

COVID-19 - Analyse et Visualisation

Réalisé par :
EL AOUAD Zineb
DIALLO Mamadou Aliou
DIALLO Sadou
DOSSO Siaka
EL HIRI Salah Eddine

20 mai 2025

Contents

1	Introduction	1
1.1	Méthodologie	2
1.2	Source des données	2
1.3	Sources d'information	2
1.4	Packages à installer (à décommenter au besoins)	3
2	Chargement, Exploration, Préparation et nettoyage des données	3
2.1	Chargement	3
2.2	Vérification et conversions des types de variables	3
2.3	Suppression des colonnes avec plus de 90 % de valeurs manquantes	3
2.4	Détection, traitement et suppression des doublons	3
2.5	ReVérification des donees et Chargement dans CovidData_Prepared (variable avec laquelle on va travailler tout le long)	4
2.6	Ajout d'une colonne type_location qui indique selon une location si c'est un Pays, un Continent, un territoires ou dépendances, une aggregation – Total location = 255	4
3	Analyses des données	4
3.1	Analyse comparative de l'évolution de la Covid 19 dans les 5 pays d'Europe les plus impactés	6
3.2	Fait le 20/05/2025	12
3.3	Avec RMarkdown	12

1 Introduction

Notre étude porte sur l'analyse comparative de l'évolution de la COVID-19 des pays d'Europe les plus impactés, que sont la **France**, l'**Allemagne**, le **Royaume-Uni**, l'**Italie** et la **Russie**. L'observation est faite sur la période de **2020 à 2024**.

Dans le contexte de notre étude, nous avons procédé par **phases** :

- La première phase a consisté au **traitement des données**.
- La dernière phase a concerné **l'analyse**.

1.1 Méthodologie

L'analyse des données a mis en évidence : - l'**observation globale** de l'évolution de la COVID-19 dans les pays du monde, - une **proportionnalité par rapport au nombre d'habitants**, - puis une **observation concentrée sur les cinq pays d'Europe** sélectionnés.

Pour mener à bien notre étude, nous avons divisé la tâche de travail en différents **niveaux d'analyse** :

- **Une analyse d'évolution temporelle** portant sur l'évolution du nombre de cas de COVID détectés.
- **Une analyse comparative** permettant de mettre en évidence :
 - la différence du nombre de cas détectés entre les pays,
 - le **taux linéaire** d'évolution.
- **Des indices d'indicateurs** pour illustrer l'impact des **mesures anti-COVID-19** sur l'évolution du virus, en corrélation avec le **niveau de rigueur** propre à chaque pays.
- **Une analyse post-COVID-19**.

1.2 Source des données

Les données utilisées proviennent du dépôt GitHub de *Our World In Data*, accessible à l'adresse suivante :

<https://github.com/owid/covid-19-data/tree/master/public/data>

Le fichier README de ce dépôt fournit des précisions sur les sources et la structure des données.

1.3 Sources d'information

- *Our World In Data* : <https://ourworldindata.org/coronavirus>
- L'article dédié aux cas de COVID-19 : <https://ourworldindata.org/covid-cases> d'OWID
- **Google** : utilisé pour la vérification des données et des tendances réelles, en complément de nos résultats
- **ChatGPT** : pour la validation des informations ainsi que pour la résolution d'erreurs et de bugs dans le code, proposition de graphe adapté à une étude

1.4 Packages à installer (à décommenter au besoins) :

2 Chargement, Exploration, Préparation et nettoyage des données

2.1 Chargement

2.2 Vérification et conversions des types de variables

2.2.1 Conversion de la date et des variables catégorielles

2.3 Suppression des colonnes avec plus de 90 % de valeurs manquantes

toutes les colonnes contenant plus de 90% de valeurs manquantes (NA) ont été retirées. Leurs noms sont conservés dans une variable `colonnes_supprimees` à des fins de traçabilité.

2.3.1 Colonnes Supprimés :

```
## [1] "icu_patients"
## [2] "icu_patients_per_million"
## [3] "hosp_patients"
## [4] "hosp_patients_per_million"
## [5] "weekly_icu_admissions"
## [6] "weekly_icu_admissions_per_million"
## [7] "weekly_hosp_admissions"
## [8] "weekly_hosp_admissions_per_million"
## [9] "excess_mortality_cumulative_absolute"
## [10] "excess_mortality_cumulative"
## [11] "excess_mortality"
## [12] "excess_mortality_cumulative_per_million"
```

2.4 Détection, traitement et suppression des doublons

2.4.1 - Detectetion

2.4.1.1 Nous avons détecté 7 770 doublons sur la combinaison location + date. Ces doublons n'étaient pas de simples répétitions, mais souvent deux lignes contenant des données partielles. **C'EST DES DONNEES FRAGMENTEES**

2.4.2 Traitement des doublons

2.4.2.1 On a fusionné chaque groupe de doublons (location, date) en une seule ligne, en conservant Pour chaque colonne, la valeur non manquante (non-NA) si elle existe. Si les deux valeurs sont NA, le NA est conservé

2.4.3 Suppression des doublons après fusion

2.4.3.1 Suppression des anciennes lignes doublées du jeu de données initial + Ajout des lignes fusionnées à leur place

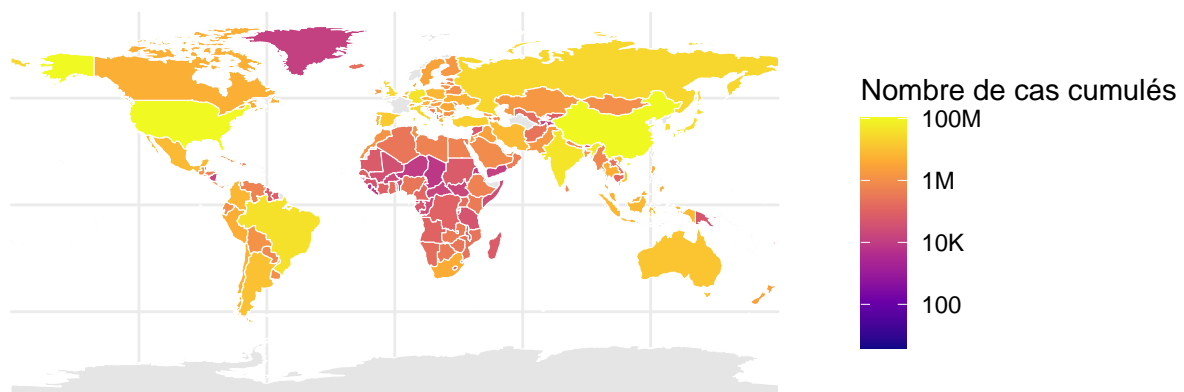
2.5 ReVérification des donees et Chargement dans CovidData__Prepared (variable avec laquelle on va travailler tout le long)

2.6 Ajout d'une colonne type_location qui indique selon une location si c'est un Pays, un Continent, un territoires ou dépendances, une aggregation – Total location = 255

3 Analyses des données

3.0.1 Nombre total de cas COVID-19 par pays sur la carte

Répartition mondiale des cas totaux de COVID-19



3.0.1.1 Interpretation : Le Graphe

Ce graphique montre la répartition mondiale des cas cumulés de COVID-19, avec un code couleur allant du violet (peu de cas) au jaune vif (très grand nombre de cas, jusqu'à 100 millions).

- **En jaune** : les pays ayant enregistré le plus grand nombre de cas cumulés (ex : États-Unis, Inde, Brésil, France).

- **En orange/rouge** : pays avec plusieurs millions de cas (ex : Canada, Espagne, Afrique du Sud, Russie).
- **En violet foncé** : pays ayant relativement peu de cas déclarés (ex : beaucoup de pays d'Afrique centrale, quelques pays d'Asie du Sud-Est).

Interprétation concrète

Les pays les plus industrialisés et les plus peuplés présentent souvent les nombres de cas les plus élevés, mais disposent également d'une capacité de dépistage plus importante.

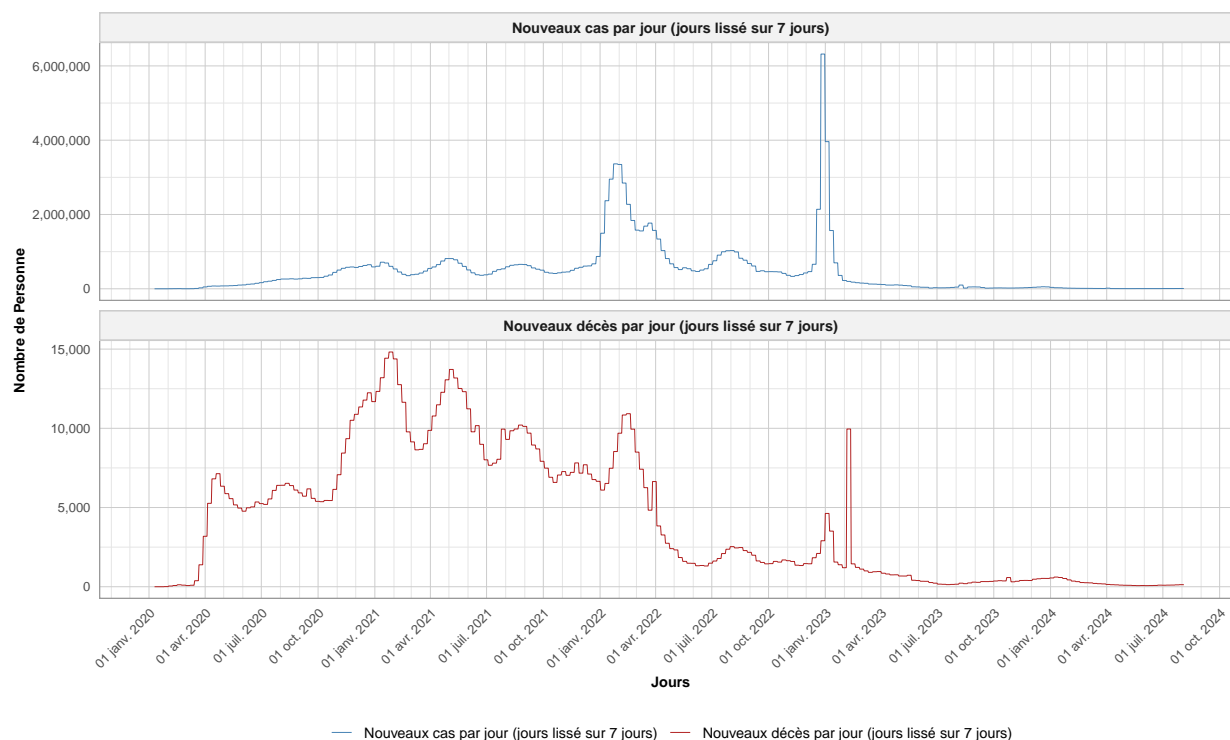
L'Afrique apparaît majoritairement en violet ou rouge foncé : cela peut traduire un nombre réel de cas plus faible, ou un sous-dépistage (moins de tests, moins de données disponibles).

Certains pays, comme la Chine ou la Corée du Nord, peuvent sembler moins touchés, mais cela peut aussi résulter des politiques de déclaration ou d'un manque de transparence dans les données.

Le graphique ne reflète pas uniquement la gravité de la situation sanitaire, mais aussi les capacités de test, les politiques de santé publique et la démographie. Il permet de visualiser rapidement les zones les plus touchées en nombre absolu.

3.0.2 Évolution Mondiale des Cas et des Décès de COVID-19

Évolution Mondiale des Cas et des Décès de COVID-19 du 05/01/2020 au 04/08/2024 (Total de Cas : 775 866 783 – Total de morts : 7 057 132)

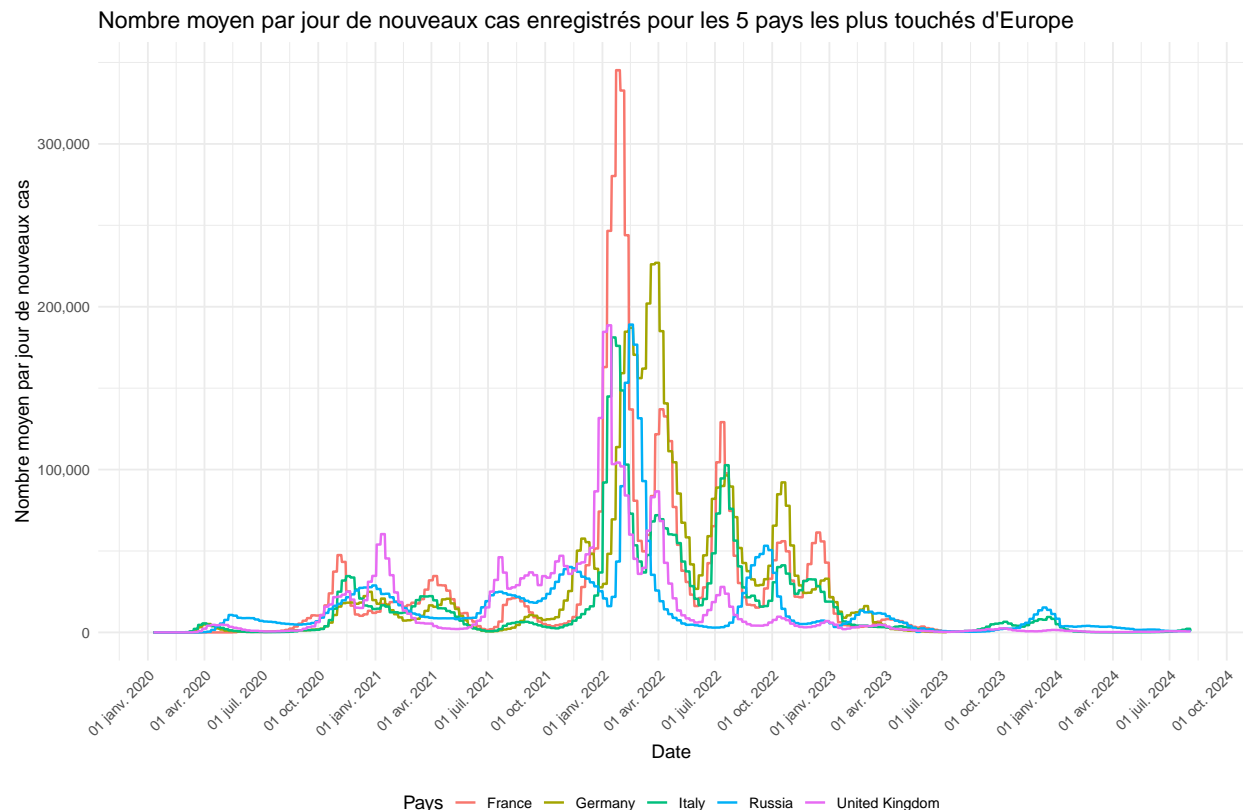


Source : Our World in Data

3.1 Analyse comparative de l'évolution de la Covid 19 dans les 5 pays d'Europe les plus impactés

3.1.1 récupérer les 5 pays d'Europe les plus touchées par le covid-19 en se basant sur le nombre total de cas enregistré pour chaque pays

3.1.2 Le graphe représentant nombre moyen par jour de nouveaux cas enregistré pour les 5 pays les plus touchés d'Europe



3.1.2.1 Évolution de la pandémie dans cinq pays européens Nous observons l'évolution de la pandémie pour chacun des cinq pays, indépendamment les uns des autres.

Les cinq pays d'Europe les plus touchés par la pandémie sont : - **France** : 20 % du taux de contamination en Europe - **Allemagne** : 19 % - **Italie** : 13 % - **Royaume-Uni** : 12 % - **Russie** : 10 %

Le taux global de contamination dans ces pays représente environ **72 %** de l'ensemble des cas détectés en Europe.

-
- Analyse par pays

- • France En début 2020, on observe une tendance moyennement importante du nombre de cas détectés, qui prend une proportion non négligeable en octobre 2020. Cette tendance reste constante jusqu'à octobre 2021 où l'on observe une croissance rapide du nombre de cas.

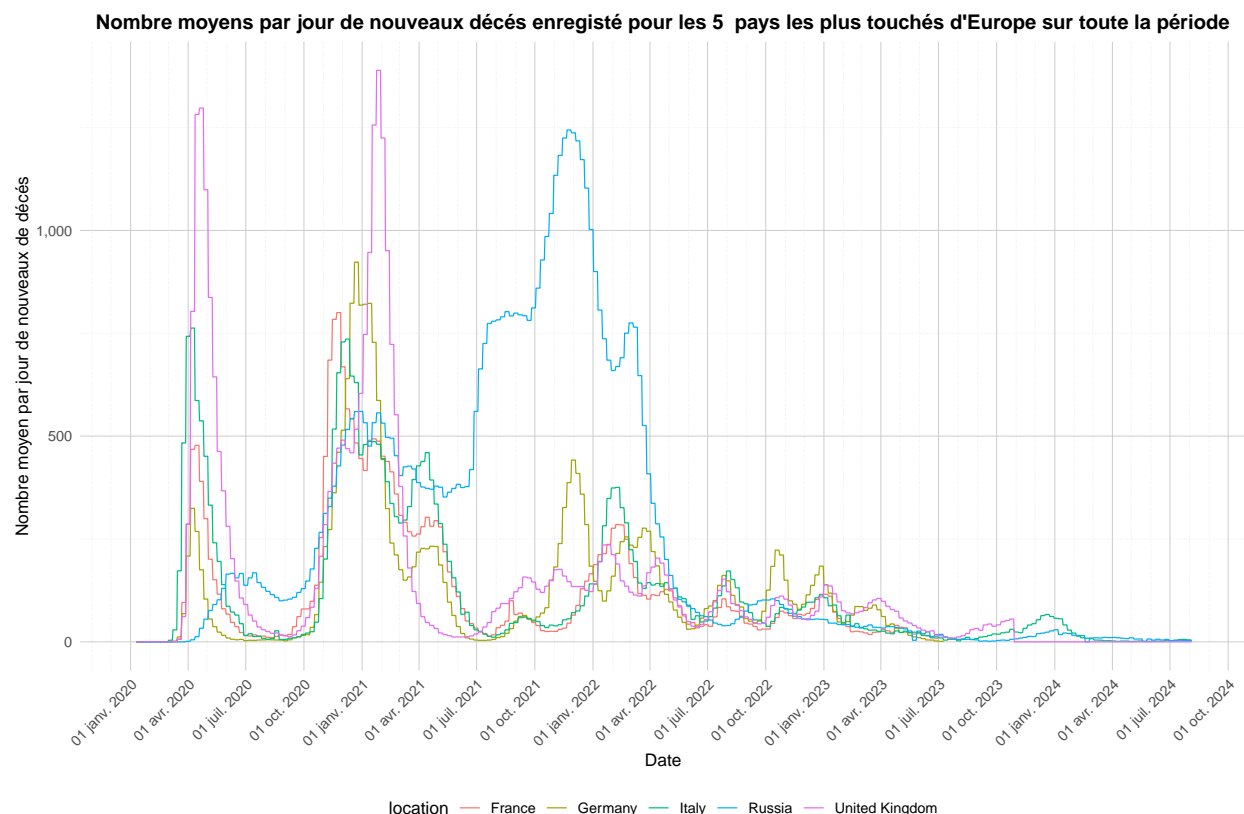
Le nombre de nouveaux cas quotidiens a atteint son pic fin 2021 et début 2022, frôlant les **350 000 cas par jour**, dépassant largement les vagues précédentes. Durant le deuxième trimestre de l'année 2022, une chute importante du nombre de cas est observée. Depuis, la situation s'améliore progressivement, bien qu'on note encore quelques petites vagues entre la fin du deuxième trimestre 2022 et 2023. À la fin de 2023, le nombre de cas s'est stabilisé à un niveau assez bas, jusqu'à devenir **négligeable début 2024**.

- • Allemagne L'Allemagne a également connu une forte vague début 2022, avec un pic autour de **200 000 cas par jour**. Par la suite, plusieurs vagues, moins importantes, se sont succédé. La courbe des contaminations est restée fluctuante, mais la **tendance générale est à la baisse à partir de 2023**.
- • Italie En début 2020, une tendance modérément importante du nombre de cas détectés est observée, qui augmente significativement en octobre 2020. Cette tendance reste stable jusqu'à octobre 2021, moment où la croissance des cas s'accélère.

Ce schéma ressemble à celui de la France, avec un pic important début 2022, atteignant près de **180 000 cas quotidiens**. Entre 2020 et 2023, plusieurs vagues se sont succédé, atteignant **100 000 cas entre juin et août 2022**, mais leur intensité a diminué avec le temps, montrant un **meilleur contrôle de la situation**.

- • Royaume-Uni Un pic très élevé a été enregistré début 2022, avec près de **200 000 cas quotidiens**. Ensuite, les cas chutent rapidement, avec seulement quelques petites vagues par la suite. Dès 2023, la courbe diminue fortement et les nouveaux cas deviennent **quasiment inexistant en 2024**, signe d'une sortie progressive de la crise.
- • Russie Les pics épidémiques sont plus étalés et **moins élevés** qu'en France ou en Allemagne. Le pic le plus marqué atteint environ **150 000 cas par jour**. Après la mi-2023, la courbe devient presque plate, ce qui pourrait s'expliquer par une **baisse réelle de la circulation du virus** ou une **modification dans la manière de tester et de rapporter les cas**.

3.1.3 Le graphe représentant nombre moyen par jour de nouveaux décès enregistré pour les 5 pays les plus touchés d'Europe



3.1.3.1 Interpretation France

En France, trois grandes vagues de décès se distinguent. La première a lieu entre mi-avril et mi-mai, avec un pic proche de 500 morts par jour. La deuxième, survenue à la fin de l'année 2020, est la plus importante, atteignant environ 800 décès quotidiens. La troisième, début 2022, avoisine les 300 morts par jour, malgré une nette baisse observée les mois précédents. Par la suite, chaque nouvelle vague de COVID-19 entraîne une hausse de mortalité moins marquée.

Allemagne

L'évolution en Allemagne est similaire à celle observée en France, avec un pic notable fin décembre 2021, autour de 900 décès par jour. Les vagues suivantes sont plus fréquentes mais moins intenses, avec une mortalité qui s'étale davantage dans le temps. À partir de la mi-2022, les augmentations deviennent moins fortes, bien que la courbe reste plus irrégulière et prolongée que dans d'autres pays.

Italie

L'Italie est l'un des premiers pays touchés, avec une première vague particulièrement meurtrière dès mars-avril 2020, atteignant près de 800 morts par jour. Une nouvelle vague significative apparaît en décembre 2020. Ensuite, la mortalité diminue, même si les vagues continuent de se succéder.

Russie

La Russie se distingue par des pics de mortalité très élevés et plus durables que dans les autres pays.

Le pic maximal est enregistré en novembre 2021, avec plus de 1 200 morts par jour. Contrairement à l'Europe de l'Ouest, la Russie reste confrontée à une mortalité élevée sur une période prolongée.

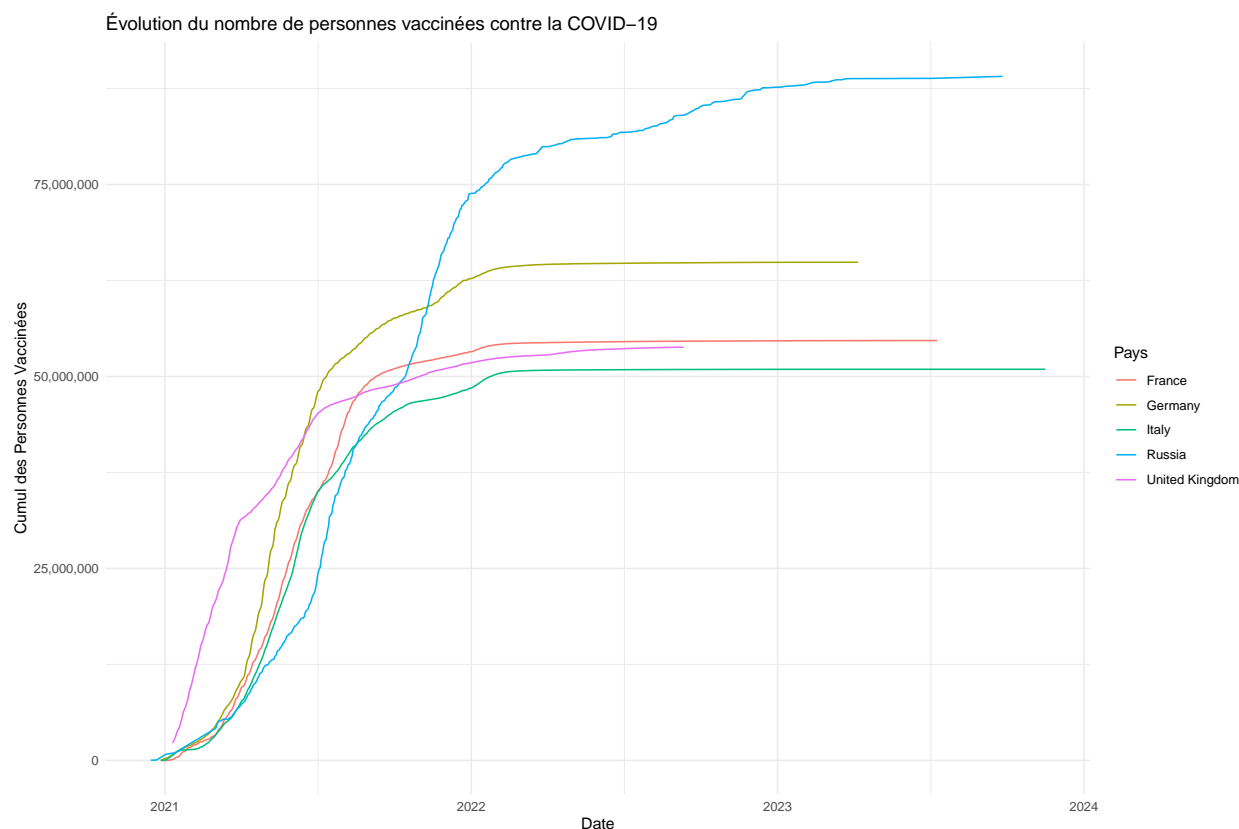
Royaume-Uni

Le Royaume-Uni connaît deux pics majeurs de décès : le premier en janvier 2021, avec près de 1 400 morts par jour — un record parmi les cinq pays analysés —, et le second fin avril 2020, avec environ 1 200 décès quotidiens. Après ces vagues, la mortalité chute rapidement. Même lors des vagues de contaminations en 2022, le nombre de décès reste relativement faible.

Synthèse comparative

La comparaison entre les cinq pays montre que la mortalité liée à la COVID-19 a culminé entre fin 2020 et début 2021, avant la généralisation des campagnes de vaccination. Le Royaume-Uni et la Russie présentent des pics particulièrement élevés, tandis que l'Italie se distingue par une première vague très précoce dès mars 2020. À partir de 2022, la tendance générale est à la baisse de la mortalité dans tous les pays.

3.1.4 Evolution des personnes ayant reçu au moins une dose de vaccin



3.1.4.1 Interpretation Les estimations de la population que nous utilisons pour calculer les paramètres de mesure par habitant sont fondées sur la dernière révision des **Perspectives démographiques mondiales** des Nations Unies.

La vaccination a été réalisée dans tous les pays, impactés ou non par la pandémie, et est devenue une obligation nécessaire pour les conditions de déplacement d'un pays à un autre sur la période

de **2020 à 2024**.

L'objectif de cette étude est d'analyser la variation des doses administrées et la masse de population ayant reçu au moins une dose de vaccin, ainsi que l'impact que ces mesures ont eu sur la propagation de la COVID-19.

Est-elle un facteur d'évolution ou un frein ?

Pour cela, nous allons observer le graphe d'évolution du taux d'injection en parallèle du graphe d'évolution de la pandémie.

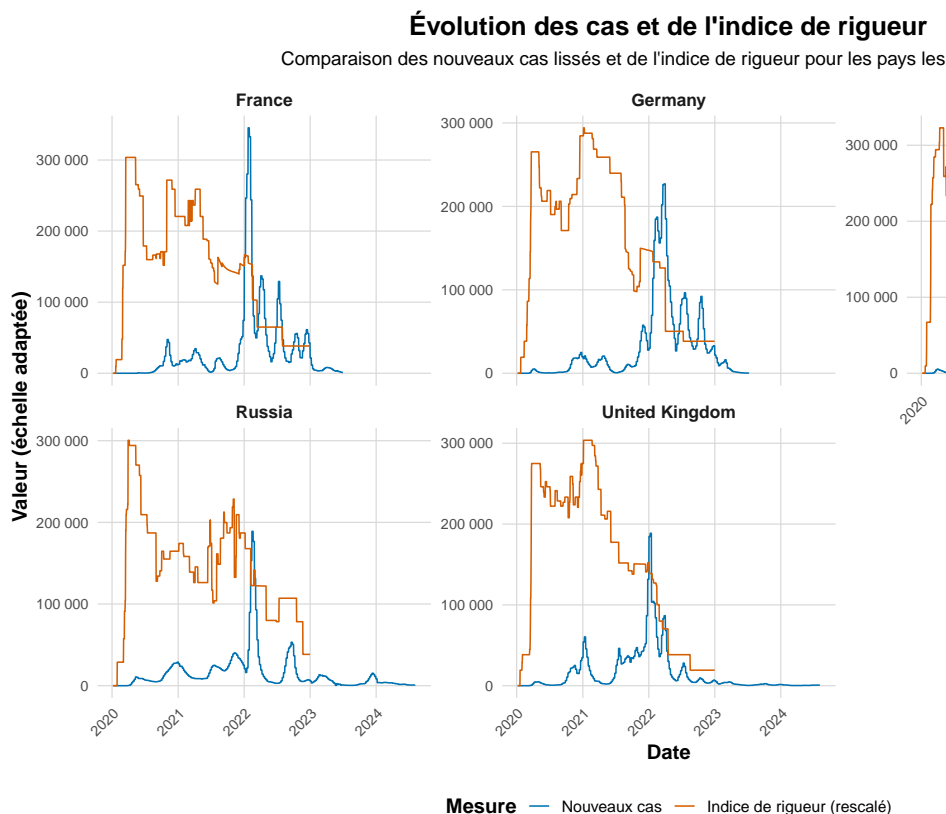
- **1,3 milliard** de doses ont été administrées en Europe de **2021 à 2024**.

Sur le graphe d'évolution du nombre de personnes ayant reçu au moins une dose de vaccin, nous observons une croissance négligeable de personnes vaccinées début 2021. Cette tendance devient progressivement importante vers la fin du premier semestre 2021, pour atteindre un pic moyen au début du premier trimestre 2022.

Après cette période, elle est restée constante pour la **France**, **l'Allemagne**, **l'Italie** et le **Royaume-Uni**.

Le nombre de personnes vaccinées en **Russie** a évolué jusqu'au début du premier trimestre 2023, puis est resté constant jusqu'en 2024.

3.1.5 Observation de l'Évolution des cas et de l'indice de rigueur selon les 5 pays européennes les plus touchés



3.1.5.1 Version sans interaction

3.1.5.2 Version avec interaction

3.1.5.3 Interpretation final : On constate que la gestion de la pandémie de COVID-19 et l'application des mesures sanitaires ont différé selon les pays européens étudiés. La France, l'Italie, l'Allemagne et le Royaume-Uni ont tous réagi de façon marquée lors des pics de contamination, en adoptant des restrictions très strictes, avec des indices de rigueur pouvant atteindre 88 au Royaume-Uni, surtout durant les années 2020 et 2021. À l'opposé, la Russie a adopté une attitude plus souple, son indice de rigueur dépassant rarement 60, même lors de vagues importantes, comme au début de 2022. En Russie, il est aussi fréquent de voir les restrictions maintenues à un niveau modéré pendant les pics, puis rapidement assouplies, parfois avant même que la situation ne s'améliore réellement. De manière générale, à partir de la mi-2022, tous ces pays ont progressivement levé les mesures, malgré quelques hausses du nombre de cas, ce qui reflète un changement de stratégie reposant davantage sur la vaccination, la gestion des hôpitaux et l'acceptation par la population. Cette analyse montre donc que chaque pays a adopté une approche différente pour faire face à la pandémie, malgré un contexte mondial similaire.

3.1.6 Analyse comparative

Le suivi de l'évolution de la COVID-19 dans les cinq pays européens les plus touchés – France, Allemagne, Italie, Russie et Royaume-Uni – montre des dynamiques variées en termes de nombre de cas et de mortalité, influencées par la progression des campagnes de vaccination. En début de pandémie, chacun de ces pays a enregistré des pics importants de contamination, avec des tendances similaires jusqu'à fin 2021, où la France et l'Italie affichent des hausses rapides, culminant respectivement à environ 350 000 et 180 000 cas quotidiens début 2022, tandis que l'Allemagne et le Royaume-Uni connaissent des pics proches de 200 000 cas par jour. La Russie, quant à elle, présente des pics plus étalés et moins élevés, avec un maximum autour de 150 000 cas par jour. Ces variations s'expliquent notamment par des différences dans la gestion de la crise et les capacités de dépistage.

Avec le déploiement massif de la vaccination dès 2021, une stabilisation puis une baisse des nouveaux cas s'observent progressivement dans la majorité des pays étudiés. En France, Allemagne, Italie et Royaume-Uni, la couverture vaccinale atteint un plateau début 2022, période à partir de laquelle la courbe des contaminations amorce une diminution notable, signe d'une efficacité du vaccin à freiner la transmission. La Russie diffère légèrement, avec une montée de la vaccination plus tardive, jusqu'au début de 2023, ce qui coïncide avec une baisse plus tardive des cas. La vaccination semble donc avoir joué un rôle clé non seulement en limitant le nombre d'infections, mais aussi en atténuant l'intensité des vagues successives.

Cette tendance se retrouve dans l'évolution de la mortalité. Les cinq pays ont connu des vagues majeures de décès entre fin 2020 et début 2021, avant la généralisation de la vaccination, avec des pics dépassant parfois les 1 000 morts par jour, notamment en Russie et au Royaume-Uni où les chiffres sont particulièrement élevés. La France et l'Italie ont suivi des trajectoires proches, marquées par une forte mortalité au début de la crise puis une diminution progressive des décès. L'Allemagne se distingue par des pics moins intenses mais plus prolongés dans le temps, certains s'étalant jusqu'en 2023. En Russie, la mortalité élevée persiste plus longtemps, reflétant une gestion plus complexe de la pandémie. Le Royaume-Uni a connu les plus hauts sommets, avec près de 1 400 décès quotidiens en janvier 2021, avant d'observer une chute rapide de la mortalité à partir de

2022. Globalement, la baisse significative des décès dans tous ces pays à partir de 2022 correspond à l'augmentation de la couverture vaccinale, ainsi qu'à l'adoption de mesures sanitaires adaptées.

Ainsi, l'analyse conjointe des cas et des décès met en lumière l'effet positif des campagnes de vaccination sur le contrôle de la pandémie. Alors que les vagues de contamination ont pu être fortes et fréquentes jusqu'au début de 2022, la vaccination a permis de limiter la gravité des cas, conduisant à une réduction progressive et durable du nombre de morts, même si la dynamique exacte varie selon le contexte national et la stratégie sanitaire de chaque pays.

3.2 Fait le 20/05/2025

3.3 Avec RMarkdown