
Projet sur R de statistiques

Précarité chez les jeunes et vieillesse de la population en Europe en 2015



Table des Matières

Introduction du jeu de données	3
Contexte	3
Méthodologie appliquée	4
Variables utilisées	4
Définition des termes	4
Variables contextuelles	4
Variables sociales	4
Variables économiques	5
Analyse descriptive de la précarité des jeunes en Europe entre les pays (2015)	5
I/ Analyse des indicateurs	5
a/ Comparaison du taux de seniors et du taux de jeunes	5
b/ Comparaison des indicateurs sociaux	8
c/ Comparaison des indicateurs économiques	9
II/ Description bivariées des variables	10
III/ Analyse en composantes principales (ACP)	10
Conclusion	17

Introduction du jeu de données

a. Contexte

Les générations successives entrent aujourd'hui sur le marché du travail avec un niveau d'études et de diplôme toujours plus élevé. Les jeunes les plus qualifiés restent les plus à même de s'insérer sur le marché de l'emploi, même si la crise de 2008 a pu avoir pour effet de retarder leur accès à un emploi durable. Par contraste, la situation relative des moins qualifiés est toujours plus dégradée, la situation économique ne faisant qu'amplifier ce phénomène. Un tiers de l'ensemble des jeunes qui entrent sur le marché du travail aujourd'hui une trajectoire d'insertion incertaine, marquée par la précarité, une discontinuité de l'emploi, une mauvaise intégration sociale voire même un faible accès aux infrastructures de santé. Nous avons voulu montrer que cette tendance n'est pas une fatalité et que ce phénomène est lié à certaines données. Pour cela, nous tentons de mettre en lumière ce que l'on appelle communément le "fossé des générations" afin de montrer les liens entre la précarité chez les jeunes et le vieillissement d'une population. Il s'agit quelque part de répondre à la question : est-ce qu'une population plus vieillissante est une population qui prend moins soin de ses jeunes ?

D'une certaine manière, nous nous proposons d'étudier le dilemme "effet d'âge" ou "effet de génération" concernant les difficultés chez les jeunes. En d'autres termes, considérant toutes les autres variables égales par ailleurs, est ce que la précarité chez les jeunes dépend de la conjoncture actuelle ou dépend-t-elle de la structure socio-économique sur les temps longs ? Si cela dépend de la conjoncture, c'est un effet de génération, c'est à dire qu'on peut attribuer cette précarité qu'à une seule (ou à un groupe de) cohorte de jeunes aujourd'hui. Si cela est toujours le cas à travers les années, alors on peut conclure que la précarité chez les jeunes est un fait systémique, c'est donc un effet d'âge qui se répète de génération en génération.

En travaillant sur les liens entre le vieillissement de la population et la précarité socio-économique des jeunes, nous voulions revisiter le "fossé des générations" : un phénomène souvent étudié en science sociales. Le mouvement de mai 68 en est un des exemples les plus marquants, autant idéologique que socio-économique. La précarité et la marginalité des jeunes, que nous définissons avec l'ONU et l'INSEE comme les personnes âgées de 15 à 24 ans, est une question importante que nous abordons par l'analyse de quatre variables sociologiques et quatre variables économique présentées ci-dessous. Cela nous touche particulièrement en tant que jeune. Nous souhaitons étudié les cohortes de 2011, 2015 et 2019 pour les pays de l'Union Européenne (UE), par manque de place, nous nous focaliserons uniquement sur 2015 (année choisie pour sa représentativité et l'exhaustivité de ses données).

b. Méthodologie appliquée

Pour commencer, nous nous sommes rendus sur le site web <https://ec.europa.eu/eurostat/web/main/home> sur lequel nous avons trouvé notre jeu de données. Le site nous propose plusieurs jeux de données portant sur la jeunesse ainsi que sur plusieurs paramètres tels que la démographie, l'éducation, l'emploi, la santé, l'inclusion sociale, la culture, la participation, le volontariat, et le monde digital. Dans notre étude, nous nous sommes intéressés à des critères sociaux, économiques, mais aussi démographiques. Nous avons donc récupéré plusieurs tableaux de données selon plusieurs variables, et nous avons notamment dû assembler plusieurs tableaux afin de réaliser notre étude sur des jeunes de 15 à 24 ans car les données présentes sur le site étaient subdivisées en sous catégories d'âge : de 15 à 19 ans et de 20 à 24 ans. Afin de rassembler ces données, nous avons eu recours au logiciel tableur Google Sheets. Afin de comparer nos données sur les jeunes avec le vieillissement de la population, nous avons également récupéré un jeu de données sur le même site web qui regroupe la population par classe d'âge. A l'aide de ce nouveau tableau, nous allons pouvoir observer si les variables socio-économiques des jeunes sont influencés par le nombre de personnes âgées en Europe.

c. Variables utilisées

Définition des termes

- Jeune : personnes âgées de 15 à 24 ans (définition INSEE et ONU).
- Seuil de pauvreté : correspond à 60 % du niveau de vie médian de la population.
- Difficulté financière : impossibilité de se payer au moins trois des dépenses ci-contre : loyer, chauffage, faire face à des dépenses inattendues, manger de la viande ou des protéines régulièrement, télévision, voiture, téléphone

Variables contextuelles

- **TJP** : Taux de personnes jeunes dans la population totale
- **TSP** : Taux de seniors dans la population totale

Variables sociales

- **T_SUICIDE** : Taux de suicides chez les jeunes par rapport au nombre de suicides totaux dans la population
- **ACC_SOINS** : Taux de jeunes qui ne peuvent pas accéder aux soins médicaux (soit trop cher soit trop loin)
- **ECH_SCO** : Taux de jeunes en échec scolaire (ayant quitté le système éducatif très tôt)
- **ENS_SUP** : Taux de jeunes qui ont atteint l'enseignement supérieur

Variables économiques

- **T_TEMP** : Taux de jeunes qui ont des contrats temporaires
- **SEUIL_PAU** : Taux de jeunes en dessous du seuil de pauvreté
- **DIF_FIN** : Taux de jeunes en difficulté financière
- **T_CHOM** : Taux de chômage chez les jeunes

Analyse descriptive de la précarité des jeunes en Europe entre les pays (2015)

I/ Analyse des indicateurs

a/ Comparaison du taux de séniors et du taux de jeunes

Afin de commencer notre étude, nous nous proposons de faire un peu d'analyse descriptive sur l'année 2015.

```
#-----  
# Chargement du jeu de données  
#-----  
Data2015<-read.csv(file = "Data2015(2).csv",header = TRUE,row.names = 1, dec=",")  
#-----  
# Analyse des indicateurs  
#-----  
#-----  
# Comparaison du taux de séniors et du taux de jeunes  
#-----  
boxplot(Data2015$TJP,main = "Pourcentages de jeunes dans les populations des pays de l'UE",  
        xlab = "Part dans la population (en %)",ylab = "Jeunes de 15-24 ans",col = "orange",  
        border = "brown",horizontal = TRUE,notch = TRUE)  
boxplot(Data2015$TJP, Data2015$TSP,  
        main = "Comparaison de la part de jeunes et de la part de séniors dans les populations de l'UE",  
        at = c(1,2),  
        names = c("Taux de jeunes en %", "Taux de séniors en %"),  
        las = 2,  
        col = c("orange","red"),  
        border = "brown",  
        horizontal = FALSE)  
#-----  
# Comparaison des indicateurs sociaux  
#-----  
t_suicide <- Data2015$T_SUICIDE  
t_acc_soins <- Data2015$ACC_SOINS  
t_ech_sco <- Data2015$ECH_SCO  
t_ens_sup <- Data2015$ENS_SUP  
norm_suicide<-rnorm(200,mean=mean(t_suicide, na.rm=TRUE), sd=sd(t_suicide, na.rm=TRUE))  
norm_soins<-rnorm(200,mean=mean(t_acc_soins, na.rm=TRUE), sd=sd(t_acc_soins, na.rm=TRUE))  
norm_scol<-rnorm(200,mean=mean(t_ech_sco, na.rm=TRUE), sd=sd(t_ech_sco, na.rm=TRUE))  
norm_ens<-rnorm(200,mean=mean(t_ens_sup, na.rm=TRUE), sd=sd(t_ens_sup, na.rm=TRUE))  
boxplot(t_suicide, norm_soins, norm_scol, norm_ens,  
        main = "Comparaison des indicateurs d'inégalités sociales pour les jeunes dans les populations de l'UE",  
        at = c(1,2,4,5),  
        names = c("Taux de morts par suicide chez les jeunes en %", "Taux de jeunes en manque d'accès aux soins en %",  
                  "Taux de jeunes en décrochage scolaire en %","Taux de jeunes atteignant l'enseignement secondaire en %"),  
        las = 2,  
        col = c("red","red","orange","orange"),  
        border = "brown",  
        horizontal = FALSE)  
#
```

```

#-----
# Comparaison des indicateurs économiques
#-----
t_temp <- Data2015$T_TEMP
seuil_pau <- Data2015$SEUIL_PAU
dif_fin <- Data2015$DIF_FIN
t_chom <- Data2015$T_CHOM
norm_temp <- rnorm(200, mean=mean(t_temp, na.rm=TRUE), sd=sd(t_temp, na.rm=TRUE))
norm_pau <- rnorm(200, mean=mean(seuil_pau, na.rm=TRUE), sd=sd(seuil_pau, na.rm=TRUE))
norm_dif <- rnorm(200, mean=mean(dif_fin, na.rm=TRUE), sd=sd(dif_fin, na.rm=TRUE))
norm_chom <- rnorm(200, mean=mean(t_chom, na.rm=TRUE), sd=sd(t_chom, na.rm=TRUE))
boxplot(norm_temp, norm_pau, norm_dif, norm_chom,
        main = "Comparaison des indicateurs d'inégalités économiques pour les jeunes dans les populations de l'UE",
        at = c(1,2,3,4),
        names = c("Taux de jeunes qui ont des contrats temporaires en %", "Taux de jeunes en dessous du seuil de pauvreté en %",
                  "Taux de jeunes en difficulté financière en %", "Taux de chômage chez les jeunes en %"),
        las = 2,
        col = c("red", "blue", "orange", "green"),
        border = "brown",
        horizontal = FALSE)
#-----
# Corrélations entre les variables
#-----
matrice_correlations<-round(cor(Data2015[2:34,1:9]),digits=2)

plot(Data2015$TJP, Data2015$TSP, xlab = "Taux de jeunes dans la population", ylab = "Taux de seniors dans la population")
title("Corrélation entre le taux de jeunes et le taux de seniors dans les populations de l'UE")
cor(Data2015$TJP, Data2015$TSP)
abline(1, 0.004, col=2)

plot(Data2015$ENS_SUP, Data2015$ECH_SCO, xlab = "Taux de jeunes qui ont atteint l'enseignement supérieur", ylab = "Taux de jeune en ?chec scolaire")
title("Corrélation entre le taux de jeunes qui ont atteint l'enseignement supérieur et le taux de jeunes en ?chec scolaire")
cor(Data2015$ENS_SUP, Data2015$ECH_SCO)

plot(Data2015$ECH_SCO, Data2015$T_TEMP, xlab = "Taux de jeune en ?chec scolaire", ylab = "Taux de jeunes qui ont des contrats temporaires")
title("Corrélation entre le taux de jeune en ?chec scolaire et le taux de jeunes qui ont des contrats temporaires")
cor(Data2015$ECH_SCO, Data2015$T_TEMP)

```

figure 1

Tout d'abord, traçons le boxplot du pourcentages de jeunes dans les populations des pays de l'UE.

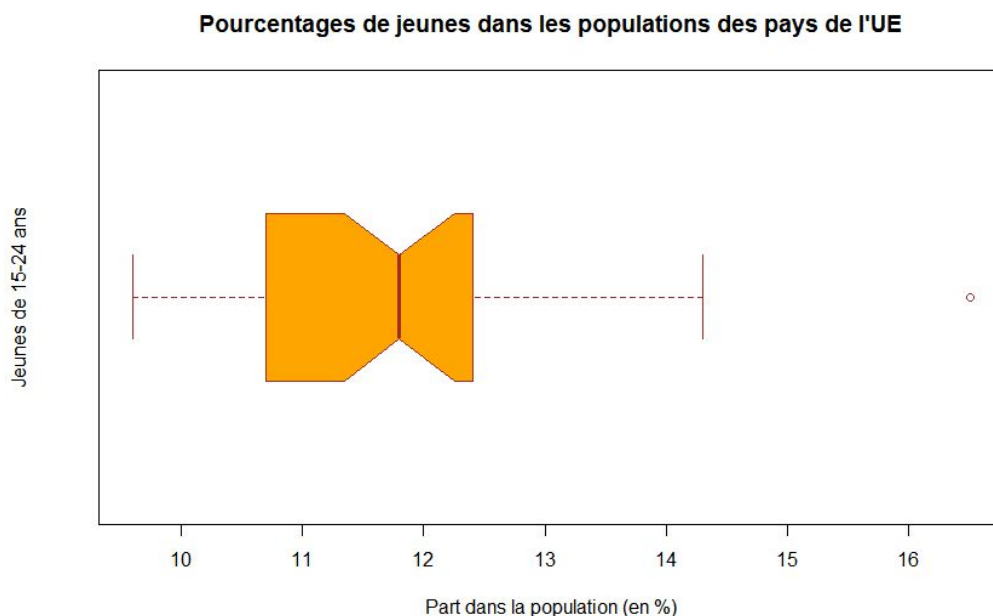


figure 2

On observe sur ce boxplot que les jeunes représentent un peu moins de 12% de la population européenne selon la médiane. Il n'y a cependant pas une grande disparité du pourcentage de jeune en fonction du pays européen car le minimum et la maximum sont ici assez proches. L'Union Européenne est assez homogène sur ce point, ce qui s'explique par l'homogénéité démographique et économique des pays membres.

Traçons le Boxplot de la comparaison de la part de jeunes et de la part de seniors dans les populations de l'UE.

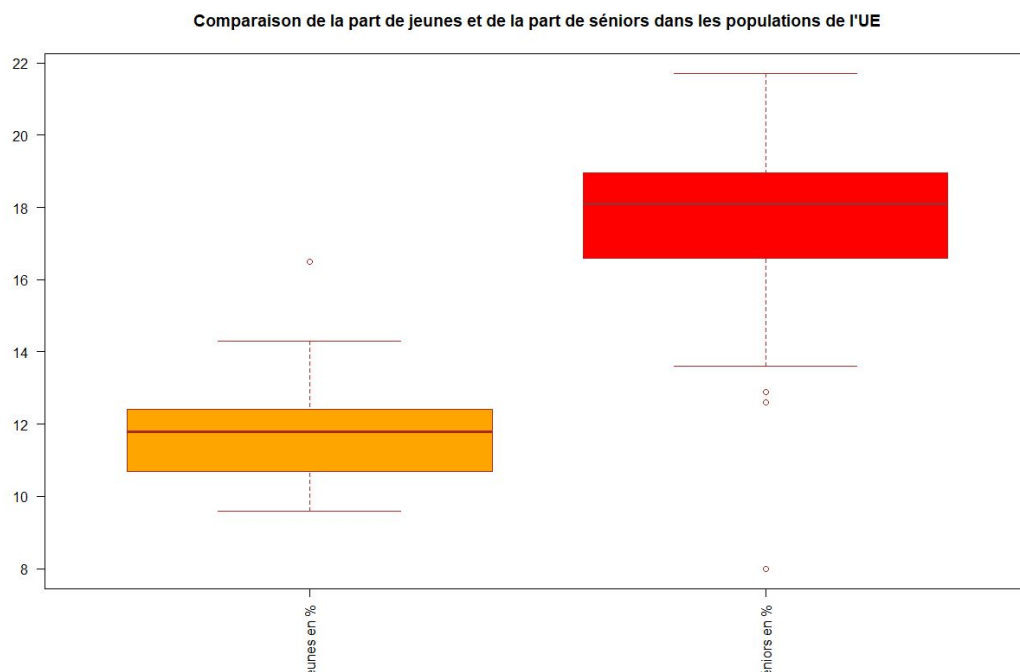


figure 3

Nous pouvons constater que la part de seniors dans l'UE est, en médiane, à peu près 18%. La population européenne est donc majoritairement plus vieille que jeune. De plus la dispersion de la part de senior est elle aussi assez faible.

Bien évidemment, une population vieille est une population qui compte moins de jeunes, les deux variables étant corrélées à $-0,77$, cela signifie que plus une population européenne est jeune, moins il y a un taux élevés de seniors comme cela est visible figure 4.

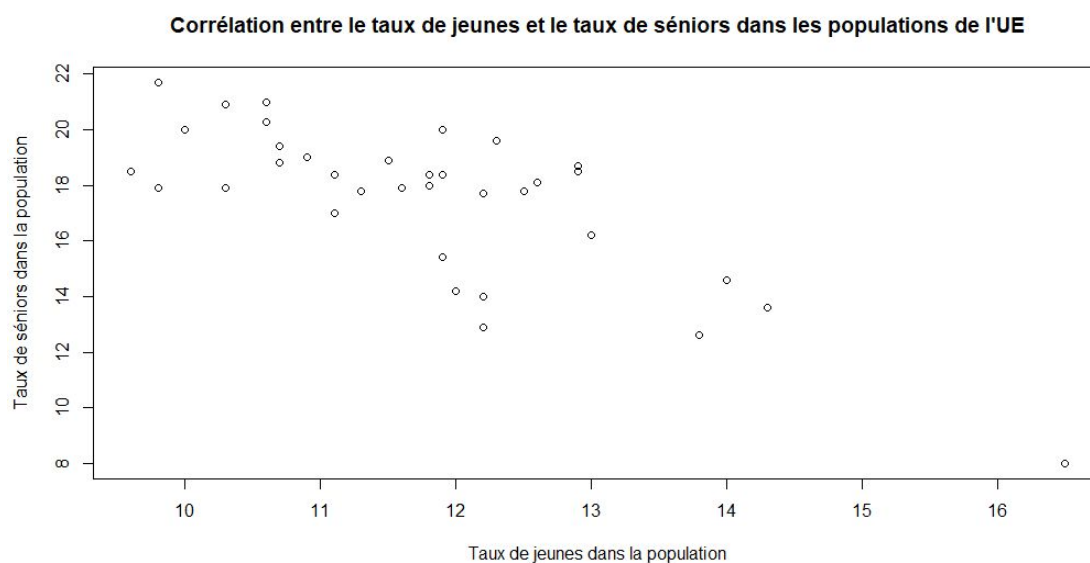


figure 4

Pour revenir à la figure 2, la dispersion des seniors étant faible, il faudra faire attention aux variables avec une forte dispersion. En effet, une grande dispersion signifierait alors une corrélation avec des variables cachées (non exprimées ici).

b/ Comparaison des indicateurs sociaux

Observons les indicateurs d'inégalités sociales. Les quatre variables observées sont, dans l'ordre, le taux de suicide, le taux d'accès aux soins, le taux de jeunes en échec scolaire et le taux de jeunes en enseignement supérieur.

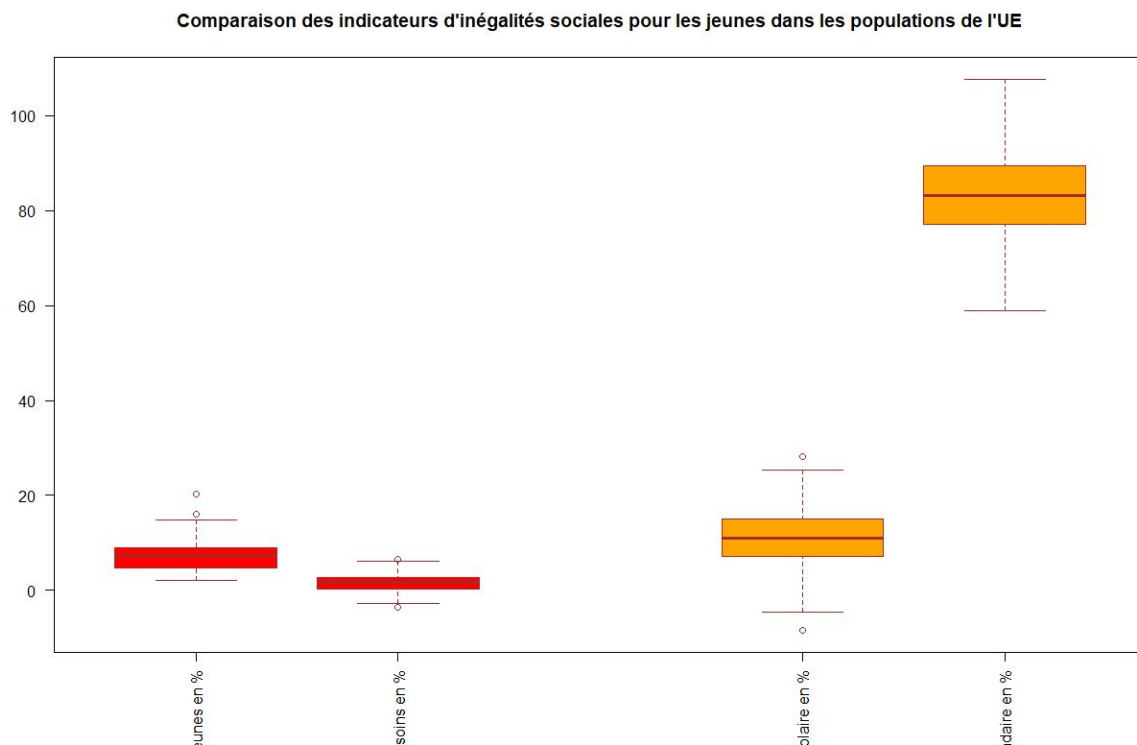


figure 5

On peut constater que les variables du taux de suicide et du taux d'accès aux soins chez les jeunes sont assez homogènes en UE en 2015. Néanmoins, le taux d'échec scolaire et le taux de jeunes dans l'enseignement supérieur est assez dispersé. Ce qui montre les inégalités entre les pays de l'UE au niveau de l'éducation, bien que le taux de senior soit relativement homogène entre ces pays. Cela signifie qu'au niveau de l'éducation au moins, on peut penser qu'il n'y a pas de lien de cause à effet entre le taux de senior d'une population et le niveau de l'éducation en Europe. Cela reste à voir pour les indicateurs de l'intégration sociale que sont le taux de suicide et le taux d'accès aux soins.

c/ Comparaison des indicateurs économiques

Observons maintenant les indicateurs des inégalités économiques, avec, dans l'ordre, le taux de jeunes en contrat temporaire, le taux de jeunes sous le seuil de pauvreté, le taux de jeunes en difficultés financières et le taux de jeunes au chômage.

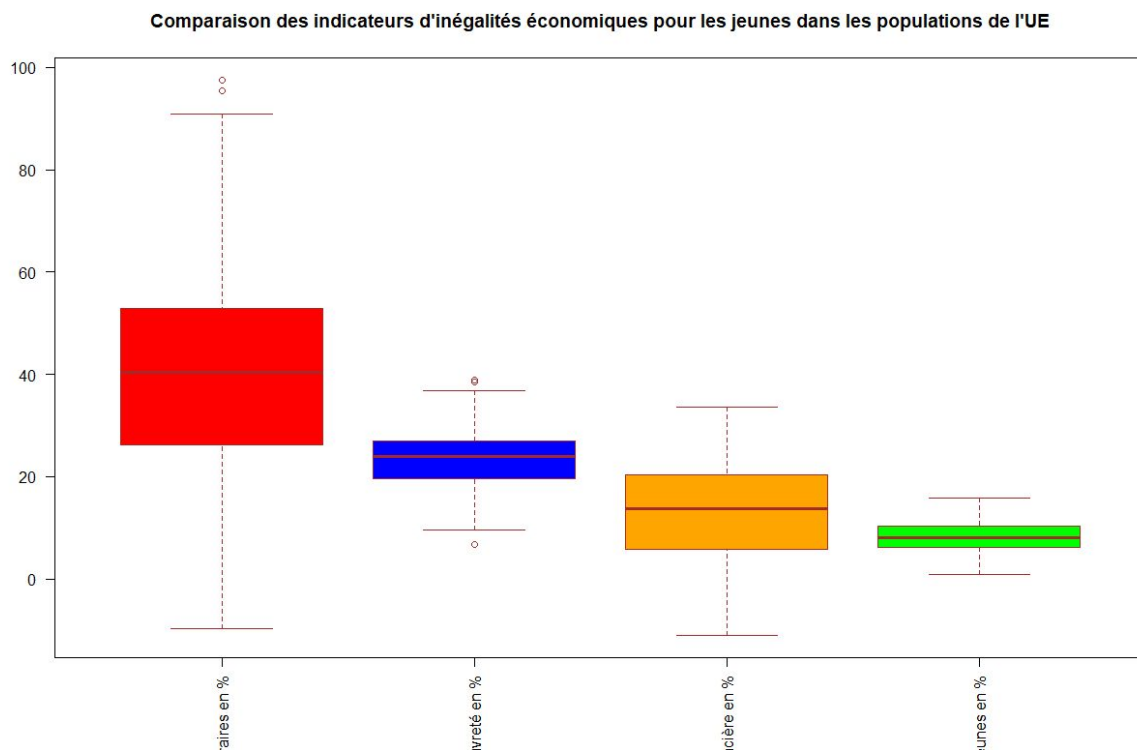


figure 6

On peut constater une très forte disparité entre les pays de l'UE au niveau du taux de jeunes en contrat temporaire. Cela peut s'expliquer par l'hétérogénéité des politiques économique, certains pays privilégiant le CDD quand d'autres privilégient le CDI. On constate que le taux de jeunes sous le seuil de pauvreté est assez élevé, et que la dispersion n'est pas très grande relativement au taux de jeunes en contrat temporaire et au taux de jeunes en difficultés financières. Finalement, le taux de jeunes au chômage est assez homogène en Europe.

II/ Description bivariées des variables

En utilisant le package Hmisc et le package corrplot, nous avons calculé la matrice des coefficients de corrélation.

```
> mcor<-cor(Data2015)
> corrplot(mcor, type="upper", tl.col="black", tl.srt=45)
```

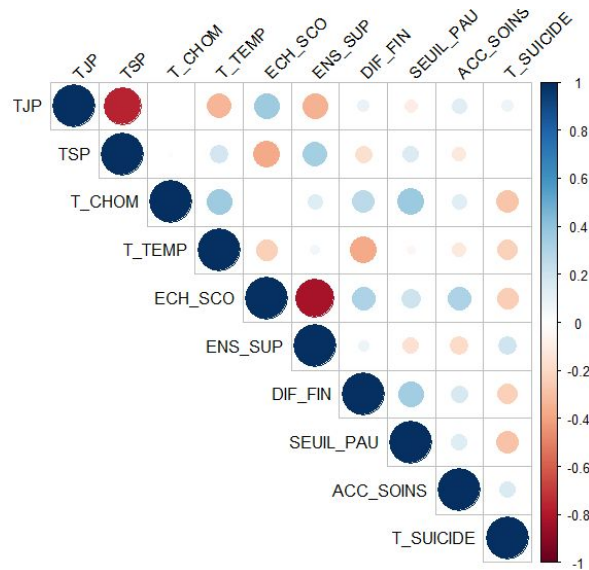


figure 7 : Représentation graphique de la matrice des coefficients de corrélation

Nous pouvons en conclure que globalement, peu de valeurs sont vraiment corrélées significativement entre elles. Tout d'abord il y a le taux de jeunes et le taux de seniors, que nous avons vu plus haut, qui est bien corrélé négativement. Ensuite nous pouvons voir qu'il y a le taux d'échec scolaire qui est bien corrélé négativement avec le taux d'enseignement supérieur. C'est logique, plus il y a de l'échec scolaire, moins est la propension des jeunes à pouvoir accéder à l'enseignement supérieur. Enfin, au vue de la variable TSP qui nous intéresse, il semble qu'il y ait une corrélation négative avec l'échec scolaire et une corrélation positive avec le taux d'enseignement supérieur : plus le taux de senior est élevé dans une population, moins il y aurait d'échec scolaire chez les jeunes et plus les jeunes accèdent à des hautes études. Cette corrélation n'est pas très élevée. Nous pouvons tout de même supposer qu'une population plus vieille est synonyme de richesse et d'un système éducatif plus ancien et plus ancré. Néanmoins, rappelons nous qu'il peut y avoir ici des variables cachées.

III/ Analyse en composantes principales (ACP)

Afin d'analyser notre jeu de données, nous avons eu recours à l'utilisation de la méthode d'ACP. Nous allons tenter de trouver d'autres corrélations entre les indicateurs économiques et sociaux chez les jeunes et le vieillissement de la population en construisant de nouvelles variables synthétiques obtenues en combinant nos variables initiales à l'aide de facteurs. Afin de réaliser cette ACP, nous avons saisi dans RStudio les lignes de codes se trouvant sur la figure 8.

figure 8

```
# Chargement du package PCAMixdata
#-----
require(PCAMixdata)
#-----
# Mise en oeuvre de l'ACP
#-----
res<-PCAMix(Data2015[1:34,2:9]) # tous les calculs de l'ACP sont stockés dans l'objet "res"
# NB : par défaut les graphiques des plans factoriels 1-2
# sont affichés à l'écran
res <- PCAMix(Data2015[1:34,2:9],graph=FALSE) # idem sans les graphiques
res # permet de voir l'ensemble des sorties numériques disponibles
#-----
# Choix du nombre d'axes à retenir
#-----
round(res$eig,digit=2) # permet d'afficher les valeurs propres et les pourcentages
# de variances expliquées par chaque axe
# Graphique de l'ébouli des valeurs propres
barplot(res$eig[,1],main="Eigenvalues",names.arg=1:nrow(res$eig))
abline(h=1,col=2,lwd=2)
round(res$quanti$cos2,digit=2)
plot(res,axes=c(3,4),choice = "cor")
plot(res,axes=c(3,4),choice="ind")
plot(res,axes=c(3,4),choice="sqload")
#-----
# Sorties numériques pour les individus et les variables
#-----
res$ind
round(res$ind$cos2,digit=3)
res$quanti
round(res$quanti$cos2,digit=3)
```

Pour commencer, nous avons déterminé les dimensions que nous allons garder pour l'ACP. On peut voir sur la figure ci-dessous que les quatre premières dimensions sont celles portant la quantité maximum de variations. Grâce au critère de Kaiser, on peut conserver les quatre premiers axes. Ces quatre axes rassemblent 77% de l'information sur nos données.

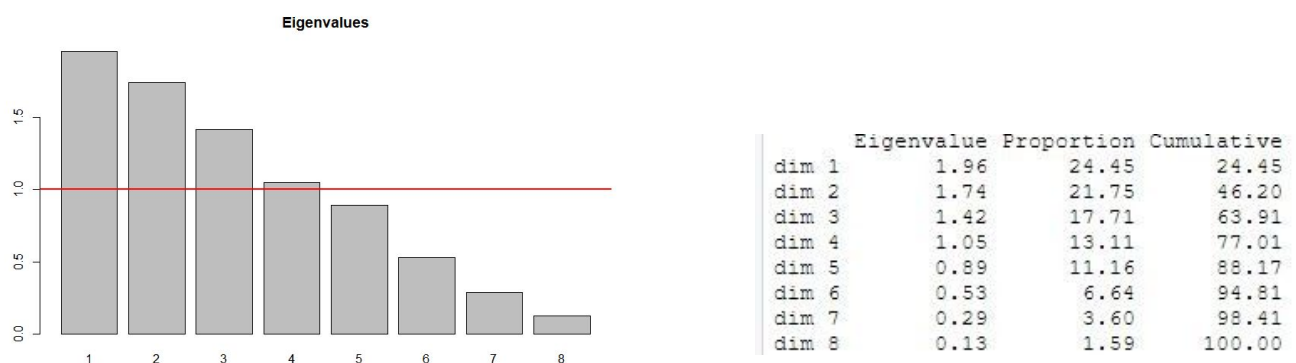


figure 9

Néanmoins, il pourrait être intéressant de regarder la dimension 5 car, on voit sur la figure 10 que 71% de l'information sur l'indicateur d'accès aux soins se trouvent sur cette dimension.

```
> round(res$quant1$cos2, digit=2)
      dim 1 dim 2 dim 3 dim 4 dim 5
TSP      0.03  0.01  0.08  0.75  0.07
T_CHOM   0.07  0.41  0.23  0.14  0.02
T_TEMP   0.06  0.00  0.81  0.01  0.00
ECH_SCO   0.70  0.16  0.00  0.00  0.01
ENS_SUP   0.36  0.49  0.06  0.01  0.02
DIF_FIN   0.14  0.46  0.21  0.00  0.02
SEUIL_PAU 0.41  0.19  0.03  0.04  0.04
ACC_SOINS 0.18  0.01  0.00  0.09  0.71
```

figure 10

On observe également sur la figure 10 que 75% de l'information sur le taux de seniors dans la population se trouve sur la dimension 4. Or, c'est la valeur à laquelle on veut comparer nos indicateurs afin de pouvoir établir, ou non, une corrélation. Il faudra donc observer les différentes dimensions en fonction de la dimension 4. Sur la dimension 1, nous retrouverons les informations sur le décrochage scolaire et le taux de jeunes sous le seuil de pauvreté. Sur la dimension 2, on va pouvoir observer les difficultés financières, l'accès à l'enseignement supérieur et le taux de chômage chez les jeunes. Et enfin, on observe sur la dimension 3 le taux de jeunes salariés avec un contrat temporaire.

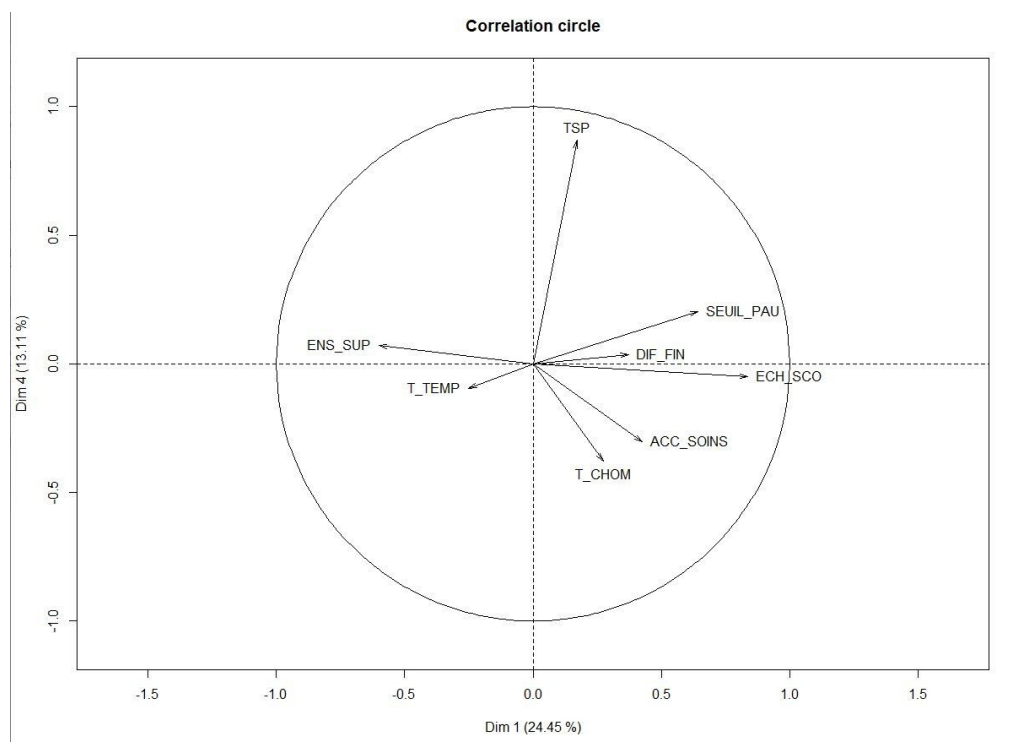


figure 11

On peut voir sur le cercle de corrélation des axes 1 et 4 (figure 11) que le taux de jeunes sous le seuil de pauvreté est légèrement corrélé avec le taux de décrochage scolaire et que celui-ci est corrélé négativement avec l'accès à l'enseignement supérieur. En effet, il est facile de concevoir que des jeunes avec une meilleure éducation auront un salaire plus élevé et donc plus de chances de vivre au-dessus du seuil de pauvreté. Néanmoins, nous ne voyons pas de corrélation avec le taux de seniors.

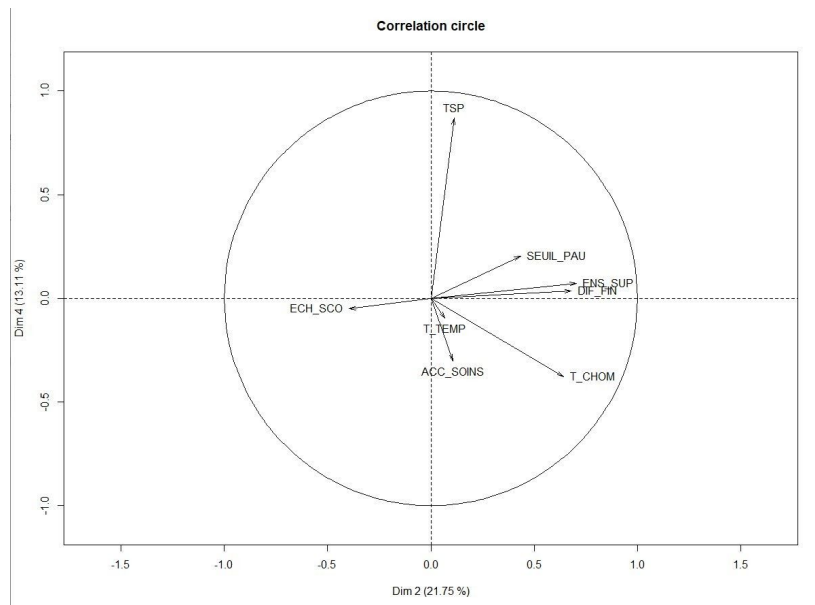


figure 12

En ce qui concerne la dimension 2 (figure 12), nous observons une forte corrélation entre les difficultés financières chez les jeunes et l'accès à l'enseignement supérieur. En effet, on peut en déduire qu'il est difficile d'avoir accès au confort financier lorsque l'on est étudiant car il est compliqué d'avoir un revenu. Là encore, on ne peut pas établir de corrélation entre le taux de seniors et les indicateurs. On peut cependant supposer que le taux de chômage est lié au taux de décrochage scolaire, malgré que ce dernier soit mal représenté sur l'axe 2.

Enfin, on voit sur la dimension 3 (figure 13) que le taux de chômage est lié au taux de contrats temporaires chez les jeunes. En effet, lorsque le marché du travail va mal, les entreprises ont tendance à favoriser les contrats temporaires. Il n'est donc pas anormal de constater cette corrélation.

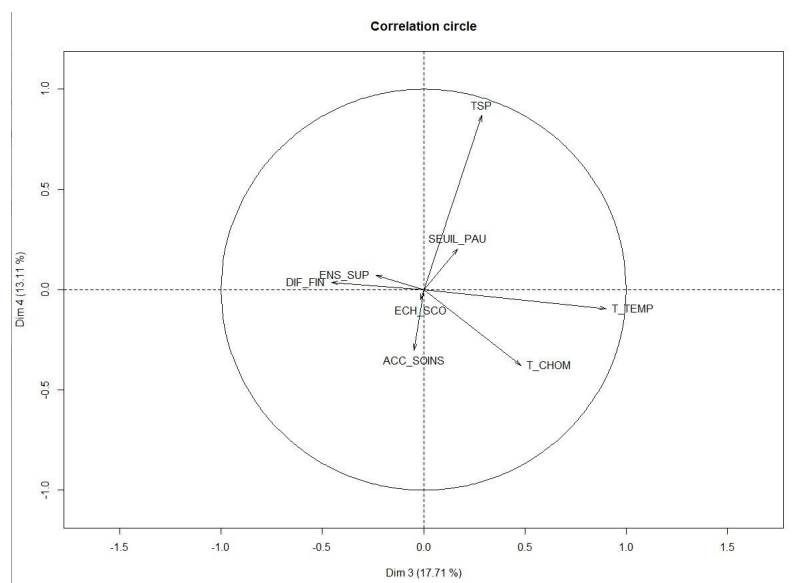


figure 13

Pour terminer, n'ayant pas trouvé de corrélations entre nos indicateurs et le taux de seniors dans les populations de l'Union Européenne, nous avons tenté d'identifier des pays qui pourraient nous éclairer sur la problématique. Nous avons donc utilisé la carte des individus.

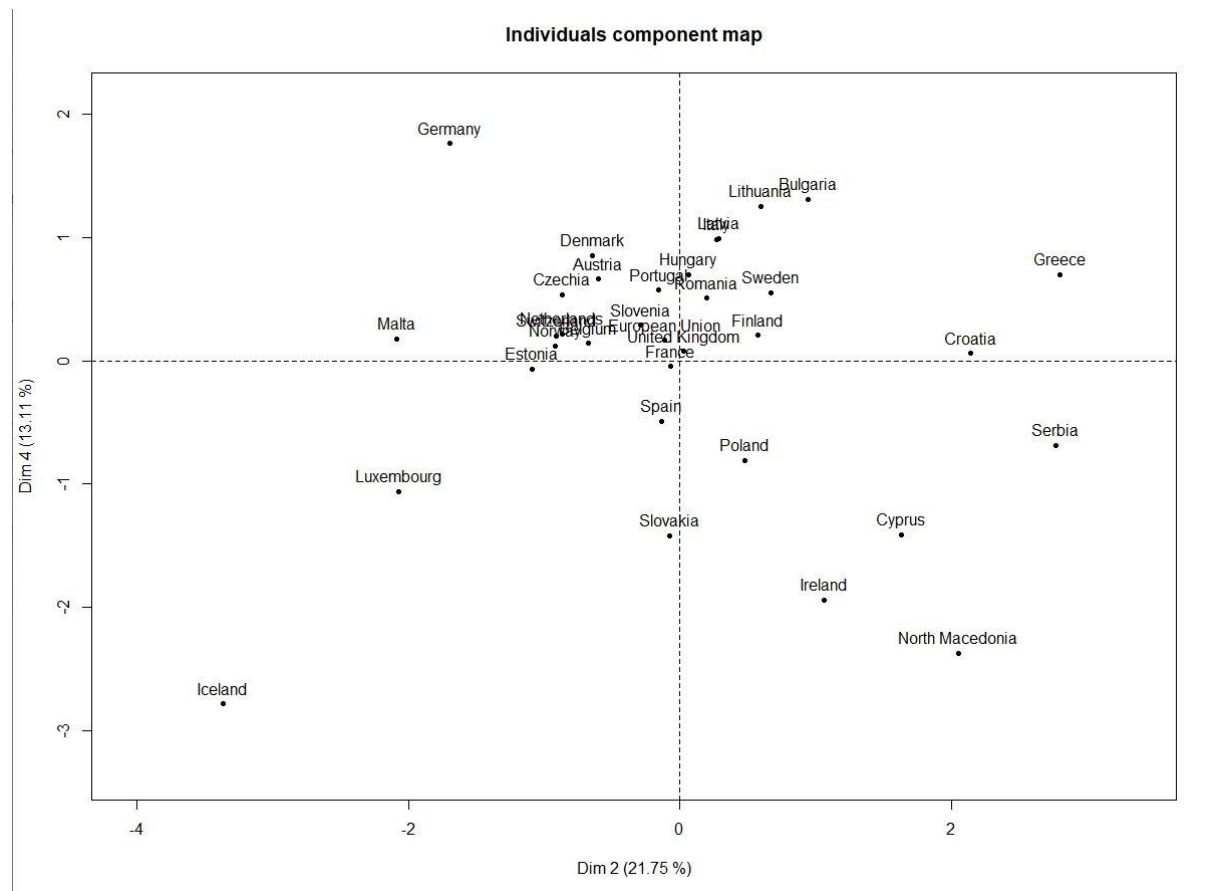


figure 14

Ce qui nous intéresse sur cette carte, ce sont les pays excentrés dans le coin supérieur droit car cette zone indique une forte proportion de seniors et d'inégalités chez les jeunes. Mais aussi le coin inférieur gauche qui indique une faible proportion de seniors et d'inégalités chez les jeunes. On retrouve alors seulement deux pays dans ces zones : la Grèce et l'Islande. On voit qu'on retrouve quelques pays dans le coin inférieur droit, ce qui indique qu'ils ont un taux de seniors faible mais beaucoup d'inégalités chez les jeunes. Enfin, beaucoup de pays se retrouvent au centre du graphe, ce qui nous montre qu'à partir de nos données, nous ne sommes pas en mesure de donner une réponse générale à la problématique posée au niveau de l'Union Européenne.

Conclusion

Pour conclure. Malgré quelques informations probantes, il nous a été impossible de démontrer un lien de corrélation entre le taux de seniors dans les populations européennes et une quelconque précarité chez les jeunes, qu'elle soit sociale ou économique. Il nous aurait fallu étudier plus de cohorte sur des temps beaucoup plus longs. De plus, une seule année d'étude ne nous a pas permis de faire varier assez la variable TSP pour en mesurer les conséquences. Nous pouvons relever quelques problèmes d'analyse à notre étude : il n'est pas raisonnable de penser qu'une variation du TSP toutes les autres variables étant égales par ailleurs nous permettrait de conclure parfaitement sur l'influence de cette dernière étant donné que beaucoup de variables ont été mises de côté. Cela laisse la place à des effets cachés.

Loin de vouloir appuyer le fossé des générations par notre travail, nous avons au contraire souhaité le mettre en lumière pour mieux le combattre. D'un côté les seniors semblent poussés à la marge de la société à mesure qu'ils vieillissent et quittent le monde de l'emploi. De l'autre, les jeunes semblent manquer de moyens et ont des difficultés à fonder leur avenir. Notre étude, finalement, montre qu'il n'existe aucun lien entre un TSP plus élevé que la moyenne européenne et une marginalisation des jeunes. Le lien entre TSP et système éducatif a été démenti par l'ACP. Nous pouvons donc répondre à notre question posée en introduction : non, une population vieillissante ne semble pas induire une précarité chez les jeunes en Europe. Il faut se tourner vers d'autres explications pour aborder cette question.