

L'IMPACT DU VACCIN CONTRE LE COVID-19 SUR LES CHIFFRES HOSPITALIERS EN FRANCE MÉTROPOLITAINE

Projet final – Analyse de données

Hugo Carlin
Ayoub Bridaoui

Table des matières

Introduction	3
Point sur la situation	4
Hospitalisation, réanimation et décès, des hôpitaux inquiets	4
Une majorité de seconde dose	5
La vaccination : déterminante pour faire baisser les chiffres hospitaliers	6
Conclusion	7
Annexes	8

Introduction

En décembre 2019 un virus inconnu est détecté sur un marché chinois dans la ville de Wuhan, au cours du mois de janvier le virus se propage et sort des frontières chinoises pour arriver en France à la fin de ce même mois. Les premiers décès sont enregistrés et le monde constate que ce virus appelé « coronavirus » ou « Covid-19 » se propage très facilement à cause de sa contagiosité. Le nombre de cas positif augmente, le nombre de décès également, ce qui inquiète les pays du monde entier. En France les premières précautions dans le but de freiner le virus sont mises en place, l'objectif est de minimiser les contacts entre les Français.

Le 17 mars 2020 à midi, le gouvernement français donne l'ordre à ses citoyens de rester chez eux sous peine de sanction, la France est confinée pour la première fois de son histoire. Après 2 mois de confinement la France est « déconfinée » le 11 mai 2020, le virus est toujours présent mais les chiffres sont moins préoccupants, les français peuvent souffler.

L'été passe, la rentrée se profile mais les chiffres liés au Covid-19 inquiètent de nouveau. Le nombre de personnes contaminées augmente, les cas de décès également, ce qui provoque une nouvelle réaction du gouvernement français. Le 30 octobre 2020 Emmanuel Macron annonce un second confinement. L'ensemble des Français est confiné avec une réglementation moins stricte mais comme pour le premier confinement la vie professionnelle de chacun est bouleversée par différentes contraintes sanitaires. À la veille de Noël, le 15 décembre 2020, les Français retrouvent leur liberté et peuvent se retrouver pour les fêtes de fin d'années, mais sous certaines conditions. Pour la première fois le gouvernement met en place un couvre-feu, ce qui restreint tout de même les individus.

Depuis cette fin d'année 2020 jusqu'à aujourd'hui, un an après, nous sommes passés par différentes phases telles que divers couvre-feux, des confinements moins strictes ou encore localisés et de nombreux variants du virus de base ont vu le jour, certains plus dangereux que d'autres. Cependant une nouvelle variable, étant censée faire baisser les chiffres hospitaliers, a vu le jour : la vaccination contre le Covid-19. Ce vaccin a pour but de minimiser les contaminations mais surtout de limiter les cas graves de Covid-19 et ainsi faire baisser les chiffres d'hospitalisation, de réanimation ou encore de décès au sein des hôpitaux. Deux ans après le début de cette pandémie, nous vivons au dépend de ces chiffres hospitaliers, s'ils sont trop haut nous devons adapter nos vies dans le but de les faire baisser et lorsqu'ils sont assez bas nous retrouvons nos libertés, malgré la constante présence des gestes barrières qui deviennent une norme à nos vies. Dans le but de faire baisser ces chiffres 76% des Français se sont fait vacciner.

À travers ce projet nous allons chercher si cette vaccination diminue véritablement les cas graves de Covid-19 et donc les chiffres concernant les hospitalisations en France métropolitaine.

Point sur la situation

Hospitalisation, réanimation et décès, des hôpitaux inquiets

L'intégralité de l'étude est réalisée le 19 décembre 2021, en pleine montée de la 5^{ème} vague portée par un nouveau variant appelé « Omicron ». Après presque 2 ans depuis le début de la pandémie en France, on a observé à plusieurs reprises des hôpitaux surpeuplés, ce qui a poussé le gouvernement à appliquer certaines mesures. Sur la *figure 1* ci-dessous on observe que 0.87% des Français métropolitains ont été hospitalisés à la suite d'une contamination, 0.16% ont été admis en réanimation et 0.14% sont décédés du Covid-19. Les différentes études réalisées ont montré que les individus subissant des complications à cause du Covid-19 comportent pour la grande majorité des comorbidités ou ont un certain âge. On trouve que 65% des patients décédés du Covid-19 comportaient des comorbidités ou encore que 75% avaient 75 ans ou plus¹, ce qui prouve une plus grande probabilité de faire une forme grave lorsqu'on présente l'une des deux caractéristiques précédemment citées. Cependant une part de la population, certes minime, contracte une forme grave sans être âgé et sans présentant de comorbidités², c'est pour cela que le vaccin a été créé, dans le but de limiter au maximum les formes graves du Covid-19 et ainsi limiter les hospitalisations, les entrées en réanimation et les décès.

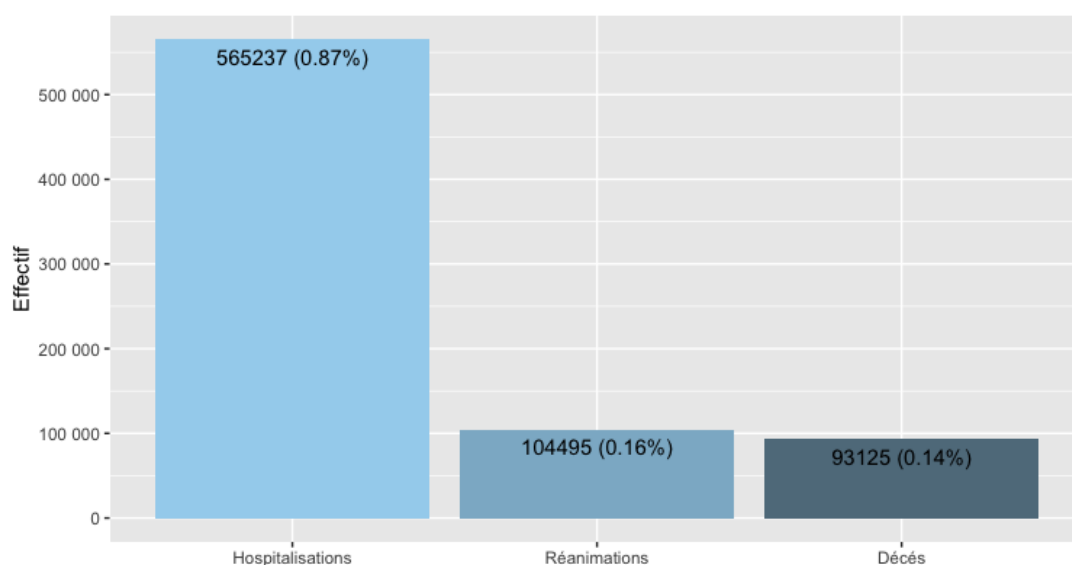


Figure 1 – Effectifs et fréquences des individus ayant subis des complications dues au Covid-19 en France métropolitaine, entre le début de la pandémie et le 19/12/2021. Grille de lecture : 565 237 français métropolitains ont été hospitalisés à cause du Covid-19 entre le début de la pandémie et le 19/12/2021, ce qui représente 0.87% de la population française métropolitaine. Source : les auteurs, selon les données de Santé Publique France.

¹ Thiébaux A. (18/11/21). Victimes du Covid en France : âge, profil, jeune, nombre. *Le journal des femmes* [en ligne] (page consultée le 27/12/21).

² Donada E. (02/02/21). Combien de moins de 18 ans sans comorbidités sont mortes du Covid-19 ? *Libération* [en ligne] (page consultée le 27/12/21).

Une majorité de seconde dose

Le 19 décembre 2021, les Français, pour la majorité, commencent leur troisième injection de vaccin contre le Covid-19. Il faut savoir que la vaccination est prioritaire pour les personnes à risque, c'est-à-dire, comme expliqué précédemment, les personnes comportant des comorbidités ou étant plus âgées. Certains individus ne veulent pas se faire vacciner et d'autre ont commencé leur vaccination plus tardivement, ce qui fait, comme on peut l'observer sur la *figure 2*, que l'on a 4 cas de vaccination. On peut voir que 76.11% des Français métropolitains ont au moins été vacciné à hauteur d'une dose. Seulement 1.04% des Français ayant reçu une dose n'ont pas encore reçu leur seconde dose, ce qui nous donne 75.05% de la population française métropolitaine étant vaccinés à hauteur de deux doses. La troisième dose a été administrée à 28.82% des Français, un écart plus conséquent qu'entre les deux premières car elle n'est disponible que depuis peu de temps. Cependant 23.89% des Français métropolitains ne sont pas vaccinés, et risqueraient, d'après les études³, de contracter des formes plus graves de Covid-19 que les individus vaccinés. On a donc différentes populations vaccinées par une, deux ou trois doses, et d'autres non vaccinés.

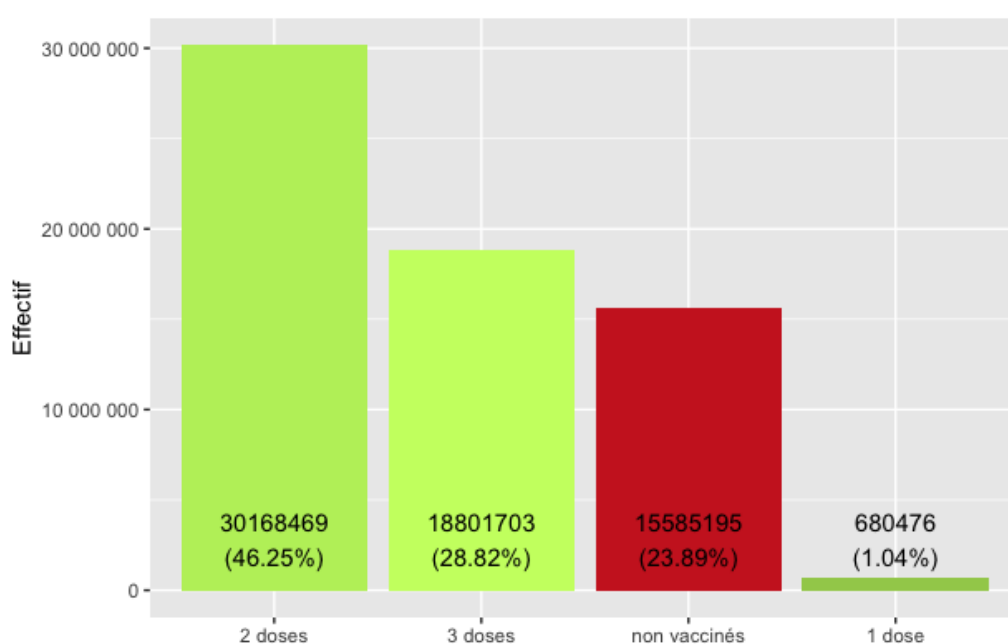


Figure 2 – Effectifs et fréquences des individus vaccinés contre le Covid-19 en France métropolitaine, entre le début de la pandémie et le 19/12/2021. Grille de lecture : 30 168 469 français métropolitains ont reçu exactement 2 doses de vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021, ce qui représente 46.25% de la population française métropolitaine. Source : les auteurs, selon les données de Santé Publique France.

³ Douin C. (11/10/21). « La vaccination réduit de 90% le risque d'hospitalisation » selon la plus vaste étude menée sur le Covid-19. *Midi Libre* [en ligne] (page consultée le 28/12/21).

La vaccination : déterminante pour faire baisser les chiffres hospitaliers

Pour reprendre la problématique de ce projet, nous cherchons à expliquer si la vaccination réduit réellement les chiffres hospitaliers dus au Covid-19 en France métropolitaine. Nous rappelons que les chiffres hospitaliers décrivent le nombre d'hospitalisations, de réanimations et de décès dus au Covid-19. À l'aide de données publiées par Santé Publique France et l'INSEE ainsi que diverses méthodes statistiques nous avons cherché à expliquer ceci entre le début de la pandémie et le 19/12/2021 (cf. *Annexe 1*). Nous rappelons également que l'étude est réalisée en analysant les données de chacun des 97 départements de France métropolitaine.

Dans un premier temps, nous observons clairement que le fait d'être non vacciné est lié aux chiffres hospitaliers dus au Covid-19, autrement dit, plus le taux de non vacciné est élevé dans les départements plus les chiffres hospitaliers le sont également (cf. *Annexe 3* et *Annexe 6*). Dans un second temps, on arrive à des résultats montrant que le fait d'être vacciné à hauteur de 3 doses à l'effet totalement inverse des non vaccinés : les départements ayant un taux élevé de personnes ayant reçu 3 doses de vaccin contre le Covid-19 présentent des chiffres hospitaliers plus faibles (cf. *Annexe 6*).

Pour aller plus loin et dans le but d'expliquer la vaccination et les chiffres hospitaliers avec d'autres variables nous nous sommes intéressés à l'âge. Après avoir remarqué des interprétations difficiles (cf. *Annexe 3* et *Annexe 12*) nous avons réussi à tirer quelques conclusions en considérant cette variable. Les personnes ayant un âge plus avancé (60 ans et plus) sont dépendants au fait d'avoir été vacciné 3 fois, ce qui s'explique par la priorité à ces derniers d'être vaccinés en premiers, nous avons donc des résultats logiques. Par ailleurs, nous avons remarqué que les non vaccinés sont liés avec le fait d'avoir des âges entre 0 et 39 ans, ce qui s'explique d'abord par le faible taux de vaccination des plus jeunes à l'heure actuelle et on peut supposer également que les individus ne voulant pas se faire vacciner sont des personnes ayant en moyenne moins de la quarantaine (cf. *Annexe 9*). En aucun cas nous n'avons observé que le fait d'être âgé est lié avec le fait d'être admis à l'hôpital, c'est pourtant le cas comme nous l'avons expliqué précédemment.

Conclusion

Le 19 décembre 2021, deux ans après le début de la pandémie mondiale du Covid-19 en France, on compte plus de 550 000 hospitalisations, plus de 100 000 admissions en réanimation et un peu moins de 100 000 décès à cause de ce virus. Ce qui représente une faible part de la population française métropolitaine mais suffisante pour s'inquiéter. La vaccination, réalisée par plus de 75% des Français a pour but de réduire l'occupation des hôpitaux et ainsi réduire le nombre de cas graves de Covid-19. Se vacciner diminue grandement les chances d'être hospitalisé, admis en réanimation ou encore décéder à cause de ce virus, les Français non vaccinés augmentent leurs chances de subir un de ces cas graves. Des variables comme l'âge se corrélaient avec le fait d'être plus vacciné que d'autres pour certaines tranches d'âges ou encore augmentent les risques de faire un cas grave de Covid-19 : principalement pour les personnes âgées. On sait que les comorbidités augmentent également ces risques, il pourrait être intéressant de s'y pencher afin de poursuivre cette étude.

Table des annexes

Annexe 1 – Description des variables utilisées pour réaliser l'étude.	9
Annexe 2 – Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant toutes les variables exploitables.....	9
Annexe 3 – Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant toutes les variables exploitables.....	9
Annexe 4 – Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant toutes les variables exploitables.....	9
Annexe 5 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables hospitalières et de vaccination.	9
Annexe 6 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables hospitalières et de vaccination	9
Annexe 7 - Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant les variables hospitalières et de vaccination	9
Annexe 8 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables d'âge et de vaccination	9
Annexe 9 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et de vaccination	9
Annexe 10 - Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et de vaccination	9
Annexe 11 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables d'âge et hospitalières.....	9
Annexe 12 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et hospitalières	9

Annexes

Annexe 1 : base de données et variables utilisées pour réaliser cette étude

- Variable de localisation :

Variable	<i>dep</i>
Type	qualitative
Description	département

Chacune des observations concerne un département français métropolitain, on en compte donc 97 au total.

- Variables sur les chiffres hospitaliers :

Variable	<i>hosp</i>	<i>rea</i>	<i>dc</i>
Type	quantitative	quantitative	quantitative
Description	Nombre d'hospitalisations dues au Covid-19 entre le début de la pandémie et le 19/12/2021	Nombre de réanimations dues au Covid-19 entre le début de la pandémie et le 19/12/2021	Nombre de décès dues au Covid-19 entre le début de la pandémie et le 19/12/2021

- Variables sur la vaccination :

Variable	<i>dose1</i>	<i>dose2</i>	<i>dose3</i>	<i>non_vacc</i>
Type	quantitative	quantitative	quantitative	quantitative
Description	Nombre d'individus ayant reçu exactement 1 dose du vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021	Nombre d'individus ayant reçu exactement 2 doses du vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021	Nombre d'individus ayant reçu exactement 3 doses du vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021	Nombre d'individus n'ayant pas reçu de doses du vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021

- Variables sur la population française :

Variable	<i>age_0_19ans</i>	<i>age_20_39ans</i>	<i>age_40_59ans</i>	<i>age_60_74ans</i>	<i>age_75+</i>	<i>pop</i>
Type	quantitative	quantitative	quantitative	quantitative	quantitative	quantitative
Description	Nombre d'individus étant âgés de 0 à 19 ans en 2021.	Nombre d'individus étant âgés de 20 à 39 ans en 2021.	Nombre d'individus étant âgés de 40 à 59 ans en 2021.	Nombre d'individus étant âgés de 60 à 74 ans en 2021.	Nombre d'individus étant âgés de 75 ans et plus en 2021.	Nombre d'individus total.

Annexe 1 – Description des variables utilisées pour réaliser l'étude. Grille de lecture : la variable *dose1* est quantitative et représente le nombre d'individus ayant reçu exactement 1 dose du vaccin contre le Covid-19 entre le début de la pandémie et le 19/12/2021. Source : les auteurs, selon les données de Santé Publique France.

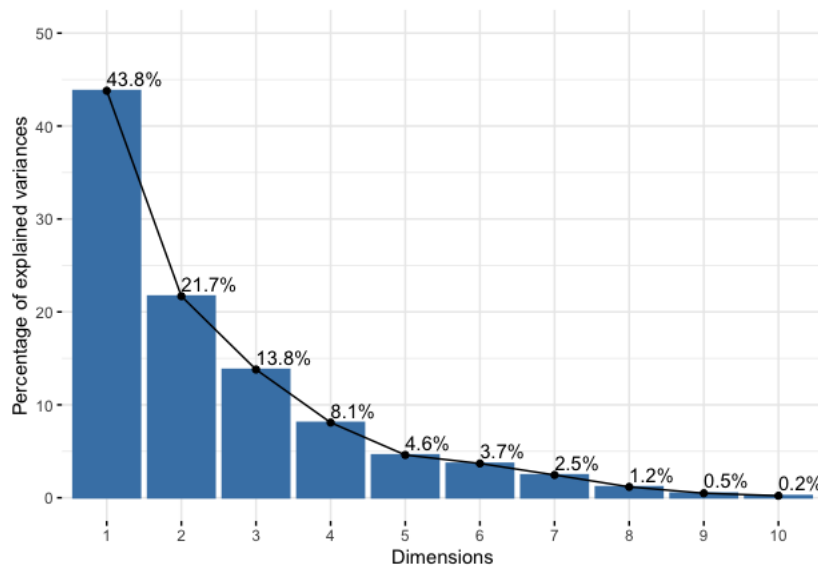
Notre base de données est le résultat de la jointure de 3 différentes bases provenant de *Santé Publique France* et de *l'INSEE* :

- Les données hospitalières dues au Covid-19 (hospitalisations, réanimations et décès) pour chaque département français métropolitain et pour chaque jour entre le début de la pandémie et le 19/12/21 (*Santé Publique France*).
- Les données de la vaccination contre le Covid-19 (1 dose, 2 doses et 3 doses reçues) pour chaque département français métropolitain et pour chaque jour entre le début de la pandémie et le 19/12/21 (*Santé Publique France*).
- Une estimation de la population française totale et selon des tranches d'âge pour chaque département français métropolitain en 2021 (*INSEE*).

La jointure a été réalisée sur le numéro des départements, ce qui nous donne une base contenant toutes les informations précédentes pour chaque département français métropolitain. Dans le but de comparer chacun des départements en supprimant l'effet taille que procure la population, nous avons passé notre base de données en taux d'incidence pour 100 000 habitants.

Le fait d'avoir uniquement des variables quantitatives, en excluant les départements qui caractérisent les « individus », la méthode que l'on va utiliser est l'Analyse des Composantes Principales qui va nous permettre de trouver des corrélations entre les différentes variables que nous avons et ainsi en tirer des conclusions.

Annexe 2 : qualité de projection de la méthode d'ACP contenant toutes les variables exploitables

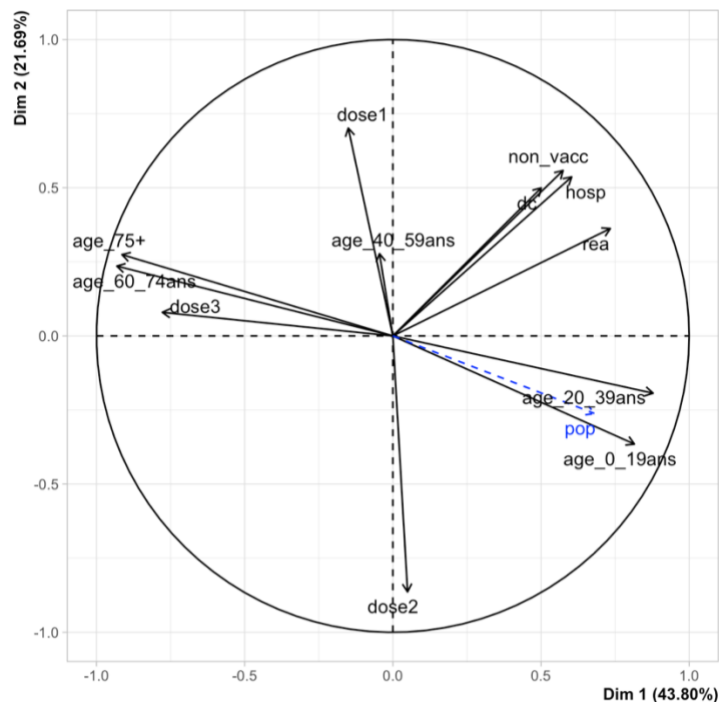


Annexe 2 – Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant toutes les variables exploitables. Grille de lecture : La dimension 1 regroupe 43.8% d'inertie. Source : les auteurs, selon les données de Santé Publique France.

Pour commencer notre étude nous réalisons une analyse des composantes principales sur l'ensemble de nos variables, la variable population n'étant pas utile elle est mise en variables quantitative supplémentaire.

L'Annexe 2 représente la qualité de projection des différentes dimensions de la méthode ACP utilisée. On peut voir que la dimension 1 regroupe 43.8% d'inertie, ce qui traduit une bonne qualité de représentation de la méthode ACP sur cette dimension. On gardera en plus de cette dimension, la dimension 2 qui est la seconde dimension la mieux représentée avec 21.7%. Si nous résumons nos 13 variables par ces deux dimensions alors on récupère $43.8 + 21.7 = 65.5\%$ de l'information contenu dans le tableau de données.

Annexe 3 : graphe des variables de l'ACP concernant toutes les variables exploitables



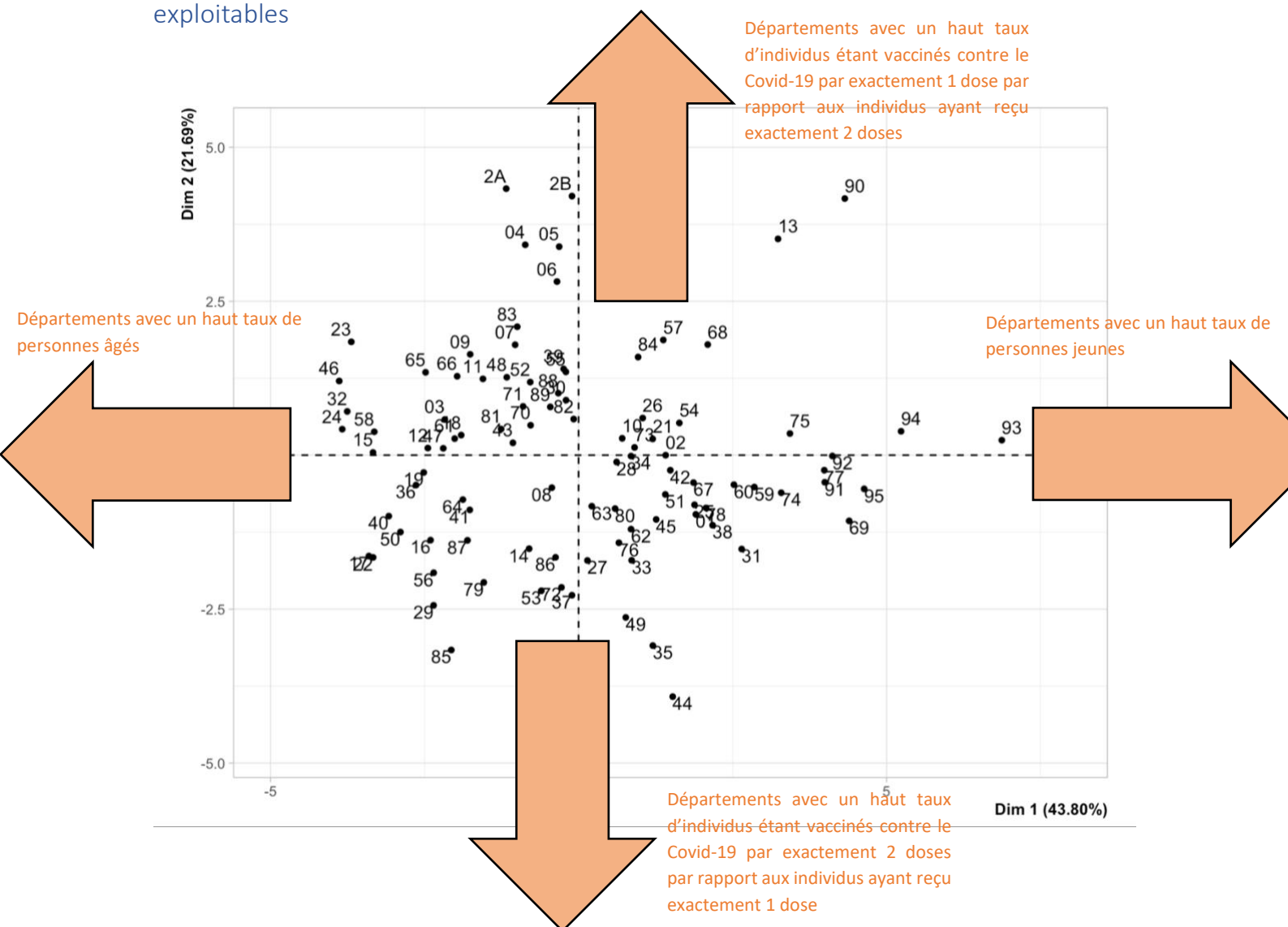
Annexe 3 – Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant toutes les variables exploitables. Grille de lecture : Les variables *age_60_74ans* et *dose3* forment un angle aigu, elles sont donc dépendantes l'une de l'autre. Source : les auteurs, selon les données de Santé Publique France.

On voit deux informations importantes sur le graphe des variables de l'ACP de l'Annexe 3. Tout d'abord la *dose3* forme un angle aigu très faible avec les variables *age_60_74ans* et *age_75+*, ce qui se traduit par une forte dépendance entre la troisième dose et ces deux catégories d'âge. Une observation qui s'explique par la priorité de vaccination de ces tranches d'âge. Ensuite, on remarque que la variable *non_vacc* est fortement corrélée avec les variables *dc*, *hosp* et *rea* : on en conclue une des principales réponses à notre problématique, que les non vaccinés est dans une catégorie très dépendant aux décès, aux hospitalisations et à la réanimation.

En ce qui concerne les axes, on remarque que les variables concernant un âge plus élevé s'expliquent sur la gauche de l'axe 1 et les variables représentant un jeune âge sur la droite de ce même axe. L'axe 2, quant à lui, nous montre une dépendance aux variables *dose1* et *dose2*, avec vers le haut les personnes vaccinées une fois et vers le bas deux fois. La variable *dose3*, fortement corrélée aux individus âgés, s'explique sur l'axe 1 avec ces derniers.

Malgré cette explication des axes, et au-delà des informations relevées, il est difficile de résumer correctement les 14 variables présentes dans ce graphe. L'axe 2 est difficilement interprétable. Les individus âgés par exemple devraient être davantage en corrélation avec les variables hospitalières car ils font partie des personnes les plus à risques et les plus hospitalisés. La catégorie des 40-59 ans représentant la moyenne d'âge des Français est représentée par une petite flèche, ce qui veut dire qu'elle est mal expliquée.

Annexe 4 : graphe des individus de l'ACP concernant toutes les variables exploitables

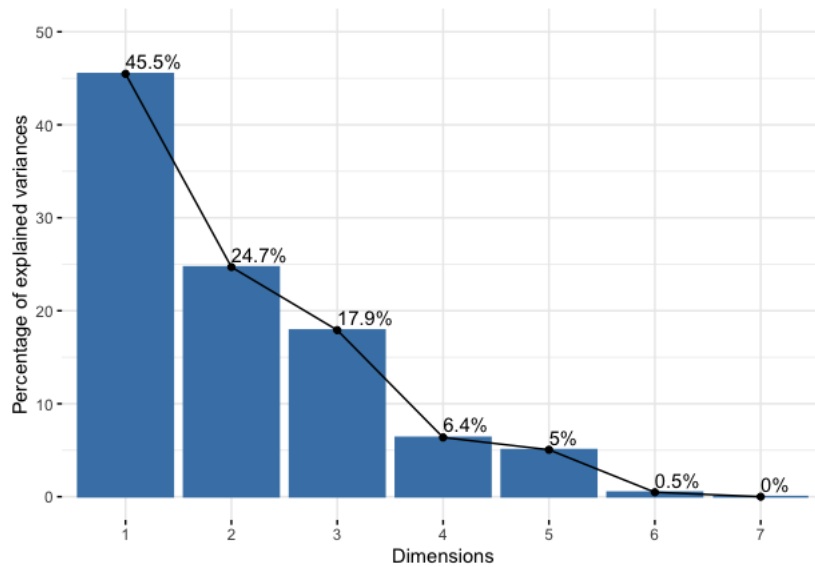


Annexe 4 – Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant toutes les variables exploitables. Grille de lecture : Les départements 90 et 13 sont proches graphiquement, ils présentent donc des similitudes au niveau de leurs variables. Source : les auteurs, selon les données de Santé Publique France.

L'axe de la dimension 1, qui ici représente principalement le taux des âges des individus dans chaque département est celui que sépare au mieux les points car il est le mieux expliqué. On peut donc interpréter que deux départements éloignés par rapport à l'axe des abscisses ont des comportements différents au niveau des âges de leurs habitants. L'axe des ordonnées, s'expliquant moins bien à l'air de représenter la différence entre taux de vaccination (une ou deux doses) des départements. Les départements en haut ont donc une différence entre les variables *dose1* et *dose2* plus grande que ceux qui se trouvent vers le bas.

On dit de deux départements proches qu'ils se ressemblent au niveau de leur comportement, cependant cette première méthode est compliquée à interpréter correctement, on a donc décidé d'ajouter des variables supplémentaires pour cibler certaines variables et en tirer des conclusions plus précises.

Annexe 5 : qualité de projection de la méthode d'ACP contenant les variables hospitalières et de vaccination

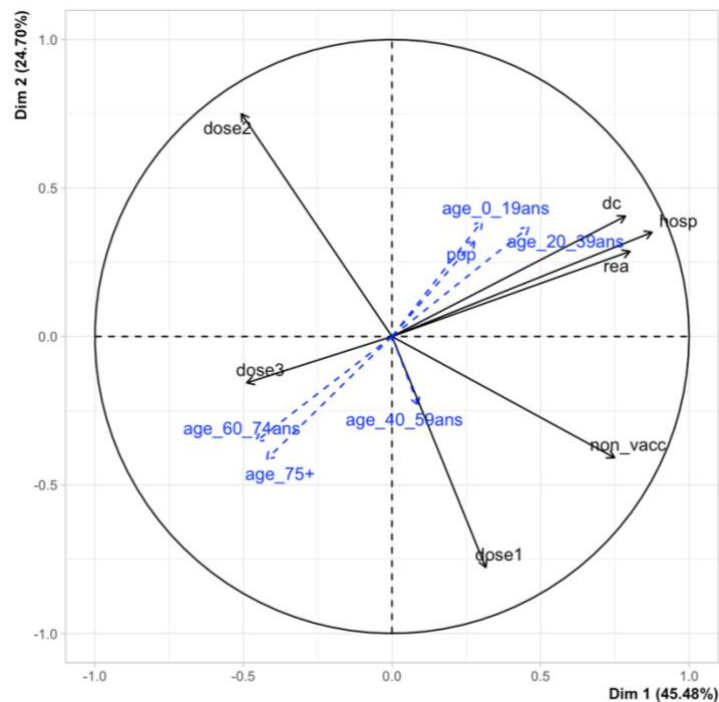


Annexe 5 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables hospitalières et de vaccination. Grille de lecture : La dimension 1 regroupe 45.5% d'inertie. Source : les auteurs, selon les données de Santé Publique France.

À présent, pour cibler notre étude sur nous réalisons une analyse des composantes principales sur les variables hospitalières et de vaccination, les autres variables sont mises en variables quantitatives supplémentaires. L'objectif est de trouver des liens existants entre les chiffres hospitaliers et la vaccination des Français.

L'Annexe 5 représente la qualité de projection des différentes dimensions de la méthode ACP décrite plus haut. On peut voir que la dimension 1 regroupe 45.5% d'inertie, ce qui traduit une bonne qualité de représentation de la méthode ACP sur cette dimension. On gardera en plus de cette dimension, la dimension 2 qui est la seconde dimension la mieux représentée avec 24.7%. Si nous résumons nos 8 variables par ces deux dimensions alors on récupère $45.5 + 24.7 = 70.2\%$ de l'information contenu dans le tableau de données.

Annexe 6 : graphe des variables de l'ACP concernant les variables hospitalières et de vaccination



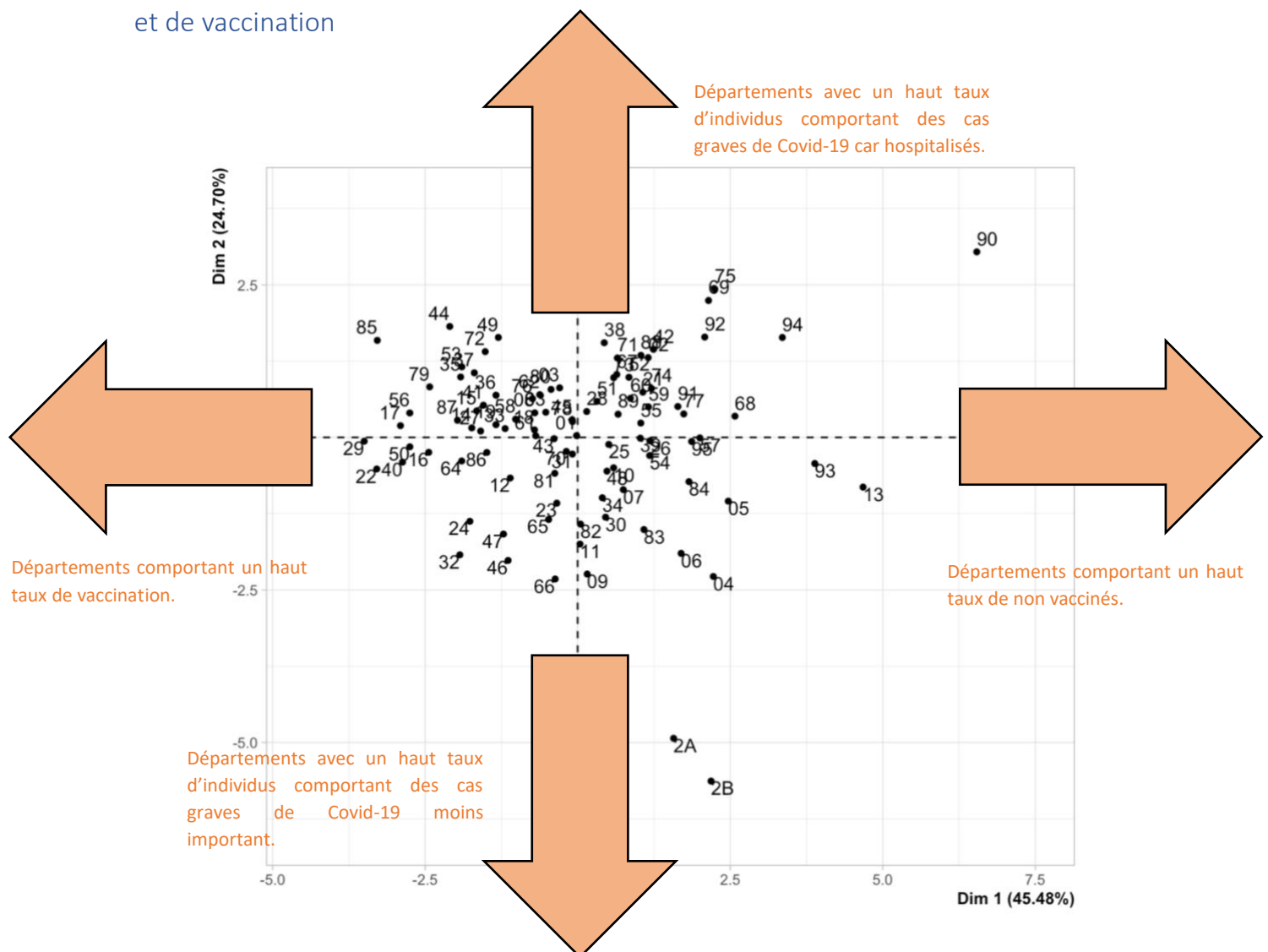
Annexe 6 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables hospitalières et de vaccination. *Grille de lecture :* Les variables *non_vacc* et *rea* forment un angle aigu, elles sont donc dépendantes l'une de l'autre. *Source :* les auteurs, selon les données de Santé Publique France.

Dans le but de cibler la relation entre les variables hospitalières et la vaccination nous avons gardé uniquement ces dernières. Le graphe de l'Annexe 6 nous montre que les trois variables concernant les chiffres hospitaliers sont dépendantes, sans surprise car elles sont bien en relation. Cette représentation est très intéressante, tout d'abord on voit que les variables *dc*, *hosp* et *rea* forment un angle plat avec ce qui se traduit par une anticorrélation entre eux : plus les départements ont un haut taux d'individus ayant reçu une troisième dose, moins les taux des chiffres hospitaliers sont importants. Au contraire, les variables hospitalières forment un angle aigu avec la variable des non vaccinés, ce qui montre une corrélation positive entre eux : plus le taux des non vaccinés est élevé plus les variables hospitalières le sont également. Cependant les variables hospitalières et les premières et secondes doses forment deux angles droits, il n'y a donc aucune relation entre elles.

L'axe des abscisses (dimension 1) montre une corrélation avec la variable *dose3* (à gauche) et *non_vacc* (à droite), il représente donc principalement le taux de vaccination. L'axe des ordonnées (dimension 2) représente un nombre plus élevé des chiffres hospitaliers vers le haut et moins élevés vers le bas.

Nous avons à présent une représentation signification de nos variables hospitalières et de vaccination, ce qui donne une première réponse à notre problématique.

Annexe 7 : graphe des individus de l'ACP concernant les variables hospitalières et de vaccination

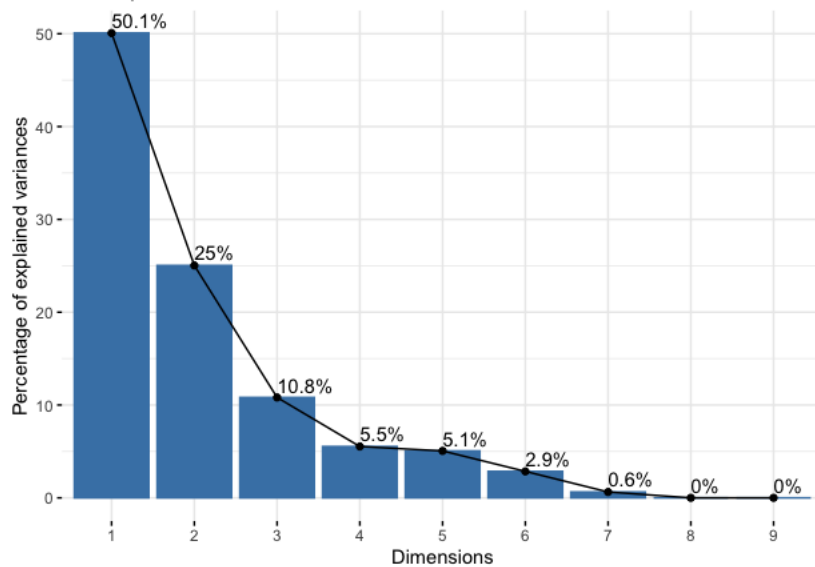


Annexe 7 - Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant les variables hospitalières et de vaccination. Grille de lecture : Les départements 93 et 13 sont proches graphiquement, ils présentent donc des similitudes au niveau de leurs variables. Source : les auteurs, selon les données de Santé Publique France.

L'axe de la dimension 1 (abscisses), représentant le taux de vaccination, est le mieux représenté. L'axe de la dimension 2 (ordonnées) représente les taux des variables hospitalières. Par exemple, la variable du département 90 en haut à droite s'exprime par : un département où le taux de non vaccinés est important avec un taux d'hospitalisations, de réanimations et de décès important, logique d'après ce qu'on a dit précédemment. La corse (2A et 2B) est la seule petite exception qui est représentée avec un taux de vaccination en dessous de la moyenne avec un taux d'hospitalisation faible.

Deux départements proches graphiquement ont des taux de vaccination et de chiffres hospitaliers similaires. Deux départements éloignés sur l'axe 1 sont différents à propos de leur comportement de vaccination, s'ils sont éloignés sur l'axe 2 ils se comportent différemment par rapport aux chiffres hospitaliers.

Annexe 8 : qualité de projection de la méthode d'ACP contenant les variables d'âge et de vaccination

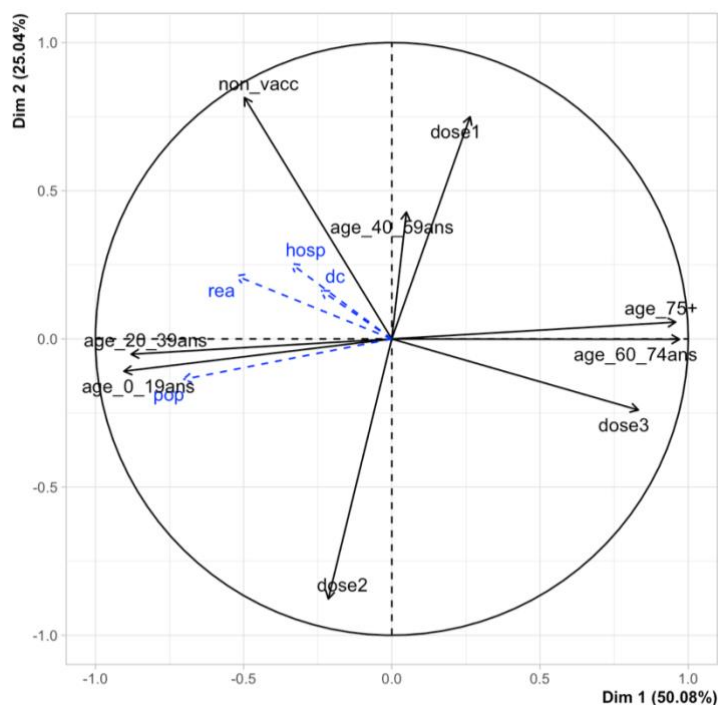


Annexe 8 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables d'âge et de vaccination. Grille de lecture : La dimension 1 regroupe 50.1% d'inertie. Source : les auteurs, selon les données de Santé Publique France.

En supplément de notre problématique, pour cibler notre étude sur les variables de vaccination selon l'âge nous réalisons une analyse des composantes principales sur les variables d'âge et de vaccination, les autres variables sont mises en variables quantitatives supplémentaires. L'objectif est de trouver des liens existants entre la vaccination et l'âge des Français.

L'Annexe 8 représente la qualité de projection des différentes dimensions de la méthode ACP décrite plus haut. On peut voir que la dimension 1 regroupe 50.1% d'inertie, ce qui traduit une bonne qualité de représentation de la méthode ACP sur cette dimension. On gardera en plus de cette dimension, la dimension 2 qui est la seconde dimension la mieux représentée avec 25%. Si nous résumons nos 11 variables par ces deux dimensions alors on récupère $50.1 + 25 = 75.1\%$ de l'information contenu dans le tableau de données.

Annexe 9 : graphe des variables de l'ACP concernant les variables d'âge et de vaccination



Annexe 9 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et de vaccination. Grille de lecture : Les variables *age_60_74ans* et *dose3* forment un angle aigu, elles sont donc dépendantes l'une de l'autre. Source : les auteurs, selon les données de Santé Publique France.

Dans le but de cibler la relation entre les variables d'âge et la vaccination nous avons gardé uniquement ces dernières. Le graphe de l'Annexe 9 nous montre comme pour l'Annexe 3 que les personnes les plus âgées sont corrélées au fait d'être vacciné à hauteur de 3 doses. La variable représentant l'âge entre 40 et 59 ans est encore une fois mal représentée, il est donc compliqué d'en tirer des conclusions. La variable *non_vacc* est corrélée positivement avec les plus jeunes et négativement avec les plus vieux, on peut donc conclure que les non vaccinés sont principalement parmi la population la plus jeune.

L'axe des abscisses (dimension 1) montre une corrélation avec la variable *age_20_39ans* et *age_0_19ans* (à gauche) et *age_75+* et *age_60_74ans* (à droite), il représente donc principalement le taux d'âge des départements. L'axe des ordonnées (dimension 2) représente un taux de vaccination plus élevé vers le bas et moins élevé vers le haut, on voit que les variables *non_vacc* et *dose1* sont corrélées avec le haut de l'axe 2 et les variables *dose2* et *dose3* avec le bas.

Nous avons à présent quelques précisions sur la corrélation entre la vaccination et l'âge des individus en France métropolitaine, le fait de cibler les variables par rapport à la première ACP nous a donc aidé à trouver plus de précision.

vaccination

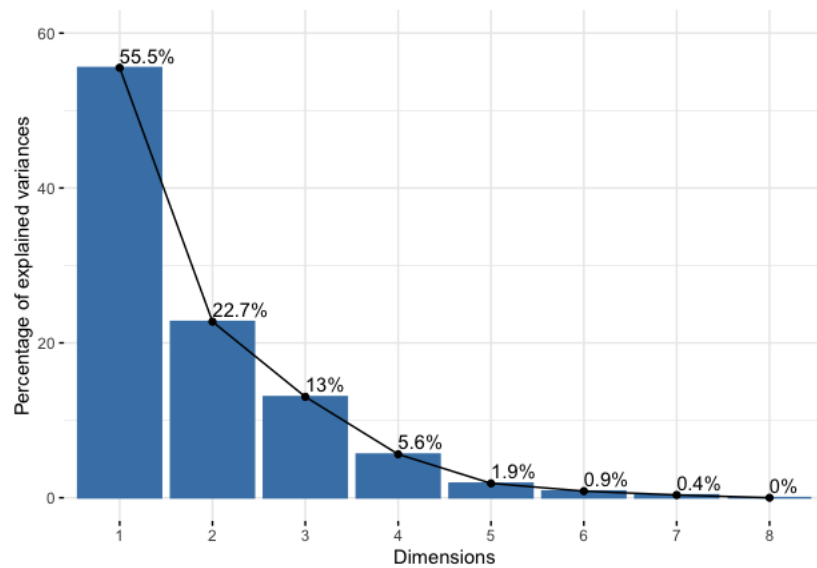


Annexe 10 - Nuage de point représentant les individus, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et de vaccination. Grille de lecture : Les départements 2A et 2B sont proches graphiquement, ils présentent donc des similitudes au niveau de leurs variables. Source : les auteurs, selon les données de Santé Publique France.

L'axe de la dimension 1 (abscisses), représentant l'âge de la population, est le mieux représenté. L'axe de la dimension 2 (ordonnées) représente les taux de vaccination. Par exemple, la variable du département 93 légèrement en haut et à gauche s'exprime par : un département où l'âge est élevé avec un taux vaccination assez faible. La corse (2A et 2B) est représentée avec un taux de vaccination faible (comme pour l'*Annexe 7*) et un âge plutôt moyen.

Deux départements proches graphiquement ont des taux de vaccination et des âges similaires dans leur population. Deux départements éloignés sur l'axe 2 sont différents à propos de leur comportement de vaccination, s'ils sont éloignés sur l'axe 1 ils se comportent différemment par rapport à l'âge.

Annexe 11 : qualité de projection de la méthode d'ACP contenant les variables d'âge et hospitalières

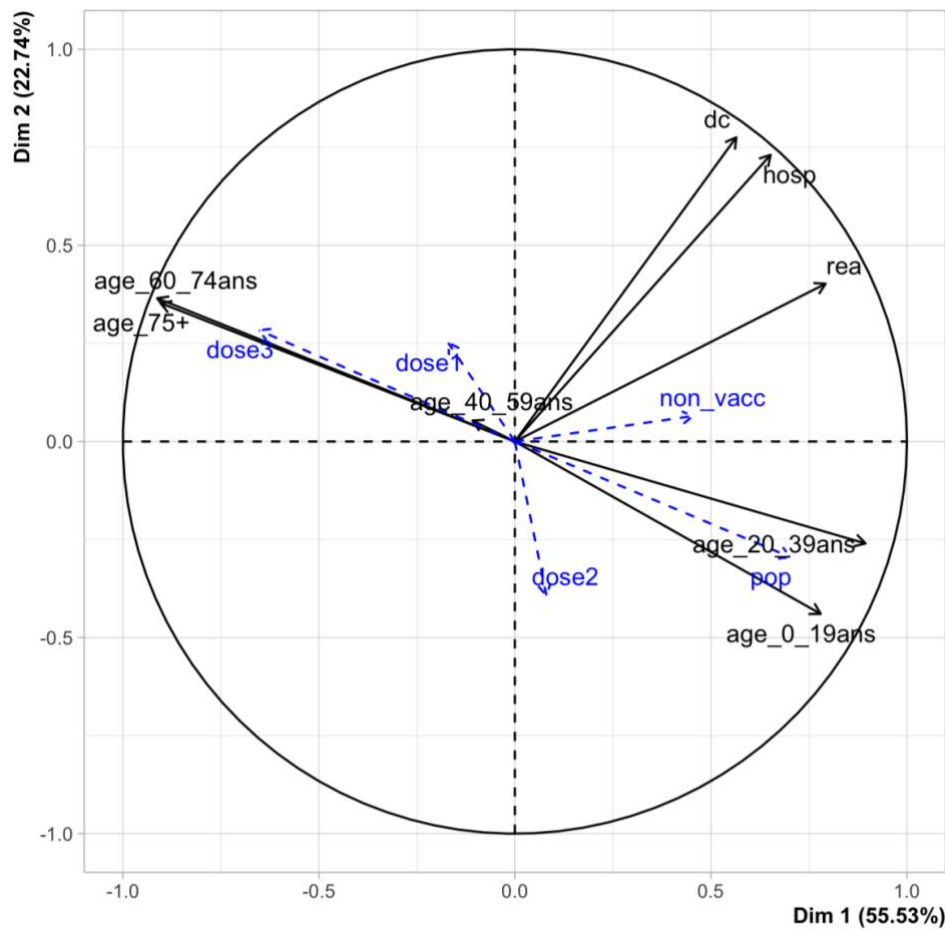


Annexe 11 - Diagramme en barre représentant le pourcentage d'inertie de chaque dimension de la méthode d'ACP contenant les variables d'âge et hospitalières. Grille de lecture : La dimension 1 regroupe 55.5% d'inertie. Source : les auteurs, selon les données de Santé Publique France.

De nouveau en supplément, pour cibler notre étude sur les variables hospitalières par rapport à l'âge nous réalisons une analyse des composantes principales sur les variables d'âge et de vaccination, les autres variables sont mises en variables quantitatives supplémentaires. L'objectif est de trouver des liens existants entre les données hospitalières et l'âge des Français.

L'Annexe 11 représente la qualité de projection des différentes dimensions de la méthode ACP décrite plus haut. On peut voir que la dimension 1 regroupe 55.5% d'inertie, ce qui traduit une bonne qualité de représentation de la méthode ACP sur cette dimension. On gardera en plus de cette dimension, la dimension 2 qui est la seconde dimension la mieux représentée avec 22.7%. Si nous résumons nos 10 variables par ces deux dimensions alors on récupère $55.5 + 22.7 = 78.2\%$ de l'information contenu dans le tableau de données.

Annexe 12 : graphe des variables de l'ACP concernant les variables d'âge et hospitalières



Annexe 12 - Cercle de corrélation des variables, sur les dimensions 1 et 2, de l'ACP concernant les variables d'âge et hospitalières. Grille de lecture : Les variables *age_20_39ans* et *age_0_19ans* forment un angle aigu, elles sont donc dépendantes l'une de l'autre. Source : les auteurs, selon les données de Santé Publique France.

Dans le but de cibler la relation entre les variables d'âge et hospitalières nous avons gardé uniquement ces dernières. Le graphe de l'Annexe 12 nous apprend rien de concret. La variable *age_40_59ans* nous pose une nouvelle fois problème et fausse la représentation, donc l'interprétation du graphe est compliquée. Nous ne tirerons aucune conclusion de cette méthode, il est donc sans intérêt de représenter le graphe des individus.