

PROJET_STAT

BODIANG Assane & AIDARA SIDI MOHAMED

23 mars 2018

Analysis of the french population since 1982/ Analyse de la population française depuis 1982

Introduction

English

Hi! This is a university project (University of Grenoble Alpes) in which we're handed the task of conducting an analysis on a data-set of our choice (Problematic/Questions/Methodology). The main goal is to have a quick glimpse at what R can do and how to use it's provided tools. Since we're students from France, we decided to work on the evolution of the country's population since 1982.

Français

Salut! Ceci est un projet d'université (Université Grenoble Alpes) dont le but est de découvrir le langage R et quelques uns de ses outils/fonctions de base. Il nous est demandé d'étudier un jeu de données au choix (Problématiques/Questions/Méthodes d'analyse entre autres). Etant donné que nous sommes étudiants en France, nous avons donc décidé d'étudier l'évolution de la population française depuis 1982.

Problématique posée: Vieillissement de la population française depuis 1982.

SOURCES: <https://www.insee.fr/fr/statistiques/1892117?sommaire=1912926>

Données:

Données 1: Composantes de la croissance démographique, France métropolitaine.

- **-p:** populations 2016, 2017, 2018, état civil 2017, soldes migratoires 2015, 2016, 2017 : résultats provisoires à fin 2017.
- **-nd:** non disponible.
- **-Champ:** France hors Mayotte jusqu'en 2014 et France incluant Mayotte à partir de 2014
- **-Source:** Insee, estimations de population et statistiques de l'état civil

Données 2: Évolution de la structure de la population, France métropolitaine.

- **-p:** résultats provisoires à fin 2017.
- **-Champ:** France métropolitaine
- **-Source:** Insee, estimations de population

Question:

Question I:

La diminution des naissances est-elle inversement proportionnelle au vieillissement de la population du fait de la faible variation du taux de mortalité d'années en années?

Question II:

La variation du ratio (mortalité/naissances) est-elle contingente de celle des tranches d'âges de la population française?

Méthodes:

Visualiser:

- La courbe d'évolution du nombre de naissances vivantes enregistrées par année (de 1982 à 2017)
- La courbe d'évolution du nombre de personnes du troisième âge par année de 1982 à 2017
- La courbe d'évolution du ratio Décès/Naissance de 1982 à 2017
- Graph bar comparatif entre le nombre de naissances vivantes et le nombre de décès
- Graph bar comparatif entre les tranches d'âges (Jeunes: moins de 20ans et Vieux: plus de 65 ans)

Ensemble des bibliothèques à charger (à installer au préalable)

Pour réaliser cette étude, nous avons utilisé les bibliothèques principalement vues en cours durant le semestre (`ggplot2`, `dplyr`, `xlsx`).

Chargement de dataframes pour nos jeux de données

En utilisant la fonction `read.xlsx()`, il est possible d'ouvrir des jeux de données au format EXCEL. Nous avons utilisé la bibliothèque `xlsx` car la communauté soutient le fait qu'elle soit l'une des plus matures et faciles à appréhender.

```
header=TRUE
df1<- read.xlsx("Composantes-de-la-croissance-démographique-France-métropolitaine.xls",1)
df2<- read.xlsx("Evolution-de-la-structure-de-la-population-France-métropolitaine.xls",1)
```

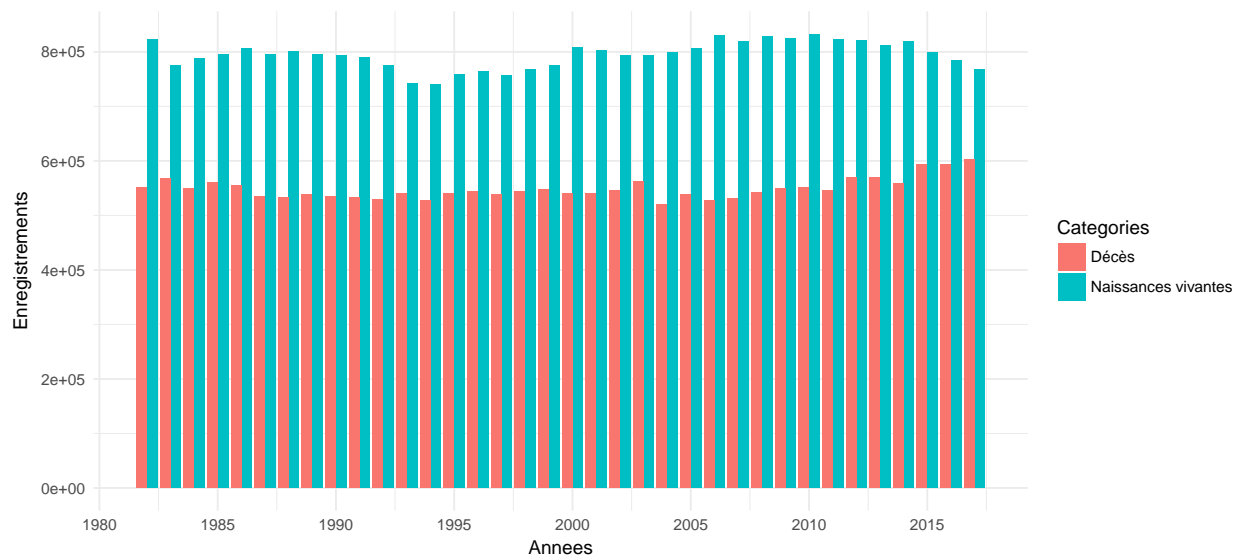
Nous avons aussi déclaré un certain nombre de variables qui vont nous aider à tracer l'ensemble de nos graphes.

```
naissances<- df1$Nombre.de.naissances.vivantes
deces<- df1$Nombre.de.deces
tranchesJeunes<-df2$de.0.a.19.ans.1
tranchesVieux<-df2$X65.ans.ou.plus.1
populationVieux<-df2$X65.ans.ou.plus
Ratio<-with(df1,(Nombre.de.deces/Nombre.de.naissances.vivantes)*100)
Annee<-df1$Annee
```

Graph bar comparatif entre le nombre de naissances vivantes et le nombre de décès

ce **graphe bar** a pour but de mettre en exergue la différence entre les enregistrements de naissances vivantes et de décès sur la période de 1982 à 2017. Les couleurs ont encore une fois fait l'objet d'un choix, choix duquel les naissances sont en bleu car augmentant la population et les décès en rouge car ayant un effet inverse sur cette dernière.

```
df3 <- data.frame(Categories=rep(c("Naissances vivantes", "Décès"), each=37),
Annees=rep(c(df1$Annee),2),
Enregistrements=c(naissances,deces))
ggplot(data=df3, aes(x=Annees, y=Enregistrements, fill=Categories)) +
geom_bar(stat="identity", position=position_dodge()) +
scale_x_continuous(breaks = round(seq(1980, 2020, by = 5),1)) +
theme_minimal()
```

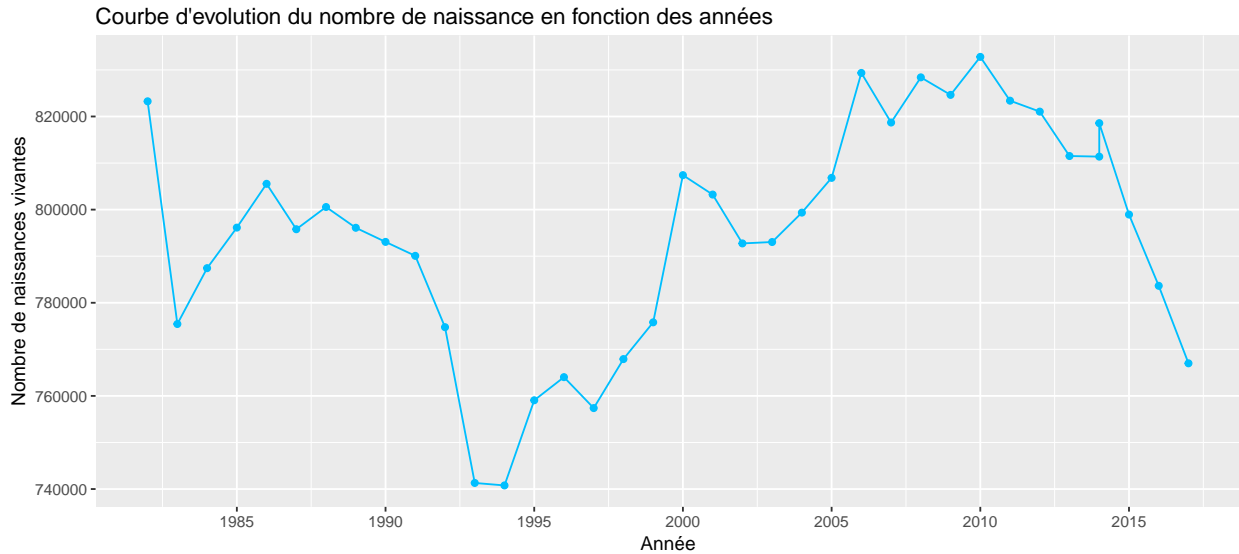


Il est serait légitime de dire que le nombre de naissances est beaucoup plus grand que le nombre de décès et cela sur toute la période de 1982 à 2017 (minimu_Naissances = 750K en 93-94 et max_Décès = 600K en 2017). Cela dit, il faut noter la quasi-constance du nombre de décès durant cette période qui tourne aux alentours de **548917.4** +/- 50K. Quant au nombre de naissances vivantes enregistrées, il est resté assez varié sur toute l'étendue de la tranche d'années étudiée.

Courbe d'évolution du nombre de naissances vivantes en fonction des années

Cette courbe montre l'évolution du nombre de naissances vivantes en fonction des années. Le choix de la modelisation en bleu vient du fait que ce tracé concerne le nombre de naissances vivantes (qui augmente donc la population totale).

```
ggplot(data=df1, aes(Annee,naissances, group =1)) +
scale_x_continuous(breaks = round(seq(1980, 2020, by = 5),1)) +
geom_point(col="deepskyblue") +
geom_line(col="deepskyblue") +
labs(x = "Année", y = "Nombre de naissances vivantes", title = "Courbe d'evolution du nombre de naissances vivantes")
```

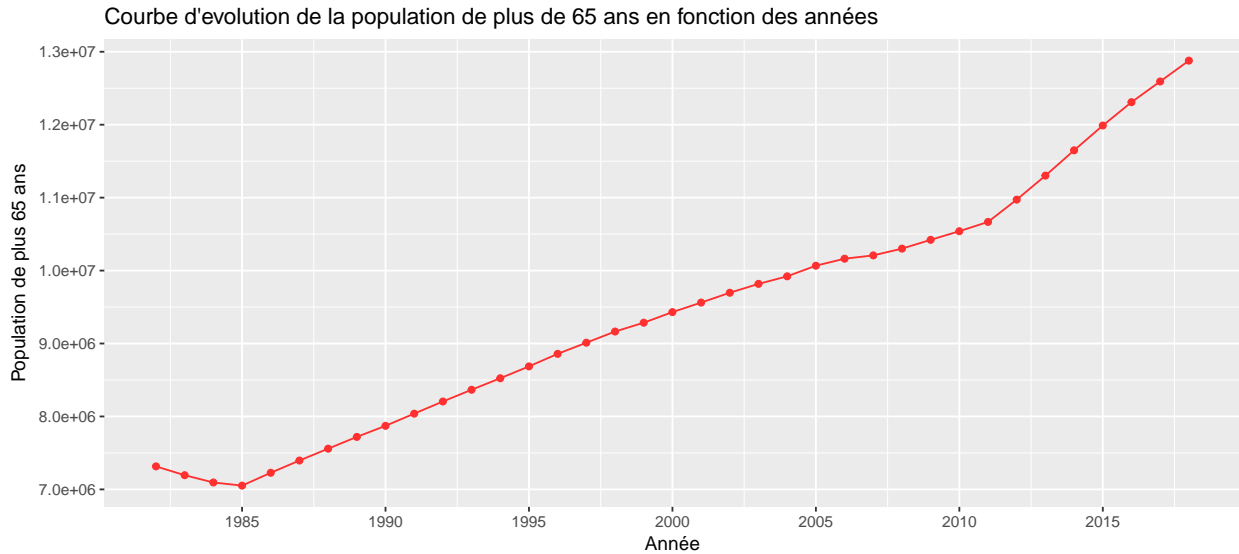


- De 1982 à 1994: nous constatons une chute du nombre de naissances vivantes par année. Entre ces deux périodes nous avons des évolutions croissantes à dents de scie entre 1985 et 1987.
- De 1995 à 2010, une évolution croissante dans l'ensemble des naissances est constatée.
- De 2010 à 2017, le nombre de naissances vivantes diminue considérablement.

Courbe d'évolution de la population de plus de 65 ans en fonction des années

La courbe ci-dessous met en évidence l'évolution de la population de plus de 65 ans en fonction des années. La couleur rouge est prise pour mettre en exergue la forte augmentation du nombre de vieillissement par année.

```
ggplot(data=df2, aes(Annee, populationVieux, group =1)) +
  scale_x_continuous(breaks = round(seq(1980, 2020, by = 5),1)) +
  geom_point(col="firebrick1") +
  geom_line(col="firebrick1") +
  labs(x = "Année",
       y = "Population de plus 65 ans",
       title = "Courbe d'evolution de la population de plus de 65 ans en fonction des années")
```

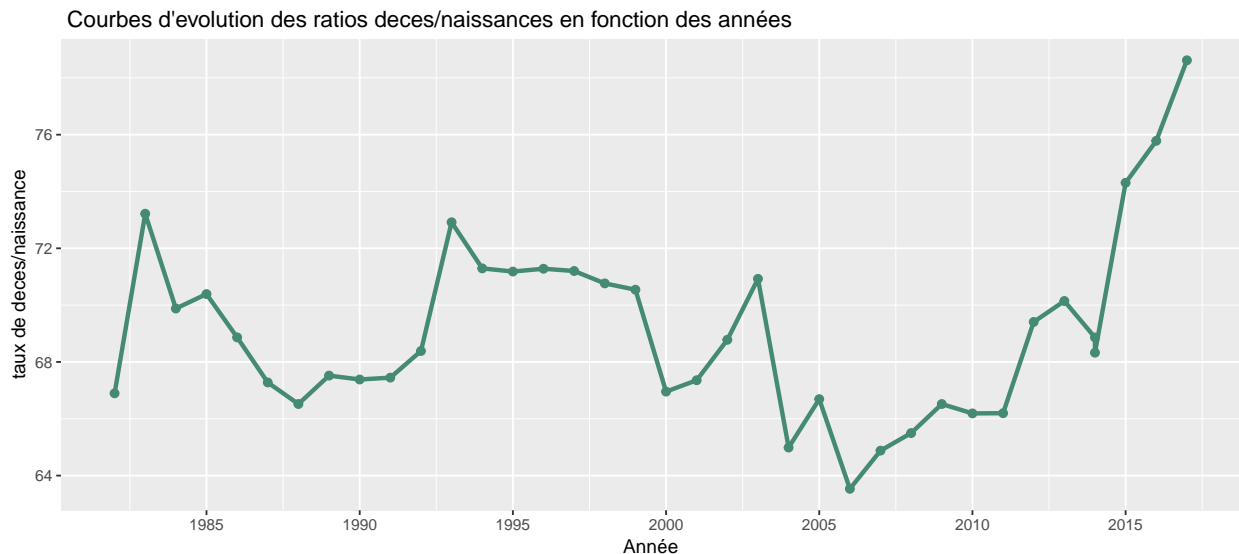


Nous constatons un vieillissement de la population au fil des années sauf entre 1982 et 1985.

Courbe d'évolution du pourcentage du nombre de deces sur le nombre de naissances par an

Le graphe ci-dessous montre l'évolution du pourcentage du nombre de deces sur nombre de naissances par an.

```
ggplot(data=df1, aes(Annee,Ratio, group =1)) +
  scale_x_continuous(breaks = round(seq(1980, 2020, by = 5),1)) +
  geom_point(col="aquamarine4", size = 2) +
  geom_line(col="aquamarine4", size = 1.2) +
  labs(x = "Année",
       y = "taux de deces/naissance",
       title = " Courbes d'evolution des ratios deces/naissances en fonction des années")
```

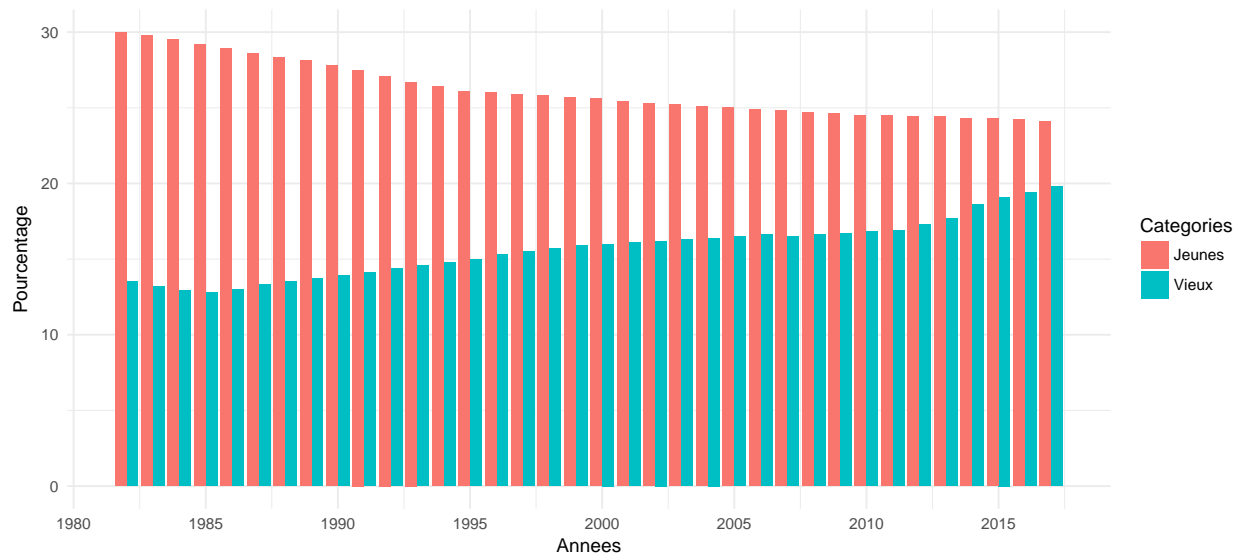


- De 1982 et 1994 et 1995 à 2010 : Nous constatons dans l'ensemble une baisse du ratio, avec une forte augmentation entre 1982 et 1983.
- De 2010 à 2017: Ce taux devient de plus en plus important et a atteint presque les 98%.

Graph bar comparatif entre les tranches d'âges (Jeunes: moins de 20ans et Vieux: plus de 65 ans)

ce **graphe bar** a été tracé afin de représenter les tranches d'âges de la population française sur la période de 1982 à 2017. Les couleurs ont encore une fois fait l'objet d'un choix, choix duquel les naissances sont en bleu car augmentant la population et les décès en rouge car ayant un effet inverse sur cette dernière.

```
df2 <- data.frame(Categories=rep(c("Jeunes", "Vieux"), each=37),  
Annees=rep(c(df1$Annee),2),  
Pourcentage=c(tranchesJeunes,tranchesVieux))  
ggplot(data=df2, aes(x=Annees, y=Pourcentage, fill=Categories)) +  
geom_bar(stat="identity", position=position_dodge()) +  
scale_x_continuous(breaks = round(seq(1980, 2020, by = 5),1)) +  
theme_minimal()
```



Réponses aux questions:

Question I:

Partant des représentations graphiques (courbes) du nombre de naissances et de la répartition des personnes du troisième âge au sein de la population, on peut prendre avec pincette l'impact inversement proportionnel de l'évolution du nombre de naissances sur l'augmentation du nombre de personnes du troisième âge. En effet, de 1985 à 2017, la population ne cesse de vieillir, indépendamment du nombre de naissances vivantes enregistré sachant que ce dernier a connu une augmentation de 1995 à 2010; ce qui selon notre hypothèse de départ, devrait diminuer le nombre de personnes âgées.

Question II:

Après analyse de la courbe d'évolution du ratio (Mortalité/Naissance) et du graphe comparatif des répartitions des tranches d'âges, il est à noter l'impact du ratio sur ces tranches d'âges. En effet, le vieillissement de la population est totalement dépendant à ce ratio. En 2017 par exemple, quand le ratio atteint son pic, la population française atteint également son summum de vieillesse en 37 ans depuis 1982.