0.006	0.007	-	0.531	0.288	0.070	0.031	0.140	0.146	1.000	0.265	0.028	-	0.297	-	-
	0.256		0.075											0.193		0.228	0.280
Police3	-	0.005	-	-	0.262	0.202	0.082	0.010	0.130	0.035	0.265	1.000	-	-	0.172	-	-
	0.080	0.092	0.041										0.001	0.082		0.069	0.103
durPolice	0.005	0.001	-	-	0.009	0.008	0.534	-	-	0.020	0.028	-	1.000	-	0.021	0.011	0.011
		0.013	0.018				0.002	0.006				0.001		0.005			
Durée	0.549	0.012	-	0.136	-	-	0.073	0.004	-	0.013	-	-	-	1.000	-	0.231	0.256
		0.023		0.387	0.312			0.090		0.193	0.082	0.005		0.167			
NSin	-	0.013	0.046	-	0.508	0.346	0.175	0.200	0.346	0.087	0.297	0.172	0.021	-	1.000	-	-
	0.201		0.003											0.167		0.162	0.221
censure	0.628	0.029	-	0.209	-	-	-	0.037	-	-	-	-	0.011	0.231	-	1.000	0.231
		0.063		0.385	0.262	0.001		0.032	0.044	0.228	0.069			0.162			
Sinistre0	0.338	0.014	0.000	0.068	-	0.066	-	0.012	-	0.000	-	-	0.011	0.256	-	0.231	1.000
				0.429		0.021		0.077		0.280	0.103			0.221			

```
corrplot(variables_quantitatives)
```

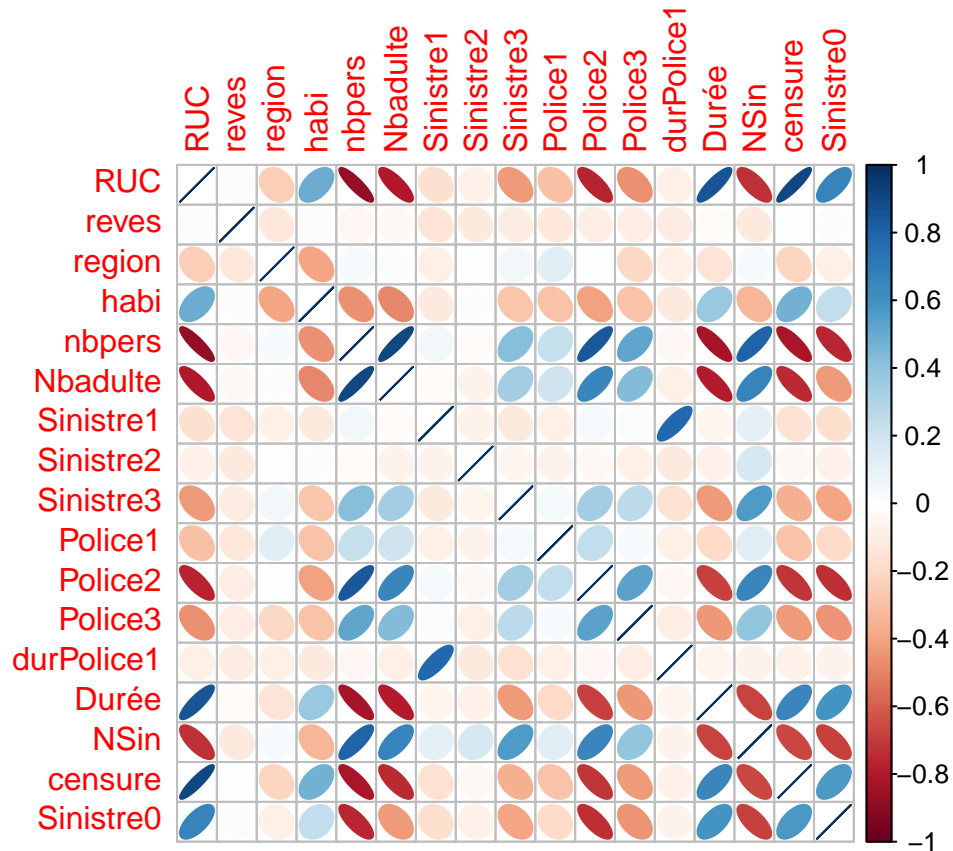


```
corrplot(cor(variables_quantitatives))
```



*#Representation des correlations : plus l'ellipse ressemble a un cercle et moins les variables sont correlees. Plus l'ellipse ressemble a une droite et plus les variables sont correlees.*

```
corrplot(cor(variables_quantitatives),method = "ellipse")
```



```
# Boξplot
# boxplot(data$column_name, main = "Boξplot of column_name", xlab = "Column name", ylab = "Values")

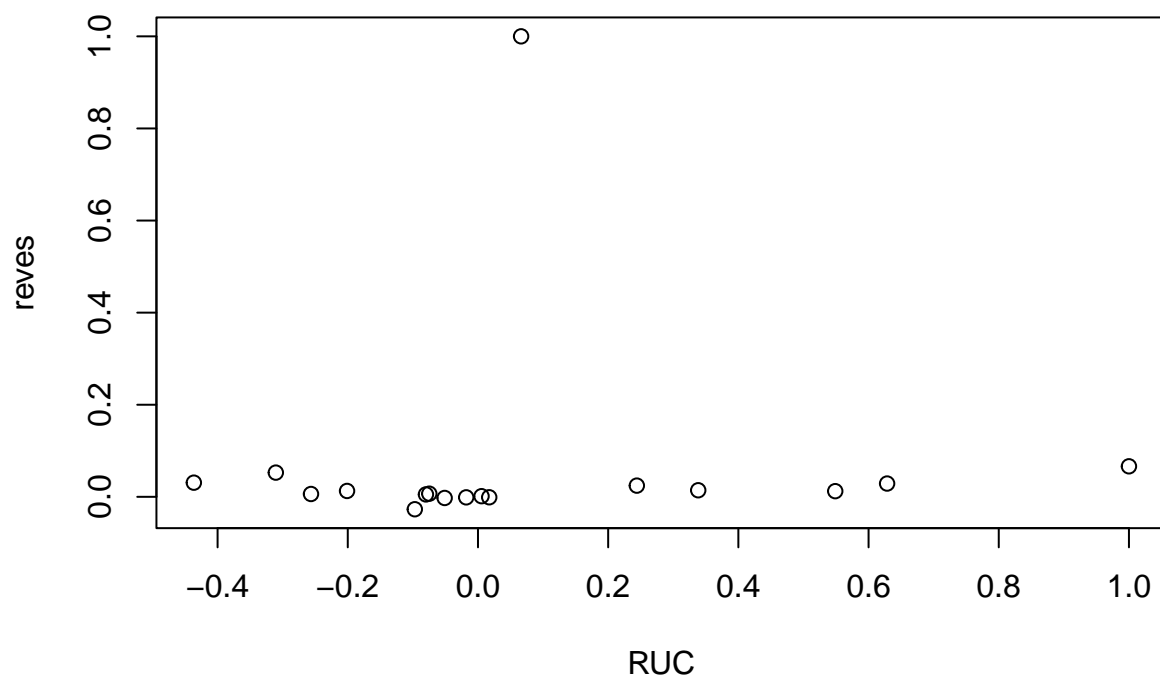
# Histogram
# hist(data$column_name, main = "Histogram of column_name", xlab = "Values", ylab = "Frequency", col =

# Estimateurs de la densit  
# density_plot <- density(data$column_name, main = "Density Plot of column_name", xlab = "Values", ylab
# lines(density_plot, col = "red")

# Statistiques basiques
# summary(data$column_name)

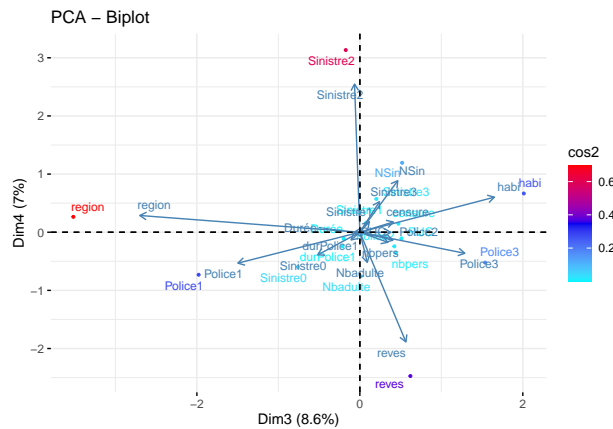
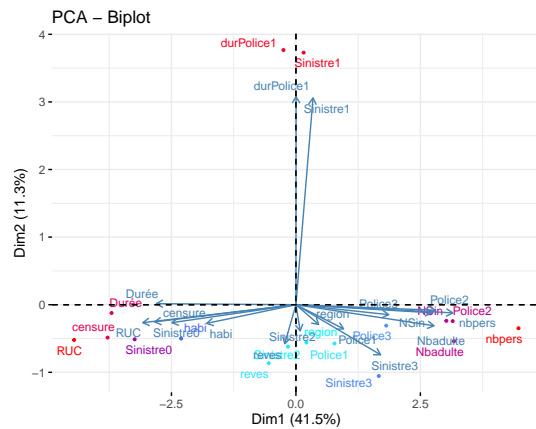
plot(variables_quantitatives)
```





```
## **Results for the Principal Component Analysis (PCA)**
## The analysis was performed on 17 individuals, described by 17 variables
## *The results are available in the following objects:
##
##   name                description
## 1  "$eig"              "eigenvalues"
## 2  "$var"              "results for the variables"
## 3  "$var$coord"        "coord. for the variables"
## 4  "$var$cor"          "correlations variables - dimensions"
## 5  "$var$cos2"         "cos2 for the variables"
## 6  "$var$contrib"      "contributions of the variables"
## 7  "$ind"              "results for the individuals"
## 8  "$ind$coord"        "coord. for the individuals"
## 9  "$ind$cos2"         "cos2 for the individuals"
## 10 "$ind$contrib"      "contributions of the individuals"
## 11 "$call"             "summary statistics"
## 12 "$call$centre"      "mean of the variables"
## 13 "$call$ecart.type"  "standard error of the variables"
## 14 "$call$row.w"       "weights for the individuals"
## 15 "$call$col.w"       "weights for the variables"
##
##   eigenvalue percentage of variance cumulative percentage of variance
## comp 1      7.05                41.49                41.49
## comp 2      1.92                11.30                52.78
## comp 3      1.47                 8.63                61.42
## comp 4      1.19                 7.01                68.42
## comp 5      1.03                 6.06                74.49
```

## comp 6	0.94	5.50	79.99
## comp 7	0.75	4.39	84.38
## comp 8	0.66	3.87	88.25
## comp 9	0.64	3.77	92.02
## comp 10	0.38	2.26	94.27
## comp 11	0.35	2.04	96.31
## comp 12	0.28	1.64	97.95
## comp 13	0.17	0.99	98.94
## comp 14	0.13	0.75	99.70
## comp 15	0.04	0.22	99.91
## comp 16	0.01	0.09	100.00



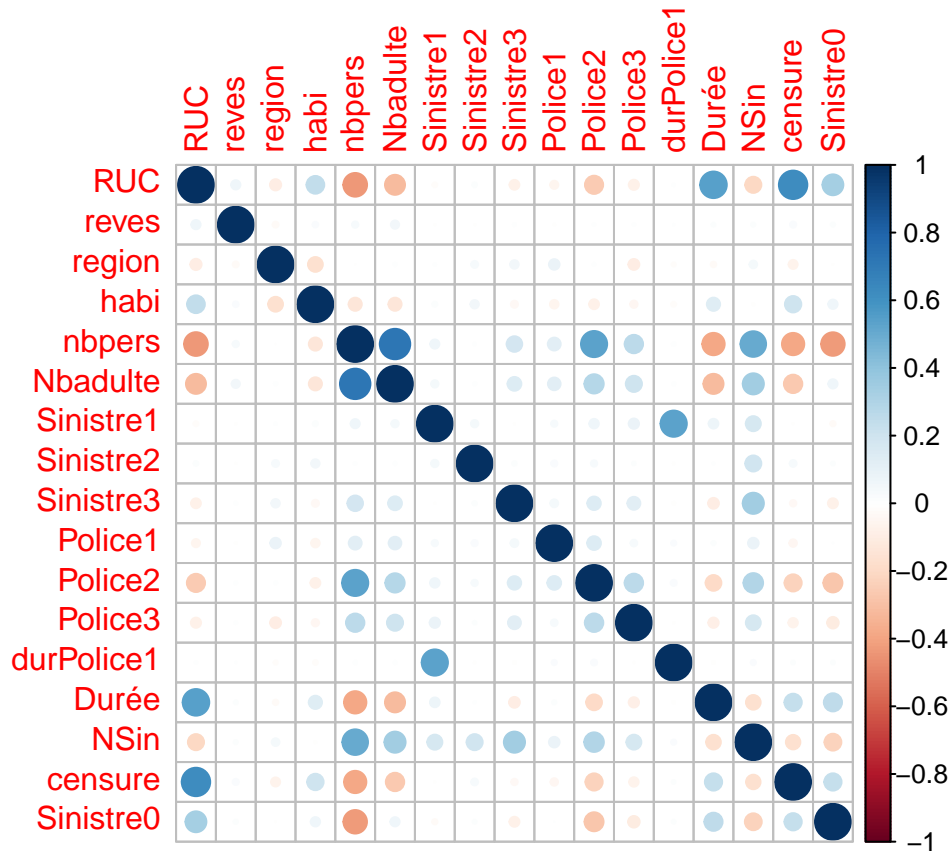
## Modélisation de Sinistre0

La méthode des MCO donne l'estimateur le plus efficient s'il n'y a pas d'endogénéité. S'il y a de l'endogénéité, OLS (MCO) va donner des résultats inconsistants. L'estimateur des variables instrumentales va être consistant, mais inefficent.

### Tester les problèmes d'endogénéité de certaines variables

```
# Selection des variables quantitatives
quant_vars <- sapply(data, is.numeric)

# Matrice de correlation
cor_matrix <- cor(data[, quant_vars])
corrplot(cor_matrix)
```



*# Tests de causalité de Granger pour toutes les paires de variables*

```
quant_vars <- as.matrix(quant_vars)
d = data[, quant_vars]
for(i in 1:(ncol(d) - 1)){
  for(j in (i + 1):ncol(d)){

    result <- grangertest(d[,i], d[,j], order = 2)

    print(paste("Granger causality test entre ", colnames(d)[i],
                " et ", colnames(d)[j], ":", result[2,4]))
  }
}
```

```
## [1] "Granger causality test entre RUC et reves : 0.0923837756616444"
## [1] "Granger causality test entre RUC et region : 0.453886215308336"
## [1] "Granger causality test entre RUC et habi : 0.0040326134264574"
## [1] "Granger causality test entre RUC et nbpers : 0.00124220485708754"
## [1] "Granger causality test entre RUC et Nbadulte : 0.0172836315703496"
## [1] "Granger causality test entre RUC et Sinistre1 : 0.709483878283177"
## [1] "Granger causality test entre RUC et Sinistre2 : 0.0951611194510821"
## [1] "Granger causality test entre RUC et Sinistre3 : 0.0981000677158548"
## [1] "Granger causality test entre RUC et Police1 : 0.737522698532588"
## [1] "Granger causality test entre RUC et Police2 : 0.0123843096814322"
## [1] "Granger causality test entre RUC et Police3 : 0.912874664962623"
## [1] "Granger causality test entre RUC et durPolice1 : 0.234660821645245"
## [1] "Granger causality test entre RUC et Durée : 0.00278908899804597"
## [1] "Granger causality test entre RUC et NSin : 0.283574612045495"
```

```

## [1] "Granger causality test entre RUC et censure : 5.81376380854141e-09"
## [1] "Granger causality test entre RUC et Sinistre0 : 0.00736870516547412"
## [1] "Granger causality test entre reves et region : 0.854231768457558"
## [1] "Granger causality test entre reves et habi : 0.106418079520996"
## [1] "Granger causality test entre reves et nbpers : 0.18968408454619"
## [1] "Granger causality test entre reves et Nbadulte : 0.36176438407316"
## [1] "Granger causality test entre reves et Sinistre1 : 0.970631057503471"
## [1] "Granger causality test entre reves et Sinistre2 : 0.934632216327838"
## [1] "Granger causality test entre reves et Sinistre3 : 0.87081127385712"
## [1] "Granger causality test entre reves et Police1 : 0.465186907224791"
## [1] "Granger causality test entre reves et Police2 : 0.366262300097807"
## [1] "Granger causality test entre reves et Police3 : 0.38259062929957"
## [1] "Granger causality test entre reves et durPolice1 : 0.99606824048349"
## [1] "Granger causality test entre reves et Durée : 0.89938550170445"
## [1] "Granger causality test entre reves et NSin : 0.956546525314893"
## [1] "Granger causality test entre reves et censure : 0.45554355536769"
## [1] "Granger causality test entre reves et Sinistre0 : 0.365219373654749"
## [1] "Granger causality test entre region et habi : 0.211123548273796"
## [1] "Granger causality test entre region et nbpers : 0.208159018409928"
## [1] "Granger causality test entre region et Nbadulte : 0.0972180984581644"
## [1] "Granger causality test entre region et Sinistre1 : 0.733192530883739"
## [1] "Granger causality test entre region et Sinistre2 : 0.115650506372673"
## [1] "Granger causality test entre region et Sinistre3 : 0.0141363787776151"
## [1] "Granger causality test entre region et Police1 : 2.76920371834171e-07"
## [1] "Granger causality test entre region et Police2 : 0.593118834178275"
## [1] "Granger causality test entre region et Police3 : 1.69581158888049e-08"
## [1] "Granger causality test entre region et durPolice1 : 0.609423624450731"
## [1] "Granger causality test entre region et Durée : 0.944334780054465"
## [1] "Granger causality test entre region et NSin : 0.0231062889761891"
## [1] "Granger causality test entre region et censure : 0.0214982891651647"
## [1] "Granger causality test entre region et Sinistre0 : 0.383659503770951"
## [1] "Granger causality test entre habi et nbpers : 3.13229775443005e-05"
## [1] "Granger causality test entre habi et Nbadulte : 4.88769676827836e-08"
## [1] "Granger causality test entre habi et Sinistre1 : 0.348720632975581"
## [1] "Granger causality test entre habi et Sinistre2 : 0.0177773059617933"
## [1] "Granger causality test entre habi et Sinistre3 : 0.131886066847617"
## [1] "Granger causality test entre habi et Police1 : 0.0438289123376355"
## [1] "Granger causality test entre habi et Police2 : 0.00991570004260132"
## [1] "Granger causality test entre habi et Police3 : 0.423185286680116"
## [1] "Granger causality test entre habi et durPolice1 : 0.399204570724914"
## [1] "Granger causality test entre habi et Durée : 4.01905405586965e-10"
## [1] "Granger causality test entre habi et NSin : 0.164590235003095"
## [1] "Granger causality test entre habi et censure : 1.60540480858891e-19"
## [1] "Granger causality test entre habi et Sinistre0 : 0.0575796813459293"
## [1] "Granger causality test entre nbpers et Nbadulte : 0.193083295586577"
## [1] "Granger causality test entre nbpers et Sinistre1 : 0.558110531140403"
## [1] "Granger causality test entre nbpers et Sinistre2 : 0.49305523182955"
## [1] "Granger causality test entre nbpers et Sinistre3 : 0.944191036071153"
## [1] "Granger causality test entre nbpers et Police1 : 0.0381762395658839"
## [1] "Granger causality test entre nbpers et Police2 : 0.38758443522889"
## [1] "Granger causality test entre nbpers et Police3 : 0.79937810796575"
## [1] "Granger causality test entre nbpers et durPolice1 : 0.272894547109501"
## [1] "Granger causality test entre nbpers et Durée : 0.0905118575346298"
## [1] "Granger causality test entre nbpers et NSin : 0.0480562871228446"

```

```

## [1] "Granger causality test entre nbpers et censure : 0.00455159012190975"
## [1] "Granger causality test entre nbpers et Sinistre0 : 0.362139149025635"
## [1] "Granger causality test entre Nbadulte et Sinistre1 : 0.358541171759482"
## [1] "Granger causality test entre Nbadulte et Sinistre2 : 0.683643172870649"
## [1] "Granger causality test entre Nbadulte et Sinistre3 : 0.954146981569278"
## [1] "Granger causality test entre Nbadulte et Police1 : 0.130836636428486"
## [1] "Granger causality test entre Nbadulte et Police2 : 0.659774129946827"
## [1] "Granger causality test entre Nbadulte et Police3 : 0.50444501848523"
## [1] "Granger causality test entre Nbadulte et durPolice1 : 0.275432907148821"
## [1] "Granger causality test entre Nbadulte et Durée : 0.0210922967848045"
## [1] "Granger causality test entre Nbadulte et NSin : 0.381605011258915"
## [1] "Granger causality test entre Nbadulte et censure : 0.00262936227638448"
## [1] "Granger causality test entre Nbadulte et Sinistre0 : 0.348576217380067"
## [1] "Granger causality test entre Sinistre1 et Sinistre2 : 0.827711518387258"
## [1] "Granger causality test entre Sinistre1 et Sinistre3 : 0.570989882925704"
## [1] "Granger causality test entre Sinistre1 et Police1 : 0.736836059511227"
## [1] "Granger causality test entre Sinistre1 et Police2 : 0.288643457072958"
## [1] "Granger causality test entre Sinistre1 et Police3 : 0.6554707617976"
## [1] "Granger causality test entre Sinistre1 et durPolice1 : 0.976563675926668"
## [1] "Granger causality test entre Sinistre1 et Durée : 0.473758676644474"
## [1] "Granger causality test entre Sinistre1 et NSin : 0.36196682034159"
## [1] "Granger causality test entre Sinistre1 et censure : 0.38432432846914"
## [1] "Granger causality test entre Sinistre1 et Sinistre0 : 0.134747038950314"
## [1] "Granger causality test entre Sinistre2 et Sinistre3 : 0.300524713276513"
## [1] "Granger causality test entre Sinistre2 et Police1 : 0.10246562700862"
## [1] "Granger causality test entre Sinistre2 et Police2 : 0.158801981749613"
## [1] "Granger causality test entre Sinistre2 et Police3 : 0.370757787884079"
## [1] "Granger causality test entre Sinistre2 et durPolice1 : 0.98055544905637"
## [1] "Granger causality test entre Sinistre2 et Durée : 0.520300685937179"
## [1] "Granger causality test entre Sinistre2 et NSin : 0.498259153148845"
## [1] "Granger causality test entre Sinistre2 et censure : 0.652951329078083"
## [1] "Granger causality test entre Sinistre2 et Sinistre0 : 0.355856344901385"
## [1] "Granger causality test entre Sinistre3 et Police1 : 0.201632295337263"
## [1] "Granger causality test entre Sinistre3 et Police2 : 0.0553060810276148"
## [1] "Granger causality test entre Sinistre3 et Police3 : 0.688841817622712"
## [1] "Granger causality test entre Sinistre3 et durPolice1 : 0.820036164981182"
## [1] "Granger causality test entre Sinistre3 et Durée : 0.0979144699157492"
## [1] "Granger causality test entre Sinistre3 et NSin : 0.0590920136482579"
## [1] "Granger causality test entre Sinistre3 et censure : 0.420641926037821"
## [1] "Granger causality test entre Sinistre3 et Sinistre0 : 0.541651999430448"
## [1] "Granger causality test entre Police1 et Police2 : 0.667133611759315"
## [1] "Granger causality test entre Police1 et Police3 : 0.00354879403230157"
## [1] "Granger causality test entre Police1 et durPolice1 : 0.785179474093198"
## [1] "Granger causality test entre Police1 et Durée : 0.132813120000447"
## [1] "Granger causality test entre Police1 et NSin : 0.0203246545252461"
## [1] "Granger causality test entre Police1 et censure : 0.204917633529259"
## [1] "Granger causality test entre Police1 et Sinistre0 : 0.846171264937825"
## [1] "Granger causality test entre Police2 et Police3 : 0.310224237087878"
## [1] "Granger causality test entre Police2 et durPolice1 : 0.055655850228255"
## [1] "Granger causality test entre Police2 et Durée : 0.277199668621766"
## [1] "Granger causality test entre Police2 et NSin : 0.0239566356380064"
## [1] "Granger causality test entre Police2 et censure : 0.0692500258718777"
## [1] "Granger causality test entre Police2 et Sinistre0 : 0.269125934485199"
## [1] "Granger causality test entre Police3 et durPolice1 : 0.653405172596898"

```

```
## [1] "Granger causality test entre Police3 et Durée : 0.0113542062495392"
## [1] "Granger causality test entre Police3 et NSin : 0.000464286639831717"
## [1] "Granger causality test entre Police3 et censure : 3.1569081679773e-05"
## [1] "Granger causality test entre Police3 et Sinistre0 : 0.0436598567098487"
## [1] "Granger causality test entre durPolice1 et Durée : 0.961630268620227"
## [1] "Granger causality test entre durPolice1 et NSin : 0.939390629133645"
## [1] "Granger causality test entre durPolice1 et censure : 0.567426085872365"
## [1] "Granger causality test entre durPolice1 et Sinistre0 : 0.0396821681991153"
## [1] "Granger causality test entre Durée et NSin : 0.572377359207435"
## [1] "Granger causality test entre Durée et censure : 0.000658970825685853"
## [1] "Granger causality test entre Durée et Sinistre0 : 0.0983144988858764"
## [1] "Granger causality test entre NSin et censure : 0.774359173306188"
## [1] "Granger causality test entre NSin et Sinistre0 : 0.647017398450364"
## [1] "Granger causality test entre censure et Sinistre0 : 0.151342377417034"
```

Si on fixe  $\alpha = 0.05$ , alors il y a une causalité entre Sinistre0 et les variables suivantes : RUC/durPolice1.

```
# Régression linéaire multiple
```

```
model1 <- lm(Sinistre0 ~ ., data = data)
```

```
# Afficher le résumé du modèle
```

```
summary(model1)
```

```
##
## Call:
## lm(formula = Sinistre0 ~ ., data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.5058 -1.4187 -0.0242  1.4905  7.9714
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.699e+01  4.738e-01  35.861  <2e-16
## pcsArtisans, comm., chefs d'ent.    1.047e-01  2.607e-01   0.402  0.6880
## pcsAutres pers. sans activite prof.    2.012e-02  2.659e-01   0.076  0.9397
## pcsCadres et prof. intellectuelles sup.    6.722e-02  2.367e-01   0.284  0.7764
## pcsEmployes   -7.427e-02  2.230e-01  -0.333  0.7391
## pcsOuvriers   -1.361e-01  2.143e-01  -0.635  0.5256
## pcsProfessions intermediaires  -1.613e-01  2.225e-01  -0.725  0.4686
## pcsRetraites  -4.910e-02  2.483e-01  -0.198  0.8432
## RUC            2.156e-04  2.186e-05   9.863  <2e-16
## csModeste      6.599e-02  2.563e-01   0.257  0.7969
## csMoyenne Inf   1.385e-01  2.043e-01   0.678  0.4978
## csMoyenne Sup   1.701e-01  1.569e-01   1.084  0.2783
## revez          -3.476e-07  3.968e-07  -0.876  0.3811
## crevpp2eme quartile  -5.551e-02  1.398e-01  -0.397  0.6914
## crevpp3eme quartile  -1.802e-01  1.959e-01  -0.920  0.3578
## crevpp4eme quartile  -2.094e-01  2.443e-01  -0.857  0.3913
## region          2.055e-02  1.282e-02   1.603  0.1090
## habi            1.353e-02  6.432e-02   0.210  0.8335
## AhabiParis + Agglomeration    1.904e-02  5.302e-01   0.036  0.9714
## AhabiUn. urb. de 10 000 a 99 999 hab.  -1.920e-01  2.782e-01  -0.690  0.4902
## AhabiUn. urb. de 100 000 hab. et +    -7.660e-02  4.377e-01  -0.175  0.8611
## AhabiUn. urb. de 2 000 a 9 999 hab.    3.983e-02  1.378e-01   0.289  0.7726
```

## AtyphNon declare	1.030e-01	2.578e-01	0.400	0.6895
## AtyphProprietaire	-2.187e-02	6.852e-02	-0.319	0.7496
## agecat41-50	-1.720e-04	9.263e-02	-0.002	0.9985
## agecat51-60	-7.366e-02	1.124e-01	-0.655	0.5123
## agecat61-96	-9.789e-02	1.666e-01	-0.588	0.5568
## AcompmCouple avec enfant(s)	-8.072e+00	1.256e-01	-64.270	<2e-16
## AcompmCouple sans enfant	1.061e-01	1.127e-01	0.941	0.3466
## AcompmPersonne seule	1.004e-01	1.627e-01	0.617	0.5372
## nbpers	1.779e-02	6.331e-02	0.281	0.7788
## enfantsPas d'enfants	NA	NA	NA	NA
## AnatMenage francais	-2.990e-01	2.048e-01	-1.460	0.1442
## AnatNon declare	-4.345e-01	2.352e-01	-1.847	0.0648
## BautoPas de vehicule	-2.069e-02	1.197e-01	-0.173	0.8627
## Nbadulte	-1.869e-02	6.861e-02	-0.272	0.7853
## Sinistre1	-5.238e-03	3.985e-03	-1.314	0.1888
## Sinistre2	2.597e-02	2.642e-02	0.983	0.3256
## Sinistre3	5.285e-03	1.170e-02	0.452	0.6515
## Police1	8.950e-03	6.040e-03	1.482	0.1384
## Police2	2.780e-03	2.726e-03	1.020	0.3079
## Police3	5.604e-03	1.304e-02	0.430	0.6674
## durPolice1	1.083e-12	9.259e-13	1.169	0.2423
## Durée	3.004e-05	6.295e-05	0.477	0.6332
## NSin	1.065e-02	9.939e-03	1.072	0.2840
## censure	8.854e-02	1.204e-01	0.735	0.4621
##				
## (Intercept)	***			
## pcsArtisans, comm., chefs d'ent.				
## pcsAutres pers. sans activite prof.				
## pcsCadres et prof. intellectuelles sup.				
## pcsEmployes				
## pcsOuvriers				
## pcsProfessions intermediaires				
## pcsRetraites				
## RUC	***			
## csModeste				
## csMoyenne Inf				
## csMoyenne Sup				
## reve				
## crevpp2eme quartile				
## crevpp3eme quartile				
## crevpp4eme quartile				
## region				
## habi				
## AhabiParis + Agglomeration				
## AhabiUn. urb. de 10 000 a 99 999 hab.				
## AhabiUn. urb. de 100 000 hab. et +				
## AhabiUn. urb. de 2 000 a 9 999 hab.				
## AtyphNon declare				
## AtyphProprietaire				
## agecat41-50				
## agecat51-60				
## agecat61-96				
## AcompmCouple avec enfant(s)	***			
## AcompmCouple sans enfant				

```

## AcompmPersonne seule
## nbpers
## enfantsPas d'enfants
## AnatMenage francais
## AnatNon declare
## BautoPas de vehicule
## Nbadulte
## Sinistre1
## Sinistre2
## Sinistre3
## Police1
## Police2
## Police3
## durPolice1
## Durée
## NSin
## censure
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.148 on 5307 degrees of freedom
## Multiple R-squared:  0.7519, Adjusted R-squared:  0.7499
## F-statistic: 365.6 on 44 and 5307 DF,  p-value: < 2.2e-16
summary(selectionAIC)

##
## Call:
## lm(formula = Sinistre0 ~ RUC + Acompm + Police1 + NSin, data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.4423 -1.4309 -0.0272  1.4871  7.8258
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.682e+01  8.389e-02  200.442  <2e-16 ***
## RUC             2.081e-04  8.701e-06   23.923  <2e-16 ***
## AcompmCouple avec enfant(s) -8.009e+00  7.598e-02 -105.410  <2e-16 ***
## AcompmCouple sans enfant    3.721e-02  8.351e-02    0.446  0.6560
## AcompmPersonne seule    -1.015e-02  1.025e-01   -0.099  0.9211
## Police1         1.050e-02  5.889e-03    1.782  0.0748 .
## NSin            1.602e-02  8.531e-03    1.878  0.0604 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.146 on 5345 degrees of freedom
## Multiple R-squared:  0.7507, Adjusted R-squared:  0.7504
## F-statistic: 2682 on 6 and 5345 DF,  p-value: < 2.2e-16

```

## IV regressions

The four kinds of variables in IV

- Y = outcome variables



- $X$  = endogenous, causal variable(s)
- $Z$  = instrument(s): doivent être exogènes, c'est-à-dire leur influence sur  $Y$  se fait seulement via leur influence sur  $X$ , la var endogène
- $W$  = any exogenous variables not including instruments