

# Projet Final Covid sujet 1

GUNDUZ Maxime ET M'PEMBELE Dieuleveut

2023-11-27

## 1. INTRODUCTION

C'est depuis la fin de l'année 2019 que le Covid-19 a fait son émergence à la vue du monde entier en partant de la Chine où les premières traces de la maladie ont été repérées pour se répandre ensuite dans le reste du monde. Il s'agit d'une maladie virale contagieuse engageant le pronostic vital, provoquée par le virus SARS-CoV-2. Cliniquement caractérisée par une atteinte respiratoire avec un tableau clinique fait essentiellement de toux, essoufflement, fièvre, asthénie physique, perte de l'odorat et/ou du goût. Les personnes les plus exposées sont les personnes âgées et les personnes ayant une comorbidité (maladie cardiovasculaire, diabète, cancer, etc.). Et bien que la plupart de personnes présentes des formes simples et qu'il existe maintenant un vaccin, la Covid-19 reste une pandémie qui a fait des ravages dans le monde entier. On dénombrait en août 2023 plus de 769,77 millions des cas et jusqu'à 6,95 millions de décès liés à la pandémie (Source : Statista, octobre 2023).

L'Europe reste l'un des continents les plus touchés par cette pandémie. L'importance de cette pandémie et de ses conséquences sur le cours de notre existence justifie de décrire son évolution dans le temps et dans l'espace. Et c'est dans cette logique que nous allons, à partir d'un recensement des données dans le monde entier, faire une analyse statistique de la pandémie de la covid-19 en termes des nombres de décès et de cas pour les pays de l'Union Européenne en 2020.

Nous allons ainsi :

- Déterminer le pays ayant recensé le plus de cas de Covid-19 et celui avec le plus de décès
- Vérifier à l'aide d'un test statistique si le rapport nombre de décès sur nombre de cas est similaire entre les pays de l'union européenne
- Montrer graphiquement l'évolution du nombre de cas, du nombre de décès et du rapport nombre de décès sur nombre de cas pour 5 pays de l'union européenne les plus touchés par la pandémie (France, Allemagne, Italie, Pologne et Espagne).
- En particulier pour la France, déterminer le mois avec le plus de cas et celui avec le plus de décès.
- Vérifier la corrélation entre le nombre de cas et le nombre de décès et entre le nombre de cas et la taille de la population.

## 2. METHODOLOGIE

Il s'agit d'une analyse quantitative rétrospective des données allant du 01/01/2020 au 14/12/2020. La population est constituée des 27 pays de l'union européenne. Et les données proviennent d'un recensement des données sur la Covid-19 pour tous les pays du monde de 2020 retrouvées sur le site du centre européen pour le contrôle et la prévention des maladies (ECDC). L'analyse des données a été réalisée en langage R à l'aide du logiciel RStudio et a fait appel à des notions de statistiques descriptives et analytiques telles que la détermination de la moyenne, de la médiane ou encore l'étude de la corrélation entre différentes variables.

## 3. RESULTATS

### 3.1. Statistiques descriptives

*Statistique descriptive des cas et des décès dus au Covid dans l'UE en 2020*

Le tableau ci-après donne une description statistique sur les cas et les décès dus au Covid-19 dans l'ensemble de pays de l'UE en 2020 :

Table 1: Statistique descriptive des cas et des décès dans l'UE

casAnnuelle	decesAnnuelle
Min. : 11101	Min. : 78
1st Qu.: 85603	1st Qu.: 883
Median : 179449	Median : 4400
Mean : 488405	Mean : 11502
3rd Qu.: 594540	3rd Qu.: 11710
Max. : 2376852	Max. : 64520

Les écart-types de distribution des cas et des décès pour les 27 pays de l'UE en 2020 sont de 639001.7 et 17686.39 respectivement.

Le pays qui a enregistré le plus de cas et celui avec le moins de cas sont respectivement la **France** et l'île de **Malte**. Et le pays qui a enregistré le plus de décès et celui ayant enregistré le moins de décès sont respectivement l'**Italie** et l'île de **Chypre**.

*Statistique descriptive pour la France sur les différents mois de 2020*

Tout comme les tableaux ci-haut, le tableau ci-après donne une description statistique sur les cas et les décès dus au Covid-19 en France en 2020 :

Table 2: Statistique descriptive des cas et des décès en France

cas	deces
Min. : 6	Min. : 0.0
1st Qu.: 19926	1st Qu.: 418.8
Median : 64193	Median : 2154.5
Mean : 198071	Mean : 4825.9
3rd Qu.: 186964	3rd Qu.: 4909.5
Max. : 886499	Max. : 21063.0

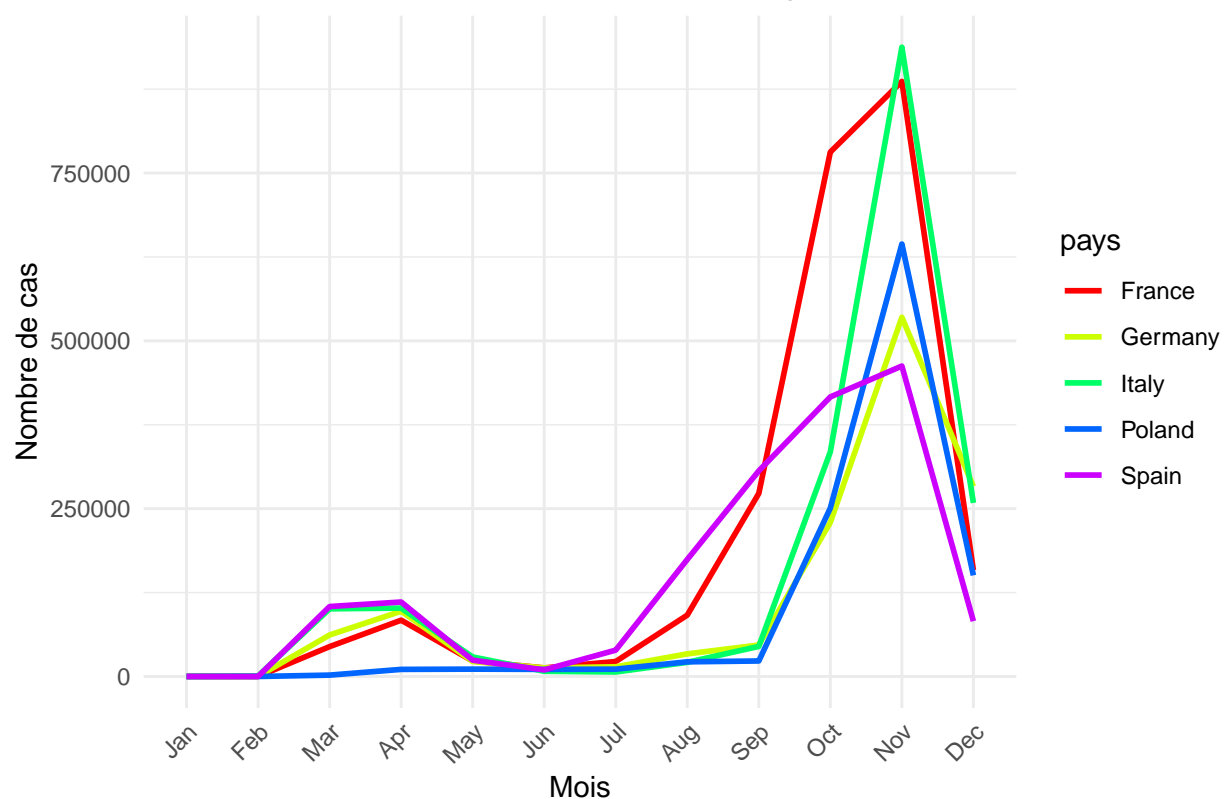
Les écart-types de distribution des cas et des décès pour les 12 mois de 2020 en France sont de 308006.1 et 6740.095 respectivement.

En France, au cours de l'année 2020 les deux mois durant lequel il a été enregistré le plus et le moins de cas de Covid-19 sont respectivement le mois de **Novembre** et le **Janvier**. Et ceux durant lequel il a été enregistré le plus et le moins de décès dus au Covid sont respectivement le mois d'**Avril** et le mois de **Janvier**.

### 3.2. Evolution du nombre de cas, nombre de décès et du rapport nombre de décès/nombre de cas

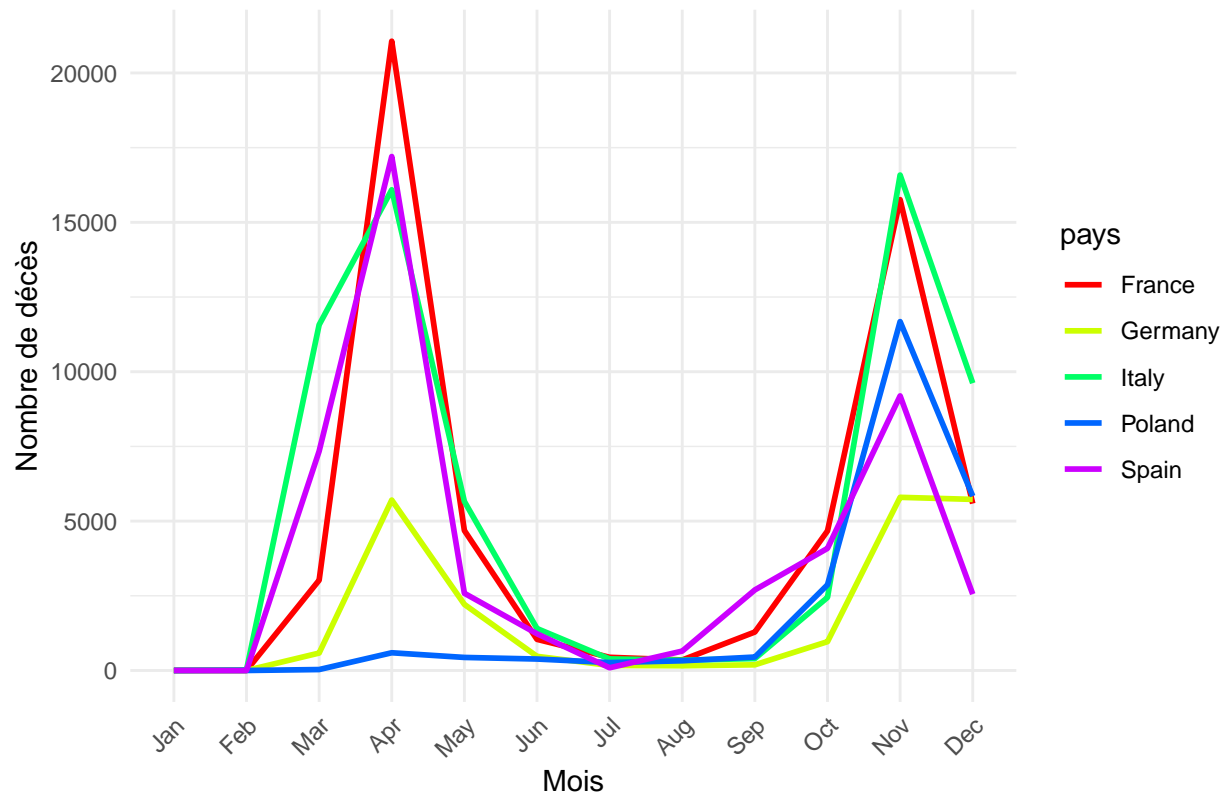
Les graphiques ci-après montrent l'évolution du nombre de cas, du nombre de décès et du rapport nombre de décès/nombre de cas en 2020 pour les cinq pays de l'UE ayant enregistré les plus grands nombres de cas de Covid-19 en 2020 à savoir la France, l'Allemagne, l'Italie, la Pologne et l'Espagne.

## Evolution du nombre de cas de Covid-19 par mois en 2020



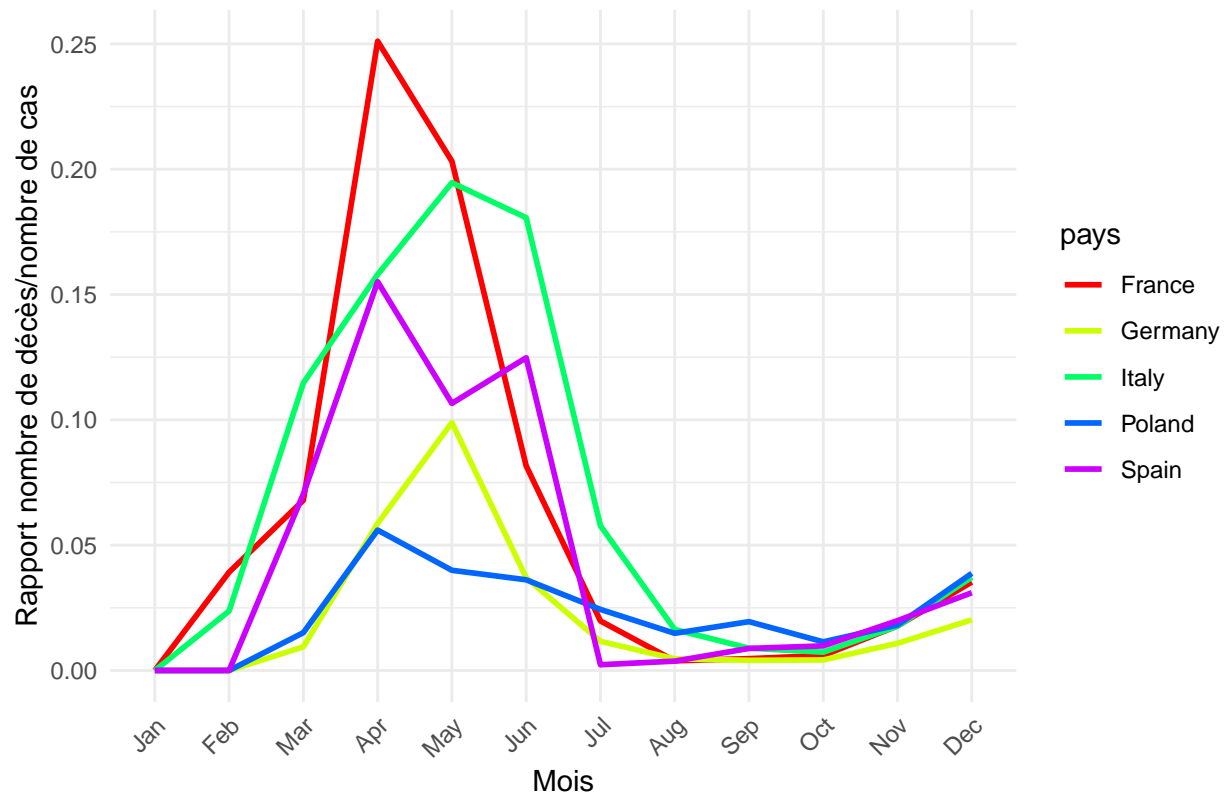
L'évolution du nombre de cas de Covid-19 est presque similaire pour ces cinq pays. On note une légère augmentation des cas entre Mars et Avril, sauf pour la Pologne. Une diminution s'en suit pour être rapidement remplacée par une nouvelle augmentation, plus importante, des cas pour les cinq pays à partir de Septembre, et même à partir de Juillet pour la France et l'Espagne.

Evolution du nombre de décès liés au Covid-19 par mois en 2020



L'évolution du nombre de décès pour ces cinq pays est caractérisée par deux pics, en Avril et en Novembre. L'évolution du nombre de décès suit une trajectoire similaire pour les cinq pays à quelques différences près. On note une augmentation à partir de Février pour atteindre le pic en Avril, ensuite une diminution importante du nombre de décès entre Mai et Août. A partir de Septembre on note nouvelle augmentation du nombre de décès bien que légèrement moins importante que la première.

### Evolution du rapport nombre de décès/nombre de cas de Covid-19 par mois



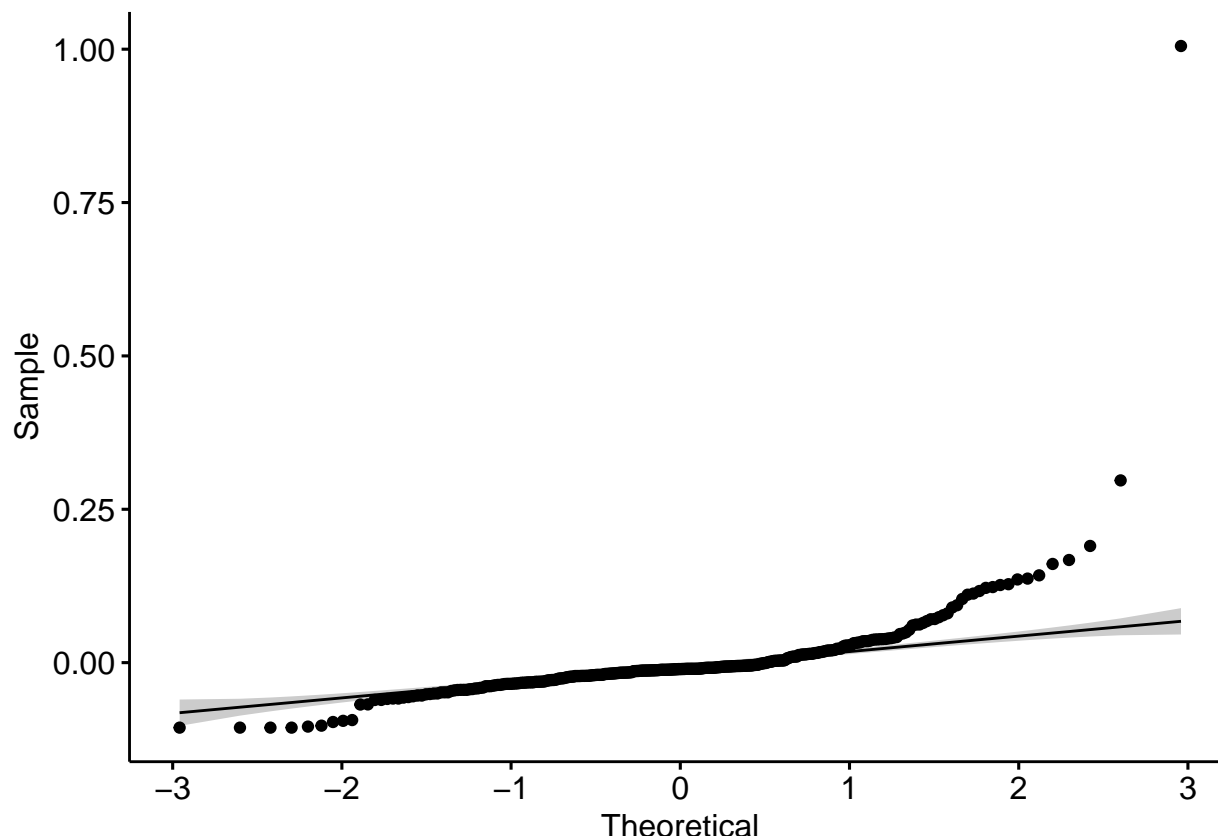
L'évolution du rapport nombre de décès/nombre de cas est caractérisée par une augmentation dès le mois de janvier. C'est entre les mois d'avril et juin que la pente devient descendante pour tomber en dessous de 0.5 entre juillet et octobre. Et à partir du mois d'octobre on note une reprise de l'ascension de cette évolution.

### 3.3. Similarité du rapport nombre de décès/nombre de cas entre les pays de l'UE

Pour vérifier si le rapport nombre de décès sur nombre de cas entre les pays de l'UE est similaire nous avons utilisé **ANOVA** qui va comparer la moyenne de rapport nombre de décès/nombre de cas des pays de l'UE. Avant de réaliser ce test nous avons vérifié la normalité de la distribution et l'homogénéité des variances.

#### *Test de normalité*

Pour vérifier si les données suivent une distribution normale nous avons fait une analyse des résidus de l'ANOVA.



#### *Test de shapiro*

Les hypothèses sont :

- $H_0$  : Les données sont normalement distribuées
- $H_1$  : Les données ne sont pas normalement distribuées

D'après le test de Shapiro, les résidus ne suivent pas une distribution normale, ( $p$ -valeur =  $2.72e-29$ , seuil d'erreur = 0.05).

Le graphique ci-haut et le test de shapiro montrent que la distribution n'est pas normale. Nous allons donc utiliser le **test de Kruskal-wallis** pour comparer les rapports nombre de décès/nombre de cas des pays de l'UE durant l'année 2020.

#### **Test de Kruskal-wallis**

D'après le test de Kruskal-wallis, il n'y a pas de différence statistiquement significative entre les groupes de pays en termes de moyenne de rapport nombre de décès/nombre de cas, ( $p$ -valeur =  $6.31e-02$ , seuil = 0.05).

### **3.4. Test de Corrélation**

#### *Corrélation Cas-Décès*

Il y a une corrélation statistiquement significative entre le nombre de cas et le nombre de décès, Le coefficient de corrélation : 0.5677285, ( $p$ -valeur =  $< 2.2e-16$ , seuil = 0.05).

#### *\*Corrélation Cas-Population*

Il y a une corrélation statistiquement significative entre le nombre de cas et la taille de la population, Le coefficient de corrélation : 0.3688516, ( $p$ -valeur =  $4.52e-282$ , seuil = 0.05).

### 3. DISCUSSION

Lors de notre étude, nous avons constaté que la France avait enregistré le plus grand nombre de cas de Covid-19, tandis que l'île de Malte en avait enregistré le moins. Cette disparité peut s'expliquer par plusieurs facteurs, tels que la population (test de corrélation cas et population) et le niveau de tourisme. La France et l'Italie font partie des cinq pays les plus peuplés de l'Union européenne et sont également parmi les trois principales destinations touristiques, alors que Malte et Chypre font partie des cinq pays les moins peuplés. Le nombre de décès dépend directement du nombre de cas (test de corrélation cas et décès) ce qui explique encore cette disparité entre ces pays en termes de décès. En plus de la taille de la population d'autres facteurs tels que la prévalence des maladies cardiovasculaires, du diabète ou du cancer, la répartition par tranche d'âge de la population ou la performance du système de santé doivent être pris en compte, ce qui pourrait expliquer par exemple que l'Italie compte le plus grand nombre de décès. En France, le mois de janvier a enregistré le moins de cas et de décès, ce qui est logique car c'était le début de la pandémie. Le mois d'avril a enregistré le plus grand nombre de décès, ce qui s'explique d'une part par l'augmentation du nombre de cas pendant cette période et d'autre part par l'impuissance des systèmes de soins face à cette menace inconnue. Le mois de novembre comme le mois avec le plus grand nombre de cas peut s'expliquer par la levée du confinement et une atténuation des mesures restrictives, mais aussi l'arrivée de la deuxième vague avec un variant beaucoup plus contagieux. Ces mêmes raisons expliquent également l'évolution du nombre de cas et de décès dans les cinq pays de l'Union européenne ayant enregistré le plus grand nombre de cas pendant toute la durée de la pandémie. En ce qui concerne les cas, on observe une augmentation en avril, avec une hausse du nombre de décès, en raison de la nature contagieuse et méconnue de la pandémie et à l'impuissance des systèmes de soins. La diminution qui a suivi est expliquée par les mesures de restriction prises par les pays (confinement, port du masque, etc.). À partir du mois de septembre, on observe une augmentation du nombre de cas due à la levée du confinement et à l'apparition du nouveau variant. Cette augmentation a également entraîné une hausse du nombre de décès, mais moins importante que lors de la première vague puisque les systèmes de soins étaient mieux préparés à gérer les cas de Covid et les premiers vaccins étaient disponibles.

### CONCLUSION

En conclusion, la pandémie de Covid-19 a eu un impact considérable dans le monde entier, avec des millions de cas et de décès enregistrés. L'Europe a été particulièrement touchée, avec des pays tels que la France, l'Italie et l'Espagne qui ont enregistré un grand nombre de cas et de décès. La corrélation entre le nombre de cas et de décès montre l'importance de contrôler la propagation du virus pour réduire le nombre de décès. Des mesures de restriction et la disponibilité de vaccins ont permis de faire face à la pandémie, bien que des défis subsistent. Il est essentiel de continuer à surveiller l'évolution de la situation et à prendre des mesures pour protéger la population et réduire l'impact de la Covid-19.