1

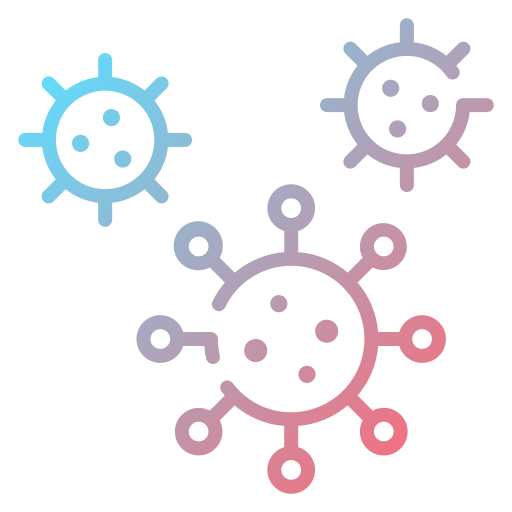
Une image contenant texte, clipart, carte de visite

Description générée automatiquement

2022-2023

**Khaoula Aroui**

Etudiante en Master statistique pour L’évaluation et la prévision



**Propagation du COVID-19 en France**

Table des matières

[Introduction 3](#_Toc126952064)

[Evolution du COVID-19 au cours de l'année et selon les zones 4](#_Toc126952065)

[A l’échelle régionale 5](#_Toc126952066)

[Visualisation d'un ensemble de départements tirés au hasard, représentants de leurs régions ? Homogénéité ou nuances ? 6](#_Toc126952067)

[Etudier deux moments clés dans le temps 7](#_Toc126952068)

[Le 12 décembre, une analyse pour évaluer l'impact de la couverture vaccinale et des caractéristiques démographiques 8](#_Toc126952069)

[Le 10 avril, une analyse pour évaluer l'impact des mesures sanitaires renforcées et des caractéristiques démographiques 10](#_Toc126952070)

[Conclusion 11](#_Toc126952071)

[Annexes 13](#_Toc126952072)

[1. Répartition globale et dispersion des indicateurs de la propagation du COVID-19 13](#_Toc126952073)

[a. Taux d’incidence 13](#_Toc126952074)

[b. Taux de positivité 14](#_Toc126952075)

[2. Analyse quantitative pour le 12 décembre 2021 (ACP) 15](#_Toc126952076)

[a. Relation entre le taux d’incidence et le pourcentage de couverture vaccinale complète 15](#_Toc126952077)

[b. Relation entre le taux de positivité et le pourcentage de couverture vaccinale complète 16](#_Toc126952078)

[c. Aperçu sur les associations linéaires entre les variables quantitatives 17](#_Toc126952079)

[d. Analyse en composantes principales (ACP), Analyse de la variance expliquée par les dimensions suggérées 18](#_Toc126952080)

[e. ACP, Corrélation entre les variables et les dimensions d’intérêt 19](#_Toc126952081)

[f. ACP, Contribution des variables aux composantes principales 20](#_Toc126952082)

[g. ACP, Qualité de représentation sur les composantes 21](#_Toc126952083)

[h. Conclusions par rapport aux contributions et qualités de représentations 22](#_Toc126952084)

[i. ACP, Interprétation et définition des dimensions 1, 2 et 3 22](#_Toc126952085)

[j. ACP, Représentation des individus sur les dimensions 1, 2 et leurs regroupements 26](#_Toc126952086)

[3. Analyse pour le 10 avril 2021, statistiques descriptives par groupe de restrictions sanitaires 27](#_Toc126952087)

[4. Code implémenté 29](#_Toc126952088)

[a. Code python sous JupiterLab 29](#_Toc126952089)

[b. Code R sous Rstudio 38](#_Toc126952090)

[Références bibliographiques 42](#_Toc126952091)

**Table des figures**

[Figure 1. Evolution du COVID-19 sur l'année 2021 selon les régions 5](#_Toc127224430)

[Figure 2. Evolution du COVID-19 sur l'année 2021 selon les départements 7](#_Toc127224431)

[Figure 3. Effet de la couverture vaccinale sur le taux d’incidence et différence entre les départements 8](file:///C:\Users\khaou\Documents\Master%20SEP\5-%20Analyse%20des%20données%20et%20data%20mining_GUATHERAT\projet%20individuel\Deliverables\Rapport_Propagation%20du%20covid4.docx#_Toc127224432)

[Figure 4. Représentation des départements selon les régions 9](#_Toc127224433)

[Figure 5. Effet de la mise en place des restrictions sanitaires renforcées du 19 mars et 25 mars et sur le taux d’incidence 10](file:///C:\Users\khaou\Documents\Master%20SEP\5-%20Analyse%20des%20données%20et%20data%20mining_GUATHERAT\projet%20individuel\Deliverables\Rapport_Propagation%20du%20covid4.docx#_Toc127224434)

[Figure 6. Distribution globale et dispersion du taux d'incidence 13](#_Toc127224435)

[Figure 7. Distribution du taux de positivité au cours de l'année 2021 14](#_Toc127224436)

[Figure 8. Diagramme de dispersion des variables Taux d'incidence et pourcentage de couverture vaccinale complète le 12 décembre 2021 15](#_Toc127224437)

[Figure 9. Diagramme de dispersion des variables Taux de positivité et pourcentage de couverture vaccinale complète le 12 décembre 2021 16](#_Toc127224438)

[Figure 10. Matrice de corrélations entre les variables quantitatives d'intérêt 17](#_Toc127224439)

[Figure 11. Répartition de la variance expliquée 18](#_Toc127224440)

[Figure 12. Représentation des variables sur le cercle des dimensions 1 et 2 23](#_Toc127224441)

[Figure 13. Représentation des variables sur le cercle des dimensions 1 et 3 25](#_Toc127224442)

[Figure 14.Représentation des départements sur les dimensions 1 et 2 selon les régions 26](#_Toc127224443)

**Liste des tableaux**

[Tableau 1. Corrélation des variables avec les composantes principales 19](#_Toc127224444)

[Tableau 2. Contributions des variables aux composantes principales 20](#_Toc127224445)

[Tableau 3. Qualité de représentation des variables 21](#_Toc127224446)

[Tableau 4. Première dimension : meilleurs contributions et qualités de représentation 22](#_Toc127224447)

[Tableau 5. Deuxième dimension : meilleurs contributions et qualités de représentation 22](#_Toc127224448)

[Tableau 6. Statistiques descriptives des indicateurs de propagation avec et sans restrictions sanitaires renforcées 27](#_Toc127224449)

# Introduction

En France, la pandémie de COVID-19 a eu un impact significatif sur la santé de la population et sur l'économie du pays. Comprendre les facteurs qui influencent la propagation du virus à l'échelle départementale peut aider à identifier les zones les plus touchées et à mettre en place des stratégies efficaces pour endiguer la transmission.

Ce rapport a pour but de décortiquer les facteurs qui ont une influence sur la propagation du COVID-19 en France à l'échelle départementale, en se focalisant sur les caractéristiques démographiques, les taux de couverture vaccinale et les mesures gouvernementales. Il examinera également les tendances et les défis liés à cette épidémie.

Afin de mieux comprendre les raisons qui ont favorisé la propagation du COVID-19 dans certaines régions plutôt que dans d'autres, des observations ont été effectuées à différents moments clés de l'évolution de la pandémie en France en 2021. Cela permettra d'éclairer les facteurs qui ont contribué à la propagation du virus à des moments particuliers de l'épidémie et de comprendre leurs tendances.

Nous avons basé notre analyse de la propagation du COVID-19 sur les données démographiques fournies par l'INSEE en 2021, c'est pourquoi nous ne considérons que les données de cette année pour notre étude.

Afin de déterminer les liens entre les différentes variables liées à la pandémie, il a été proposé de mener des études transversales pour repérer les tendances dans une population et identifier les facteurs de risque potentiels pour une région ou un département donné.

Pour comprendre de manière plus complète l'impact des vaccinations sur la propagation du COVID-19, des observations ont été réalisées vers la fin de l'année 2021, après l'ouverture de la vaccination à différentes tranches d'âge de la population. Cela permettra d'évaluer si les vaccins ont eu un effet sur la diminution de la propagation du virus dans les zones les plus touchées, et si la stratégie de vaccination a été efficace pour limiter la transmission du virus.

Des observations ont été réalisées après la mise en place des restrictions sanitaires renforcées dans certains départements afin de comprendre l'impact des mesures gouvernementales sur la propagation de la COVID-19. Cela correspond à une période de 20 jours ou moins, selon le département, après la mise en action de restrictions renforcées pour 16 départements en 19 mars 2021 et l'ajout de 3 autres départements en 25 mars 2021. Cela nous permettra de déterminer si les mesures gouvernementales ont eu un impact spécifique sur la limitation de la propagation du virus dans certaines régions ou départements.

Les informations utilisées pour cette étude ont été principalement récoltées auprès de trois sources : le ministère des Solidarités et de la Santé, La Santé Publique France et l’INSEE. Outre les informations démographiques récoltées ou produites pour l'année 2021, il s'agit de données quotidiennes sur les variables liées aux thèmes en question.

Notre étude s'est concentrée sur les départements de la France métropolitaine, notre population cible se compose donc de ses 96 départements.

Après un travail de préparation des données, nous disposons de trois sources pour l'analyse : une première qui présente un suivi des départements sur l'année 2021 et qui permet également de visualiser la chronologie temporelle de la propagation et de repérer les moments clés à étudier. Les deux autres sources correspondent chacune à une date précise et fournissent des données départementales.

Les variables qui nous intéressent sont les suivants :

* Le taux d’incidence dans le département : un indicateur de l’activité épidémique, correspond au nombre de personnes testées positives (RT-PCR et test antigénique) pour la première fois depuis plus de 60 jours rapporté à la taille de la population. Il est exprimé pour 100 000 habitants.
* Le taux de positivité des test virologiques dans le département : correspond au nombre de personnes testées positives (RT-PCR et test antigénique) pour la première fois depuis plus de 60 jours rapporté au nombre total de personnes testées positives ou négatives sur une période donnée ; et qui n‘ont jamais été testées positives dans les 60 jours précédents.
* Le taux d’occupation hospitalière dans le département : l s’agit de la proportion de patients atteints de COVID-19 en réanimation, en soins intensifs, ou en unité de surveillance continue rapportée au total des lits de réanimation.
* Le taux de couverture vaccinale dans le département : nombre de personnes complètement primovaccinées par population.  Les personnes avec une primo-vaccination complète sont notamment : personnes vaccinées par deux doses par les vaccins nécessitant deux doses (ex : vaccins Pfizer, Moderna ou Astra-Zeneca), personnes vaccinées par une dose par les vaccins nécessitant une seule dose (ex : vaccin Janssen), personnes vaccinées par une dose par les vaccins Pfizer, Moderna ou Astra-Zeneca en cas d’antécédent de COVID-19, personnes immunodéprimées pouvant recevoir 3 à 4 doses.
* Le taux de couverture du rappel vaccinal dans le département : Couverture vaccinale dans la population du 1er rappel du vaccin contre la COVID-19, (i.e. couverture primo-vaccination COVID19 complète et avec une dose de rappel)
* La densité de la population dans le département
* La part des femmes et des hommes de la population dans le département
* La part des différentes classes d’âges dans la population du département (trois classes : moins de 24 ans, de 25 à 59 ans et plus de 60 ans)

# Evolution du COVID-19 au cours de l'année et selon les zones

En premier lieu, notre objectif était d'analyser la répartition globale des indicateurs de propagation du COVID-19 au cours de l'année, évaluer leur degré de dispersion, identifier les tendances et repérer les évolutions clés à différents moments dans l'ensemble de la population cible.

Une analyse de l’activité épidémique au cours de l’année 2021 visible dans **l’**[**annexe 1**](#_Répartition_globale_et) a révélé une claire augmentation de l'activité épidémique observée en avril, suivie d'une diminution remarquable des cas positifs en septembre. Cependant, il est important de noter que ces tendances varient considérablement entre les départements, ce qui souligne la nécessité d'une analyse détaillée à l'échelle locale. Il est également remarquable que durant l'hiver, une forte augmentation des cas a été repérée, mais cette fois-ci avec une beaucoup moins grande variabilité des taux d'incidence entre les départements par rapport aux autres dates.

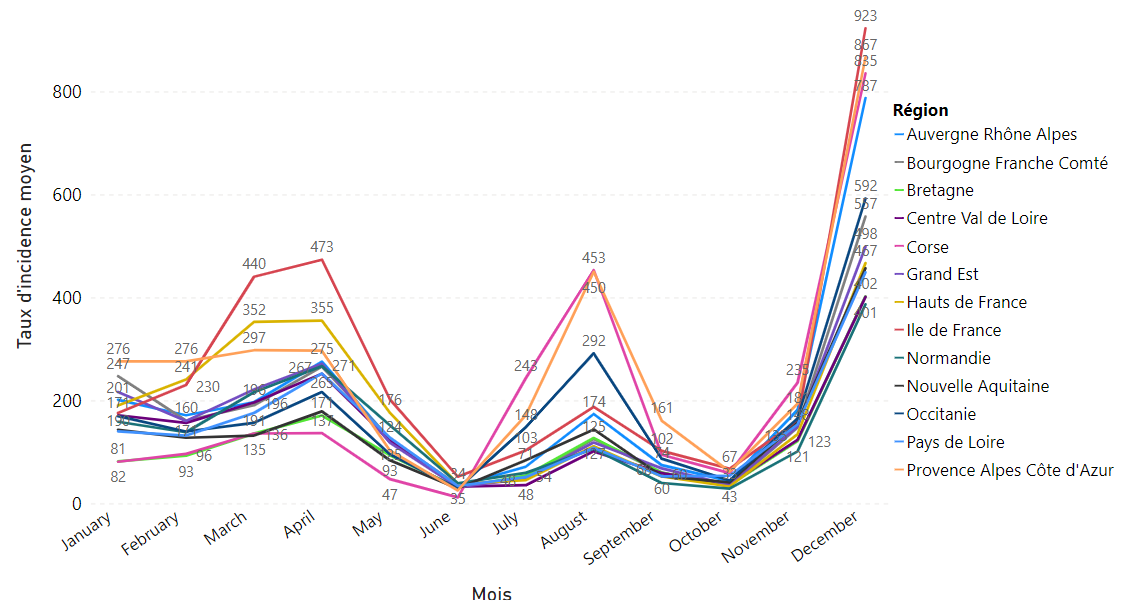
De plus une analyse équivalente de l’ensemble des valeurs des taux de positivités prises par les départements à des moments espacée dans l’année a montré une même tendance générale et variabilité. Toutefois il est à souligner qu’un taux de positivité élevé peut indiquer une forte propagation du virus dans une communauté, tandis qu'un taux de positivité faible peut indiquer que la plupart des tests sont effectués chez des personnes qui ont peu de chances d'être infectées. Il est important de noter que le taux de positivité ne donne pas une image complète de la situation épidémiologique. Il est donc important de le considérer en combinaison avec d'autres indicateurs comme le taux d’incidence pour obtenir une image complète de l'évolution de la pandémie.

Ces résultats montrent l'importance de continuer à surveiller de près l'évolution de la pandémie à différents niveaux géographiques

Ensuite, nous voulons creuser plus profondément pour repérer les variations à l’échelle départementale et régionale et identifier les moments où ces différences sont les plus marquées. Cela nous permettra, dans l’étape suivante, d’effectuer une analyse transversale dans le temps et de déterminer les facteurs contributifs à ces changements. Cette analyse détaillée à différents niveaux géographiques nous aidera à mieux comprendre les facteurs favorisant la propagation du COVID-19.

A l’échelle régionale

Figure 1. Evolution du COVID-19 sur l'année 2021 selon les régions



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : En moyenne 473 nouveaux cas pour chaque 100 000 personnes ont été repéré en avril pour l’ensemble des départements de la région Ile-de-France.

En général, pour toutes les régions, l'année 2021 a été marquée par une hausse des cas de COVID-19 au début de l'année, suivie d'un pic au printemps. Cependant, on a observé une diminution des cas au début de l'été, avant une nouvelle augmentation en août. Cette tendance s'est poursuivie avec une diminution à l'automne, suivie d'une nouvelle hausse des cas en hiver, liée à l'apparition de nouvelles variantes du virus.

Il est également important de noter que les taux d'incidence varient considérablement selon les régions et selon les périodes considérées.

Le pic observé en avril est visible à l'échelle régionale à des niveaux clairement différents. Cela suggère une forte augmentation de la transmission du virus dans certaines régions par rapport à d'autres.

Certaines régions, comme les Hauts-de-France et l'Ile-de-France, ont connu une augmentation marquée de l'activité épidémique depuis mars, atteignant également les pics les plus élevé du mois d’avril. Cependant, ces régions ont réussi à réduire considérablement le nombre de nouveaux cas, à un niveau similaire à celui des autres régions. Cela peut être attribué aux mesures sanitaires renforcées mises en place par le gouvernement dans ces régions, notamment à partir du 20 mars.

Au fait, tout au long de la pandémie, dans certaines régions, les autorités ont pris des mesures strictes pour limiter la propagation du virus, comme des verrouillages et des restrictions de voyage, ce qui a entraîné une baisse des cas. D'autres régions ont eu des taux de transmission plus élevés et ont dû mettre en place des mesures plus strictes pour endiguer la propagation

Ces mesures auraient probablement contribué à la réduction de la transmission du virus dans ces régions.

Dans les analyses qui suivent nous allons investiguer l’influence de ces mesures, tout particulièrement, après 20 jours de la mise en place des restrictions du 19 mars.

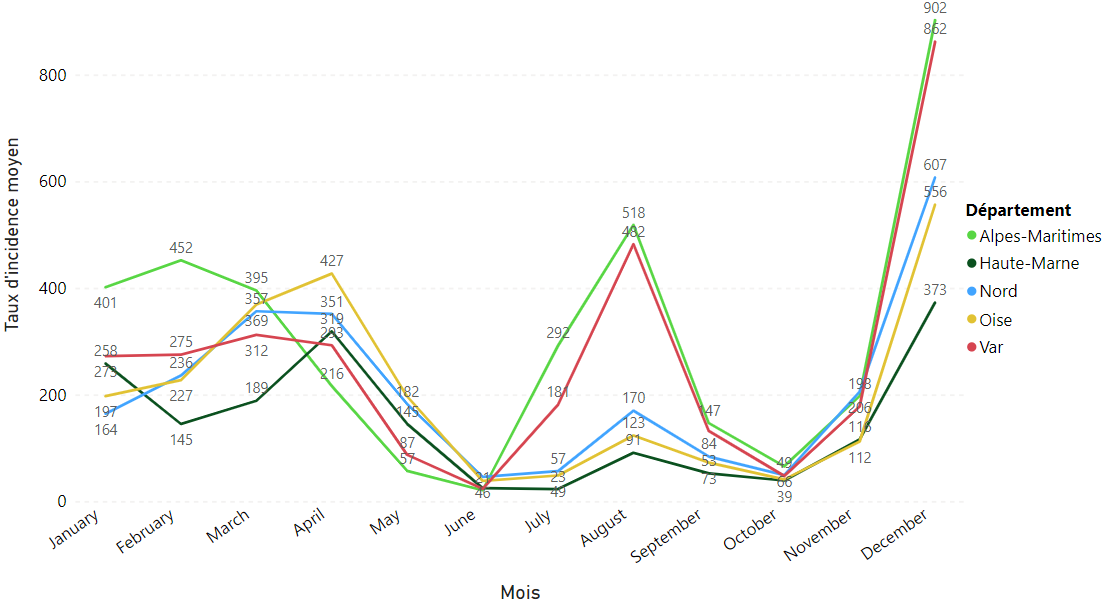
Notons que ces différences entre régions, sont clairement fidèles aux résultats des analyses précédentes et expliquent notamment la variabilité introduite par ces derniers.

Les données agrégées à l'échelle régionale permettent de voir les tendances générales, du fait qu’elles regroupent plusieurs départements chacune. Cependant, les analyses à l'échelle départementale peuvent révéler des nuances entre ces derniers.

Visualisation d'un ensemble de départements tirés au hasard, représentants de leurs régions ? Homogénéité ou nuances **?**

Notre étude comprend 96 départements. Pour une meilleure visualisation de l'évolution de l'épidémie dans ces départements, nous avons choisi de représenter un échantillon aléatoire. Le but était de vérifier si l'observation macroscopique précédente a pu vraiment décrire ce qui s'est passé à l'échelle microscopique ; au niveau des départements.

Figure 2. Evolution du COVID-19 sur l'année 2021 selon les départements



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 de data.gouv.fr.

**Cohorte** : Un échantillon aléatoire des départements de la France Métropolitaine.

**Lecture** : En moyenne 452 nouveaux cas pour chaque 100 000 personnes ont été repéré en février pour le département Alpes-Maritimes.

Nous constatons une réflexion quasi similaire à l’échelle départementale des tendances observées à l'échelle régionale de l'évolution des taux d’incidence, mais il existe des différences notables entre les départements d'une même région. Par exemple, les départements du Var et des Alpes-Maritimes, qui font partie de la région Provence-Alpes-Côte d'Azur, ont connu des évolutions contrastées entre février et mars. Le département Var a enregistré une augmentation de 13% au nombre de cas, atteignant en moyenne 312 nouveaux cas pour chaque 100 000 personnes en mars, tandis que les Alpes-Maritimes ont connu une diminution, enregistrant un taux d'incidence moyen de 395 pour chaque 100 000 personnes qui correspond à 12% moins que la valeur enregistrée en février.

Afin de comprendre les facteurs qui peuvent causer une propagation plus importante du COVID-19 dans certains départements par rapport à d'autres, et pour évaluer si ces facteurs ont des impacts différents selon les départements, nous allons étudier deux points clés dans le temps. Cela nous permettra de déterminer les facteurs et d'analyser leur influence sur la propagation de la pandémie.

# Etudier deux moments clés dans le temps

L'objectif de cette section est de révéler les facteurs déterminant les différences repérées précédemment entre les départements et d'évaluer leur influence.

#### Le 12 décembre, une analyse pour évaluer l'impact de la couverture vaccinale et des caractéristiques démographiques

Figure 3. Effet de la couverture vaccinale sur le taux d’incidence et différence entre les départements

Une image contenant carte

Description générée automatiquementUne image contenant carte

Description générée automatiquement

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et de la couverture vaccinale de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Le 12 décembre 2021, le taux d'incidence en Ain est relativement élevé par rapport aux autres départements. Cependant, le pourcentage de couverture vaccinale complète est relativement faible dans ce département. À la Vendée, une grande partie de la population a été complètement vaccinée.

En observant ces deux cartes, on peut déduire qu'il existe un lien entre le taux de couverture vaccinale complète d'un département et son taux d'incidence du COVID-19. Les départements avec les taux de couverture vaccinale les plus élevés affichent généralement les taux d'incidence les plus bas.

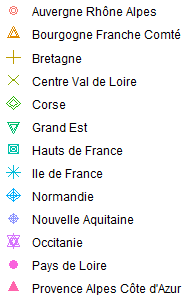
Le 12 décembre 2021, les départements du sud-est ont été particulièrement touchés par la pandémie, enregistrant les taux d'incidence les plus élevés et les taux de couverture vaccinale les plus faibles. Contrairement aux départements du nord-ouest qui ont enregistré les plus fortes proportions de couverture vaccinale à cette date ayant en revanche la plus faible activité épidémique.

Cette potentielle relation entre les taux de couverture vaccinale et l'activité épidémique a été étudiée de manière plus approfondie grâce à une analyse quantitative détaillée dans l'[annexe 2](#_1_ère_analyse). Cette analyse nous a permis de mieux comprendre l'influence des campagnes de vaccination, ainsi que d'autres variables, sur la propagation de la pandémie dans les différents départements.

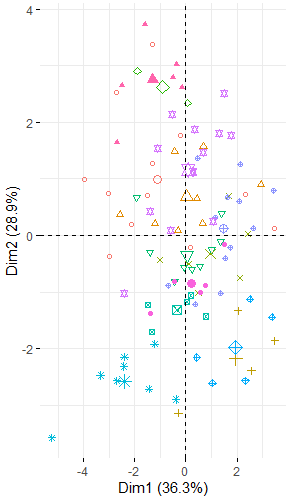
Nous en avons pu résumer les résultats dans ce qui suit.

Figure 4. Représentation des départements selon les régions

Région



**Taux d’occupation hospitalière, de positivité et d’incidence élevés**



**Couverture Vaccinale et de Rappel vaccinal faibles, population jeune et taux d’incidence élevé**

**Couverture Vaccinale et de Rappel vaccinal élevées, population âgée et taux d’incidence faible**

**Taux d’occupation hospitalière, de positivité et d’incidence faibles**

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Les départements d’île de France ont des couvertures vaccinales faible et des populations jeunes.

Ce graphe met en évidence l’influence de la couverture vaccinale et l’âge de la population sur la propagation du COVID-19.

On voit clairement que les départements de l'Île-de-France, ayant une population plus jeune, ont une couverture vaccinale et un taux de rappel vaccinal plus faibles par rapport aux autres départements. Les départements de la région Occitanie, avec une population plus âgée, sont parmi les plus impactés par les campagnes de vaccinations et de rappel vaccinal.

Les départements de Normandie et de Bretagne, avec une population moins jeune, sont également touchés par les campagnes de vaccinations.

Les départements des régions de sud de Provence-Alpes-Côte d'Azur et d’Auvergne-Rhône-Alpes semblent avoir les taux d'occupation hospitalières les plus élevés, particulièrement pour le second département qui affiche également des taux d'incidence et de positivité élevés. Les départements du Grand Est ont un comportement proche de la moyenne.

#### Le 10 avril, une analyse pour évaluer l'impact des mesures sanitaires renforcées et des caractéristiques démographiques

Figure 5. Effet de la mise en place des restrictions sanitaires renforcées du 19 mars et 25 mars et sur le taux d’incidence

Une image contenant carte

Description générée automatiquementUne image contenant carte

Description générée automatiquement

Restrictions sanitaires

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Le 10 avril 2021, malgré la mise en place des restrictions sanitaire depuis plus de 20 jours, le taux d'incidence en Oise est relativement élevé par rapport aux autres départements.

Nous rappelons que le 19 mars 2021, il y a eu une mise en place de restrictions sanitaires renforcées dans les départements à forte activité épidémique.

Nous pouvons clairement constater que, bien qu'il y ait déjà 20 jours depuis la mise en pratique des mesures restrictives, on observe que les départements en question n'ont pas nécessairement des situations sanitaires optimales.

Nous pouvons également observer que les mesures prises pour contrôler la situation épidémique varient en efficacité selon les départements. Par exemple, Alors que la Somme présente un taux d'incidence relativement faible, celui de l'Oise est relativement élevé Malgré les mesures sanitaires communes imposées le 19 mars.

Cela suggère que les mêmes mesures peuvent donner des résultats différents selon les départements. Une analyse approfondie est nécessaire pour valider ces conclusions.

Une analyse descriptive a été effectuée pour comparer les deux ensembles de départements, ceux qui ont subi les restrictions sanitaires renforcées du 19 mars jusqu’au 10 avril et ceux qui n'ont pas été concernés.

Des différences significatives ont été constatées. Cependant, malgré l'application de ces mesures, la situation de la propagation du COVID-19 dans ces départements reste toujours préoccupante. Cela peut suggérer que les mesures sanitaires renforcées dans ces départements n'ont pas été suffisamment efficaces pour réduire significativement la propagation. Il est important de rappeler que les départements concernés se situent principalement dans les régions d'Île-de-France, de Hauts-de-France et de Normandie.

On peut penser que le manque d'application et de surveillance des restrictions aient pu jouer un rôle dans ce résultat, ou encore que les départements concernés présentent des caractéristiques particulières, telles qu'une population plus jeune qui n'applique pas systématiquement les mesures.

Cela montre que d'autres facteurs peuvent également influencer les taux de propagation du COVID-19, et que les restrictions sanitaires ne sont peut-être pas les seuls facteurs à prendre en compte pour évaluer leur efficacité.

# Conclusion

La pandémie de COVID-19 a été un événement mondial qui a mis à l'épreuve l'état de préparation du gouvernement face à un virus hautement infectieux. Nous avons étudié, dans le cadre de ce projet, les principaux facteurs ayant contribué à la propagation rapide au niveau régional et départemental, ainsi que l'efficacité des procédures gouvernementales pour faire face à la menace du COVID-19.

Une étude sur l’année 2021 sur les indicateurs de propagation a montré un comportement similaire entre les régions d'un point de vue de l’évolution dans le temps, avec des pics de taux d'incidence pour la plupart de la région au printemps, un second pic à la fin de l'été, et un pic énorme l'hiver suivant avec l'apparition des nouvelles variantes du virus. D'un autre côté, nous avons remarqué des différences entre les régions, et donc entre les différents départements, tel que certaines régions ont atteint un taux d'incidence élevé tandis que d'autres ont maintenu des niveaux contrôlables.

Nous avons approfondi les moments critiques de la propagation du COVID-19 en France, en deux sections. Dans la première, nous avons constaté que la couverture vaccinale et l'âge de la population exercent une influence significative sur la propagation du virus. Les départements avec une population plus jeune ont montré un taux de couverture vaccinale plus faible et une plus grande propagation du COVID-19 que les départements avec une population plus âgée, qui ont été plus impactés par les campagnes de vaccination. La deuxième section s'est concentrée sur les restrictions sanitaires imposées par le gouvernement. Nous avons constaté que ces mesures n'ont pas été suffisamment efficaces pour réduire considérablement la propagation du COVID-19 dans les départements concernés. Cela peut être dû au fait que ces départements ont une population plus jeune qui ne respecte pas systématiquement les mesures sanitaires, ou que d'autres facteurs, tels que la position géographique, peuvent également influencer la propagation du virus.

En conclusion, notre analyse suggère que les caractéristiques sociodémographiques des communautés et la couverture vaccinale ont un impact significatif sur la propagation du COVID-19 dans les différents départements.

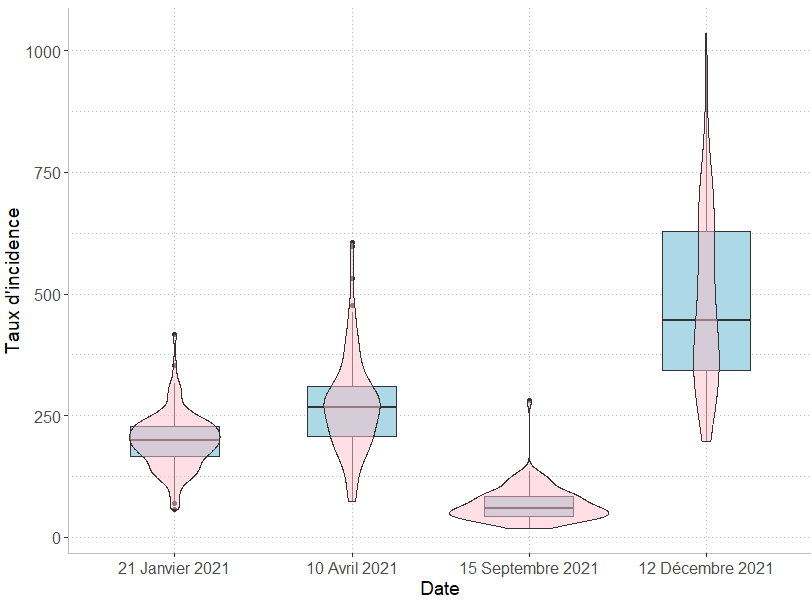
Il est important de mentionner que autres facteurs peuvent également influencer la propagation, et que les restrictions sanitaires et les couvertures vaccinale et l’âge de la population ne sont peut-être pas les seuls facteurs à prendre en compte pour évaluer complètement les facteurs impliqués dans la propagation du COVID-19.

# Annexes

## Répartition globale et dispersion des indicateurs de la propagation du COVID-19

### Taux d’incidence

Figure . Distribution globale et dispersion du taux d'incidence



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Le 10 avril 2021, la médiane des taux d'incidence des départements en France métropolitaine était de 265. Comme cette médiane se situait presque au milieu de la boîte, cela signifie que la plupart des départements de la France métropolitaine ont enregistré des valeurs proches de 265 nouveaux cas pour chaque 100 000 personnes dans les 60 jours précédents.

Le premier quantile pour cette même date est de 207 et le troisième quantile est de 310.

4 valeurs aberrantes sont repères pour cette date avec des taux d’incidences supérieurs à 400.

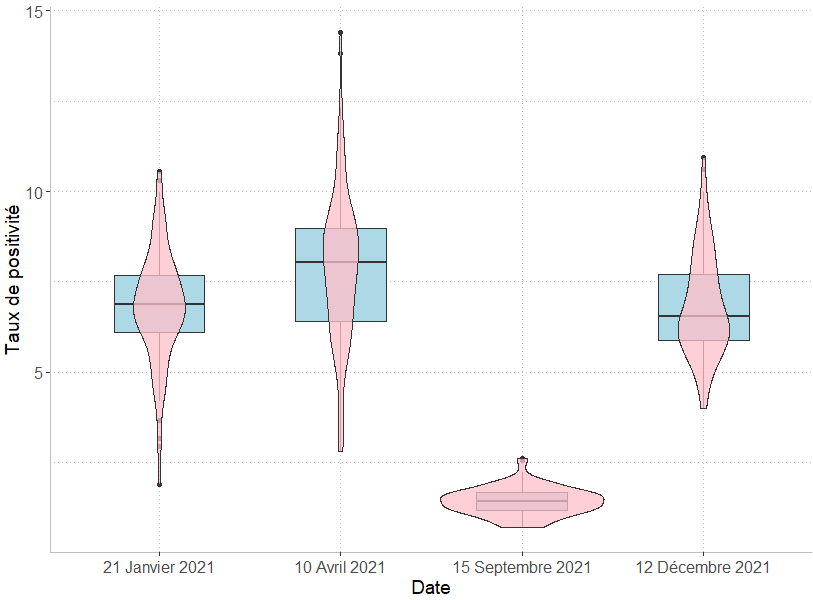
Les valeurs en 15 septembre 2021 présentent une distribution beaucoup plus allongée par rapport aux autres dates. Cela indique une variabilité remarquable des valeurs par rapport aux autres dates.

Cette visualisation combinée nous permet de suivre l'évolution de la distribution et de la dispersion des taux d'incidence dans les départements de France métropolitaine au cours du temps. Les variations des plages de valeurs et des médianes indiquent des différences de l’ensemble des valeurs des taux d'incidence des départements sur l’année. Les différences de distributions observées entre les différentes dates étudiées suggèrent également des variations des tendances sur l’année.

On observe une augmentation de l'activité épidémique en avril, suivie d'une diminution remarquable des cas positifs en septembre, avec une grande variabilité entre les départements. Une forte augmentation est repérée pour l'hiver avec une beaucoup moins grande variabilité des taux d’incidences entre les départements par rapport aux autres dates.

### Taux de positivité

Figure 7. Distribution du taux de positivité au cours de l'année 2021



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Le 21 janvier 2021, la médiane des taux de positivité des départements en France métropolitaine était de 6.8 %. Comme cette médiane se situait au milieu de la boîte, cela signifie que la plupart des départements de la France métropolitaine ont enregistré des valeurs proches de 6% sur les 60 jours précédents. C’est-à-dire que pour la plupart des départements, sur 100 personnes testées, 6 sont infectées.

Le premier quantile pour cette même date est de 6% et le troisième quantile est de 7.7%.

6 valeurs aberrantes sont repères pour cette date avec des taux.

Les valeurs en 12 décembre 2021 présentent une distribution beaucoup moins étroite des taux de positivité par rapport aux autres dates. Ce qui montre beaucoup moins de variabilités dans les données par rapport aux autres dates.

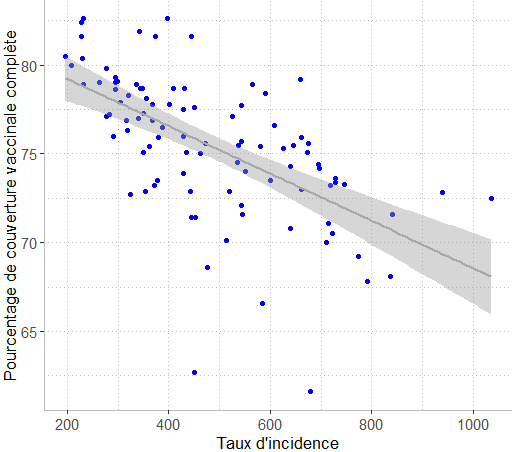
Cette visualisation combinée nous permet de suivre l'évolution de la distribution et de la dispersion des taux de positivité dans les départements de France métropolitaine au cours du temps. Les variations des plages de valeurs et des médianes indiquent des changements dans les taux d'incidence dans ces départements. Les différences de distributions observées entre les différentes dates étudiées suggèrent également des variations des tendances.

On observe une augmentation de taux de positivité en avril, suivie d'une diminution remarquable des en septembre, avec beaucoup plus de variabilité des données pour cette date. Une forte augmentation est enregistrée pour l'hiver.

## Analyse quantitative pour le 12 décembre 2021 (ACP)

### Relation entre le taux d’incidence et le pourcentage de couverture vaccinale complète

Figure 8. Diagramme de dispersion des variables Taux d'incidence et pourcentage de couverture vaccinale complète le 12 décembre 2021



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

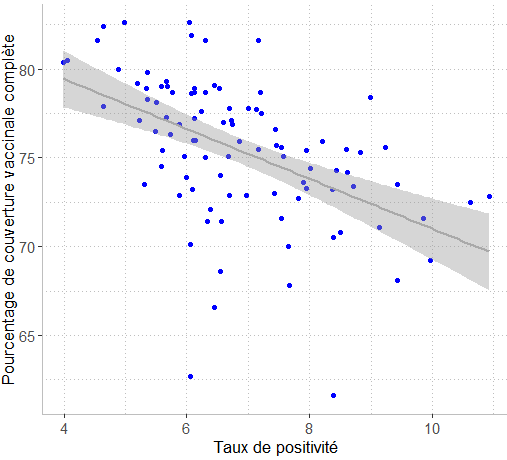
**Lecture** : Le 12 décembre 2021, il a été observé que le département qui a enregistré le taux d’incidence le plus faible (soit 200 nouveaux cas pour chaque 100 000 personnes) a également eu un pourcentage de couverture vaccinale particulièrement élevé, dépassant les 80%.

La forme générale de la distribution de la majorité des points est linéaire de direction négative.

En examinant le nuage des points, on peut observer que la plupart des points sont regroupés autour d'une ligne de tendance descendante. La densité des points est également plus élevée à proximité de cette ligne, ce qui indique une corrélation linéaire négative forte entre le taux d'incidence et le pourcentage de couverture vaccinale complète. Cependant, il est important de noter que certaines valeurs aberrantes sont également présentes, qui se situent à l'écart de la tendance générale. Ces valeurs peuvent correspondre à des départements spécifiques ayant des caractéristiques distinctes des autres, et qui peuvent altérer la relation globale. Il est donc important de vérifier cette relation en utilisant des tests statistiques appropriés pour évaluer la significativité de cette corrélation et pour comprendre les causes de cette relation observée.

### Relation entre le taux de positivité et le pourcentage de couverture vaccinale complète

Figure 9. Diagramme de dispersion des variables Taux de positivité et pourcentage de couverture vaccinale complète le 12 décembre 2021



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr.

**Cohorte** : Les départements de la France Métropolitaine.

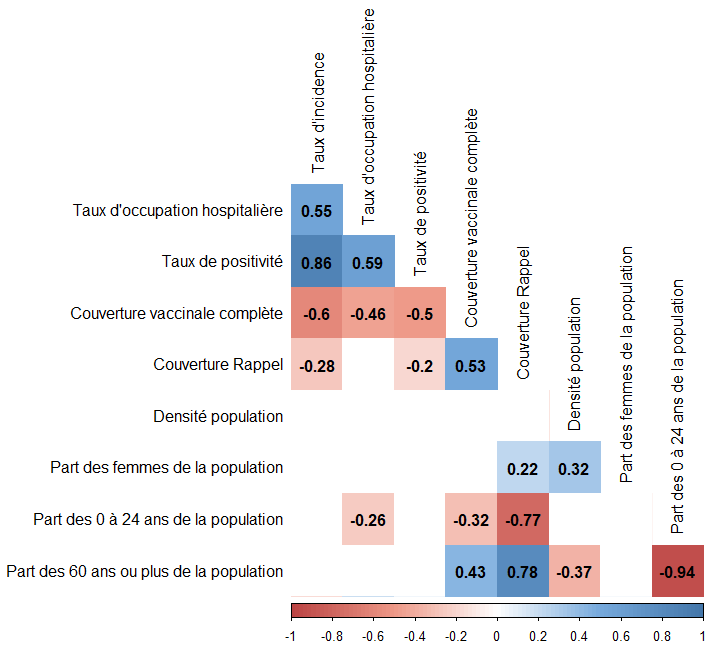
**Lecture** : Le 12 décembre 2021, il a été observé que le département qui a enregistré le taux de positivité le plus fort (soit 11%) a également eu un pourcentage de couverture vaccinale relativement moyen, ne dépassant pas les 75%.

La forme générale de la distribution de la majorité des points est linéaire de direction négative.

En analysant le nuage de points, on peut constater que la plupart des points sont regroupés autour d'une tendance linéaire décroissante. La densité des points est également plus élevée à proximité de cette tendance, ce qui suggère une forte corrélation négative entre le taux de positivité et le pourcentage de couverture vaccinale complète. Cependant, il est important de noter qu'il y a également des valeurs aberrantes qui se situent en dehors de cette tendance. Ces valeurs peuvent être liées à des départements ayant des caractéristiques particulières qui peuvent affecter la relation globale. Il est donc essentiel de vérifier cette relation en utilisant des tests statistiques appropriés pour évaluer la signification de cette corrélation et pour comprendre les causes de cette relation observée.

### Aperçu sur les associations linéaires entre les variables quantitatives

Figure . Matrice de corrélations entre les variables quantitatives d'intérêt



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Le coefficient de corrélation entre les variables « Couverture Rappel » et « Part des 60 ans ou plus de la population » est de 0.78.

Les coefficients présentés ci-dessus indiquent des résultats statistiquement significatifs à un niveau de confiance de 95%. Les cases vides indiquent que les coefficients calculés pour ces relations ne sont pas statistiquement significatifs à un niveau de confiance de 95%.

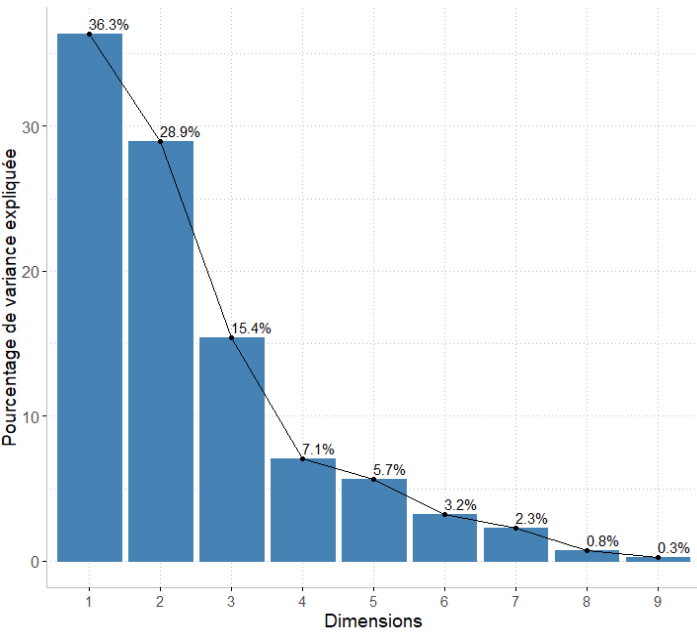
L’analyse de la matrice de corrélations révèle une corrélation positive entre la couverture de rappel et la proportion de personnes de 60 ans ou plus dans la population (0.78). Il est également possible de constater une corrélation négative forte entre la couverture de rappel et la proportion de personnes âgées de 0 à 24 ans dans la population (-0.77). Cela indique que le rappel a touché une grande partie de la population âgée, mais n'a pas concerné les plus jeunes, le 21 décembre 2021.

Des corrélations négatives relativement forte sont visibles entre la couverture vaccinale et les variables taux d’incidence (-0.6), taux de positivité (-0.5) et taux d’occupation hospitalière (-0.46), ce qui semble indiquer une relation directe entre la couverture vaccinale et la réduction de la propagation du virus.

Aucune relation significative n'a été observée entre les variables démographiques (âge et sexe) et les indicateurs de propagation du virus. Il serait nécessaire d'effectuer une analyse qualitative plus approfondie pour comprendre davantage ces relations potentielles.

### Analyse en composantes principales (ACP), Analyse de la variance expliquée par les dimensions suggérées

Figure 11. Répartition de la variance expliquée



**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La première dimension explique 36.3% de la variabilité des données relatives à la propagation du virus, aux vaccinations et aux caractéristiques démographiques des départements.

36,3% de la variabilité des données relatives à la propagation du virus, aux vaccinations et aux caractéristiques démographiques des départements de la France métropolitaine au 12 décembre 2021 est expliquée par la première dimension de l'analyse en composantes principales, près de 29% par la deuxième et 15,4% par la troisième. Ces composantes contiennent l'essentiel de l'information importante des données, représentant plus de 80% de la variance totale. Dans les étapes suivantes, il est crucial de considérer l'interprétation de ces composantes en termes des variables d'origine. Ensuite, nous allons utiliser ces dimensions pour visualiser la répartition des individus dans leurs espaces respectifs. Cela nous permettra de mieux comprendre les relations entre les individus et les variables d'origine, ainsi que d'identifier les groupes ou tendances présentes dans les données.

### ACP, Corrélation entre les variables et les dimensions d’intérêt

Tableau 1. Corrélation des variables avec les composantes principales

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Dimension 1 | Dimension 2 | Dimension 3 |
| Taux d'incidence | -0.66 | 0.61 | -0.3 |
| Taux d'occupation hospitalière | -0.29 | 0.78 | -0.02 |
| Taux de positivité | -0.52 | 0.75 | 0.02 |
| Couverture vaccinale complète | 0.80 | -0.28 | -0.06 |
| Couverture de rappel vaccinal | 0.84 | 0.33 | -0.19 |
| Densité population | -0.21 | -0.35 | -0.71 |
| Part des femmes de la population | 0.12 | -0.02 | -0.86 |
| Part des 0 à 24 ans de la population | -0.69 | -0.64 | 0.09 |
| Part des 60 ans ou plus de la population | 0.77 | 0.59 | 0.04 |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La variable "Couverture Rappel" est fortement corrélée positivement avec la première dimension, avec un coefficient de corrélation de 0.84.

La forte corrélation positive entre la première dimension et la couverture vaccinale complète (0.8) et celle du rappel vaccinal (0.84) indique que cette dimension explique une grande partie de la variation des proportions de la population qui ont reçu les vaccins et leurs rappels dans les départements. Cela peut suggérer que cette dimension est liée aux facteurs qui influencent la couverture du rappel vaccinal, tels que l'adhésion à la vaccination, les campagnes de sensibilisation et les barrières à la vaccination (pouvant inclure l’accès limité aux centre de vaccination (par exemple pour les habitants des zones rurales ou éloignées), les barrières psychologiques (tel que les personnes ayant des croyances anti-vaccination réticentes à se faire vacciner) et les barrières administratives (comme les exigences de preuve d'identité ou de statut de santé).

Il est également intéressant de noter qu'il y a une forte corrélation (0.77) entre cette même dimension et la part des personnes âgées de 60 ans ou plus dans la population, ce qui pourrait indiquer que cette population est particulièrement touchée par les campagnes de vaccination.

De l'autre côté, il y a une corrélation négative entre la dimension 1 et le taux d'incidence avec un coefficient assez élevé de -0.66 ainsi qu'avec le taux de positivité (-0.52). Cela pourrait indiquer que cette dimension est liée aux facteurs qui réduisent la transmission du virus et la maladie, tels que les mesures de distanciation sociale et les campagnes de sensibilisation. Il est à souligner aussi que cette dimension est également fortement négativement corrélée avec la proportion de la population jeune âgée de 24 ans ou moins. Cela indique que cette proportion de la population peut avoir un impact négatif sur la couverture vaccinale. Il est possible que cette tranche d'âge ne soit pas ciblée par les campagnes de vaccination ou qu'elle soit moins encline à se faire vacciner.

En revanche, la deuxième dimension est fortement corrélée positivement avec les indicateurs de propagation du virus, avec une force et une direction de corrélation similaires à celles de la première dimension avec les variables exprimant les proportions des groupes d'âge de la population.

La troisième dimension présente de fortes corrélations négatives avec les variables précisant la densité de population du département et la proportion de femmes dans cette population.

### ACP, Contribution des variables aux composantes principales

Tableau 2. Contributions des variables aux composantes principales

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Dimension 1 | Dimension 2 | Dimension 3 |
| Taux d'incidence | 13.28 | 14.14 | 6.46 |
| Taux d'occupation hospitalière | 2.64 | 23.34 | 0.04 |
| Taux de positivité | 8.27 | 21.65 | 0.02 |
| Couverture vaccinale complète | 19.68 | 3.01 | 0.25 |
| Couverture de rappel vaccinal | 21.40 | 4.2 | 2.73 |
| Densité population | 1.3 | 4.74 | 36.65 |
| Part des femmes de la population | 0.44 | 0.02 | 53.06 |
| Part des 0 à 24 ans de la population | 14.69 | 15.48 | 0.64 |
| Part des 60 ans ou plus de la population | 18.31 | 13.42 | 0.14 |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La variable « Couverture Rappel » contribue à 21,4% à la première dimension.

Les variables ayant les contributions les plus élevées à la première dimension sont, dans l'ordre, la couverture de rappel vaccinal (21,4%), la couverture vaccinale complète (19,68%), la proportion des personnes âgées de 60 ans ou moins de la population (18,31%), la proportion de la population jeune (14,69%) et le taux d'incidence (13.28%). Ces variables sont les plus corrélées à cette dimension, selon la section précédente, ce qui soutient notre démarche d'analyse.

La deuxième dimension est fortement constituée par les taux d’incidence, de positivité et d’occupation hospitalière.

Quant à la troisième dimension, elle est exclusivement constituée par la densité de la population dans le département en question et la par des femme de cette population.

### ACP, Qualité de représentation sur les composantes

Tableau 3. Qualité de représentation des variables

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Dimension 1 | Dimension 2 | Dimension 3 |
| Taux d'incidence | 0.43 | 0.37 | 0.09 |
| Taux d'occupation hospitalière | 0.09 | 0.61 | 0 |
| Taux de positivité | 0.27 | 0.56 | 0 |
| Couverture vaccinale complète | 0.64 | 0.08 | 0 |
| Couverture Rappel | 0.7 | 0.11 | 0.04 |
| Densité population | 0.04 | 0.12 | 0.51 |
| Part des femmes de la population | 0.01 | 0 | 0.74 |
| Part des 0 à 24 ans de la population | 0.48 | 0.4 | 0.01 |
| Part des 60 ans ou plus de la population | 0.6 | 0.35 | 0 |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La couverture de rappel vaccinal est très bien représentée par la première de dimension avec un cos 2 proche de 1 (0.7).

Les variables les mieux représentées sur la première composante principale sont celles qui indiquent la proportion de la population ayant reçu la vaccination complète et celle ayant reçu la vaccination de rappel, ainsi que la proportion de la population âgée de 60 ans ou plus.

Les résultats obtenus sur la qualité de représentation des variables sur les dimensions de l'ACP sont en accord avec les relations déduites dans les sections précédentes. Les variables les plus contributives aux axes sont les mieux représentées.

En résumé, une variable est bien représentée sur une dimension de l'ACP, est également fortement liée à cette dimension et contribue donc significativement à l'explication de sa variance.

C’est dans cette optique qu’on va exploiter ces résultats pour expliciter la signification des dimensions obtenues et regrouper les départements en fonction de ces dimensions.

### Conclusions par rapport aux contributions et qualités de représentations

Tableau . Première dimension : meilleurs contributions et qualités de représentation

|  |  |
| --- | --- |
|  | Bonne à forte contribution |
| Qualité de représentation moyenne | * Taux d'incidence * Part des 0 à 24 ans de la population |
| Bonne qualité de représentation | * Couverture de rappel vaccinal * Couverture vaccinale complète * Part des 60 ans ou plus de la population |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : les couvertures vaccinales, les part des classes d’âge et le taux d’incidences définissent le premier axe.

Tableau . Deuxième dimension : meilleurs contributions et qualités de représentation

|  |  |
| --- | --- |
|  | Bonne à forte contribution |
| Moyenne à bonne Qualité de représentation | * Taux d'occupation hospitalière * Taux de positivité * Taux d'incidence |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

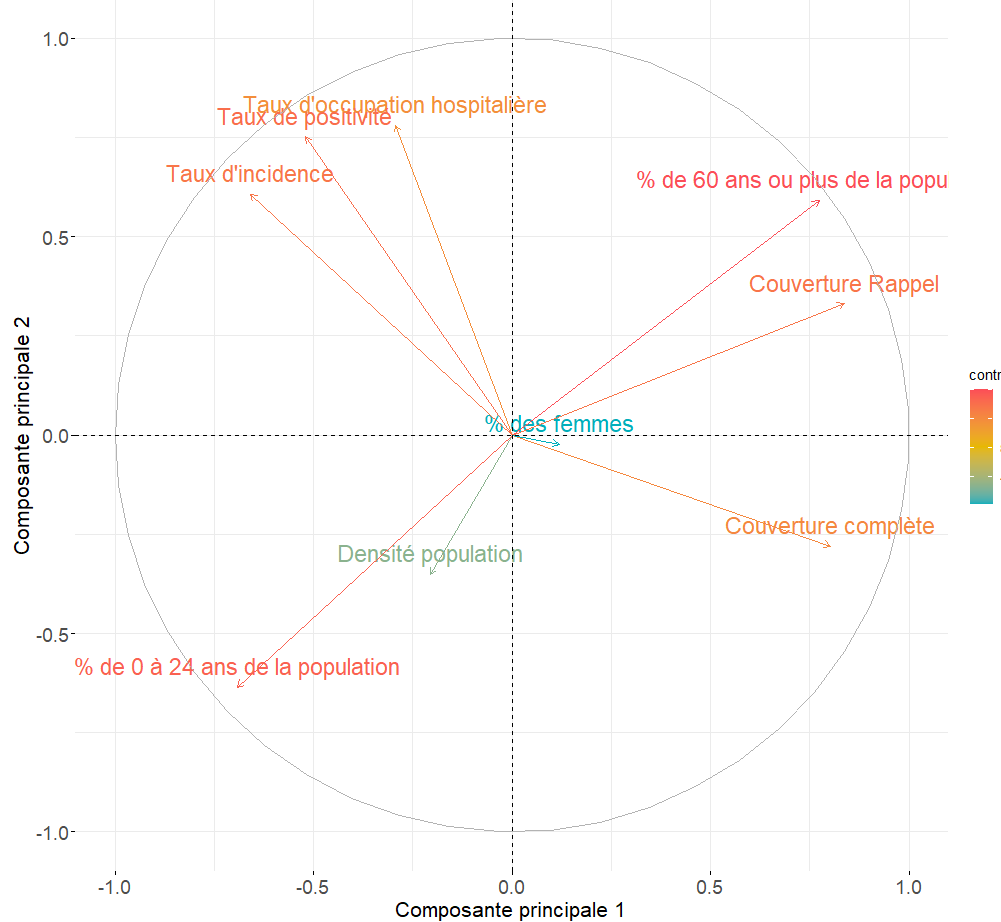
**Lecture** : les taux de positivité, d’occupation hospitalière, et d’incidence définissent le premier axe.

### ACP, Interprétation et définition des dimensions 1, 2 et 3

Pour mieux comprendre la relation entre les variables et les individus (dans ce cas, les départements), il est crucial de bien déterminer les dimensions choisies et surtout de confirmer les conclusions précédentes.

Cela permet de mieux comprendre les patterns cachés dans les données et de regrouper les individus en fonction de ces relations.

Figure 12. Représentation des variables sur le cercle des dimensions 1 et 2



**Contribution à Dim 1**

16

12

8

4

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La contribution de la variable « Couverture Rappel » à la première dimension dépasse 16% tandis que la « Densité population » est inférieur à 4%. La variable « Taux d’occupation hospitalière » est fortement positivement corrélée à la deuxième dimension du fait qu’elle est proche de celle-ci et positionnée en haut du cercle.

Les indicateurs taux d’incidence, taux de positivité et taux d’occupation hospitalière sont proches les unes des autres, d’où elles sont corrélées positivement. Càd si l’une augmente, l’autre augmente également dans une proportion similaire.

La position du taux d’incidence à gauche près du bord du cercle de corrélation et proche de la dimension 1 indique une corrélation négative avec celle-ci et une forte contribution à la même dimension.

Le Taux d’occupation hospitalière quant à lui est fortement correlé ainsi que le taux d’incidence et le taux de positivité avec la dimension 2.

De même, la couverture du rappel vaccinal et la proportion de personnes âgées de 60 ans ou plus sont fortement corrélées et sont positionnées en haut à droite près du bord du cercle et de la première dimension, indiquant leurs fortes contributions et leurs corrélations positives avec la même dimension.

La proportion de la couverture vaccinale complète est également positvement fortement correlée à la première composante.

Le deuxième axe est fortement corrélé négativement avec la proportion de la population jeune (24 ans ou moins). Sa contribution à ce même axe est également remarquable, puisqu'elle dépasse les 16%.

Ainsi, On peut conclure que les zones avec des couvertures vaccinales et de rappel vaccinal élevées se situe à droite.

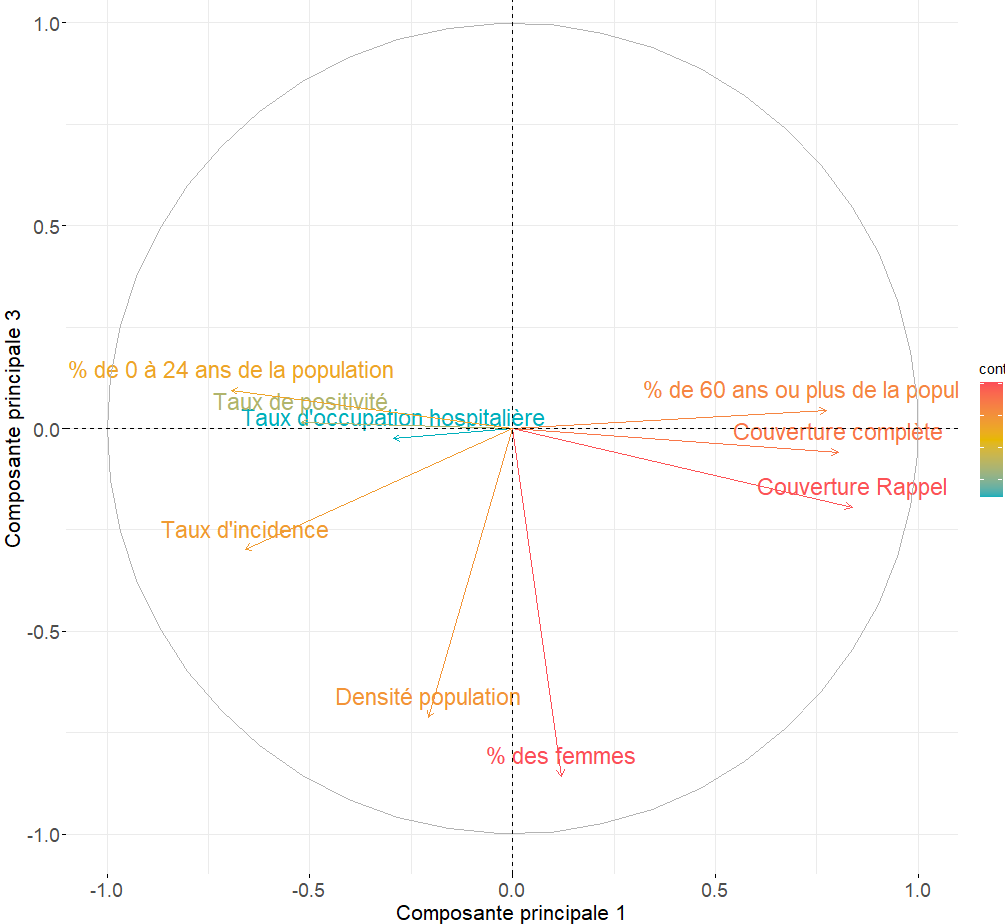
Les communautés plus âgées se situent à droite en haut.

Les zones avec des taux d’incidence et de positivité élevée se situent à gauche et en haut.

Les zones avec une population plutôt jeune se situent en bas à gauche.

Une définition plus claire dans [la section j de l’annexe.](#_ACP,_Représentation_des)

Figure 13. Représentation des variables sur le cercle des dimensions 1 et 3



16

12

8

4

**Contribution à Dim 1**

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : La contribution de la variable « Couverture Rappel » à la première dimension dépasse 16% tandis que celle de la variable « Taux d’occupation hospitalière » est inférieur à 8%. La variable « % des femmes de la population » est fortement négativement corrélée à la troisième dimension du fait qu’elle est proche de celle-ci et positionnée en bas du cercle.

Ce graphe ne révèle pas de conclusions significatives en plus de celles introduites précédemment sur la première dimension. Même si la densité de population et la part des femmes dans la population présentent une forte corrélation avec la troisième dimension, il n'y a pas de conclusion qui peut être tirée concernant l’effet du sexe ou la densité de population sur la propagation du COVID-19 dans les différents départements.

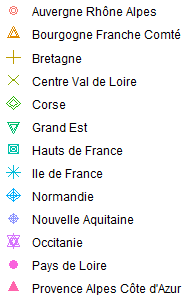
### ACP, Représentation des individus sur les dimensions 1, 2 et leurs regroupements

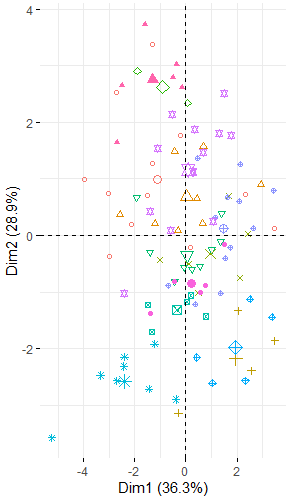
La visualisation des individus sur les dimensions choisies aide également à mieux comprendre ces relations et à identifier les groupes d'individus similaires. Le regroupement des individus en fonction de leur position sur les dimensions

Figure 14.Représentation des départements sur les dimensions 1 et 2 selon les régions

Taux d’occupation hospitalière, de positivité et d’incidence élevés

Région





Couverture Vaccinale et de Rappel vaccinal faibles, population jeune et taux d’incidence élevé

Couverture Vaccinale et de Rappel vaccinal élevées, population âgée et taux d’incidence élevé

Taux d’occupation hospitalière, de positivité et d’incidence faibles

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Les départements de la région Occitanie ont des communautés plutôt âgées et des proportions de couverture vaccinale élevées. La taille des individus dans le graphe représente une mesure visuelle de l'importance, elle peut être déterminé par la fréquence d'apparition.

Une interprétation détaillée de cette représentation est fournie dans la section   
**[le 12 décembre](#_Le_12_décembre,)**[.](#_Le_12_décembre,)

## Analyse pour le 10 avril 2021, statistiques descriptives par groupe de restrictions sanitaires

Tableau . Statistiques descriptives des indicateurs de propagation avec et sans restrictions sanitaires renforcées

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Restrictions sanitaires renforcées | non | oui | Total | p-value |
| Taux d'incidence | **n** | **77** | **19** | **96** | **<0.001** |
| **mean ± SD** | **238.91 ± 70.65** | **399.86 ± 103.28** | **270.76 ± 100.84** |
| **median** | **245.99** | **398.76** | **267.59** |
| **[Q1 ; Q3]** | **[189.91 ; 286.24]** | **[327.8 ; 458.7]** | **[207.13 ; 310.06]** |
| **[min ; max]** | **[73.13 ; 398.97]** | **[240.04 ; 606.35]** | **[73.13 ; 606.35]** |
| **N Manquantes** | **0** | **0** | **0** |
| Taux d'occupation hospitalière | **n** | **77** | **19** | **96** | **<0.001** |
| **mean ± SD** | **93.28 ± 18.69** | **138.07 ± 24.35** | **102.14 ± 26.71** |
| **median** | **97** | **151.4** | **100** |
| **[Q1 ; Q3]** | **[82.7 ; 106.1]** | **[118.25 ; 154.25]** | **[86.1 ; 116.1]** |
| **[min ; max]** | **[61.9 ; 120.4]** | **[94.2 ; 157.1]** | **[61.9 ; 157.1]** |
| **NA** | **0** | **0** | **0** |
| Taux de positivité | **n** | **77** | **19** | **96** | **<0.001** |
| **mean ± SD** | **7.4 ± 1.8** | **9.91 ± 2.16** | **7.89 ± 2.12** |
| **median** | **7.59** | **9.43** | **8.04** |
| **[Q1 ; Q3]** | **[6.25 ; 8.64]** | **[8.75 ; 11.24]** | **[6.41 ; 8.97]** |
| **[min ; max]** | **[2.81 ; 11.11]** | **[5.92 ; 14.41]** | **[2.81 ; 14.41]** |
| **N Manquantes** | **0** | **0** | **0** |

**Source** : Les auteurs, selon les données sur les indicateurs de suivi de COVID-19 et les couvertures vaccinales de data.gouv.fr et les données sur les caractéristiques démographiques de l’INSEE.

**Cohorte** : Les départements de la France Métropolitaine.

**Lecture** : Il existe une différence marquée entre le taux d'incidence moyen et son écart type dans les départements qui ont subi des restrictions sanitaires renforcées depuis le 19 mars 2021 (399.86 ± 103.28)et ceux qui n'ont pas été affectés par ces décisions (238.91 ± 70.65). Cette différence est statistiquement significative (p-valeur<0.001).

Il est possible de visualiser clairement que pour les trois indicateurs de la propagation du COVID-19, il existe des différences significatives entre les départements qui ont subi des restrictions sanitaires renforcées pendant 20 jours et ceux qui n'ont pas été concernés par cette décision. Cependant, contrairement à ce que l'on pourrait s'attendre, même après l'application de ces mesures renforcées, les taux dans ces départements restent nettement les plus élevés et au-dessus de la moyenne pour la totalité des départements de la France métropolitaine. Par exemple, le taux d'incidence moyen pour ces départements est de 399,86±103,28 pour 100 000 habitants, alors que les départements non concernés, qui sont en nombre de 77, ont un taux d'incidence moyen de 238,91±70,65 et une médiane de 245,99.

Il est important de noter que ces départements se situent principalement dans les régions d'Île-de-France, de Hauts-de-France et de Normandie.

## Code implémenté

### a. Code python sous JupiterLab

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant table

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant table

Description générée automatiquement

Une image contenant texte

Description générée automatiquement 



Une image contenant texte

Description générée automatiquement



Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant table

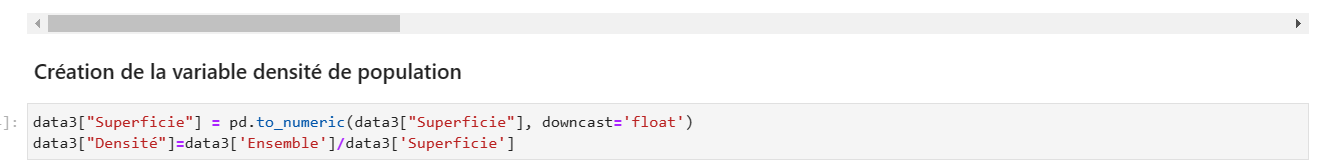
Description générée automatiquement

Une image contenant table

Description générée automatiquement

Une image contenant table

Description générée automatiquement



Une image contenant table

Description générée automatiquement

Une image contenant table

Description générée automatiquement

Une image contenant texte

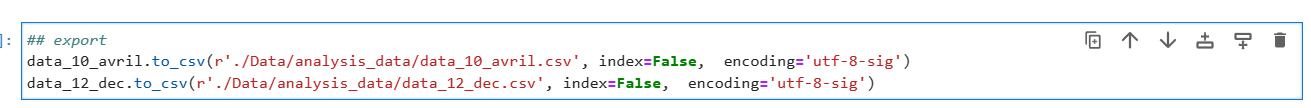
Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement



### Code R sous Rstudio

Pour l’importation des données issues de ce qui précède

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

Une image contenant texte

Description générée automatiquement

# Références bibliographiques

1. "Info Coronavirus COVID-19 - Les actions du Gouvernement." Gouvernement.fr. [En ligne]. Disponible à l'adresse : <https://www.gouvernement.fr/info-coronavirus>
2. "Population des départements français." Insee. [En ligne]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/>
3. "Population par département − France, portrait social." Insee. [En ligne]. Disponible à l'adresse : <https://www.insee.fr/fr/statistiques/>
4. "Département français : liste, carte, région, préfecture." regions-departements-france.fr. [En ligne]. Disponible à l'adresse : <https://regions-departements-france.fr/departement/>
5. "Indicateurs de suivi de l’épidémie de COVID-19." data.gouv.fr. [En ligne]. Disponible à l'adresse : <https://www.data.gouv.fr/fr/datasets/indicateurs-de-suivi-de-lepidemie-de-covid-19/>
6. "Données relatives aux personnes vaccinées contre la COVID-19 (VAC-SI)." data.gouv.fr. [En ligne]. Disponible à l'adresse : <https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19-vac-si/>
7. "A Complete Guide to Violin Plots." Chartio. [En ligne]. Disponible à l'adresse : <https://chartio.com/learn/charts/violin-plots/>