



# Analyse Décès et Vaccination du Covid-19 (Janvier 2020- Décembre 2022)



SQL (Microsoft SQL Server)  
Excel pour les visualisations

Source de donnée :

## Coronavirus (COVID-19) Deaths

This page provides data on the number of confirmed deaths from COVID-19.

 <https://ourworldindata.org/covid-deaths>



J'ai téléchargé les données sur le site ci-dessus de début 2020 à fin 2022 pour observer l'évolution du Covid-19 sur ces 2 années qui ont été les plus importantes.

J'ai séparé le fichier téléchargé en deux base de données distinctes, l'une concernant les données liées à la contamination et aux décès et l'autre à la vaccination. Elles ont comme point commun la date et la localisation. Le but était d'être confronté à la liaison de plusieurs bases de données.

 [Code SQL](#)

## Questions

Les questions portent sur les points général de l'épidémie à savoir des informations sur le nombre et taux de décès ainsi que sur la vaccination et son impact général.

### Décès :

- Combien de décès en tout sur cette période ?
  - Pour combien de pays recensés ?
- Quels sont les 2 pays ayant été les plus touchés ?
  - En nombre de décès absolu ?
  - En taux de décès par rapport à la population ?

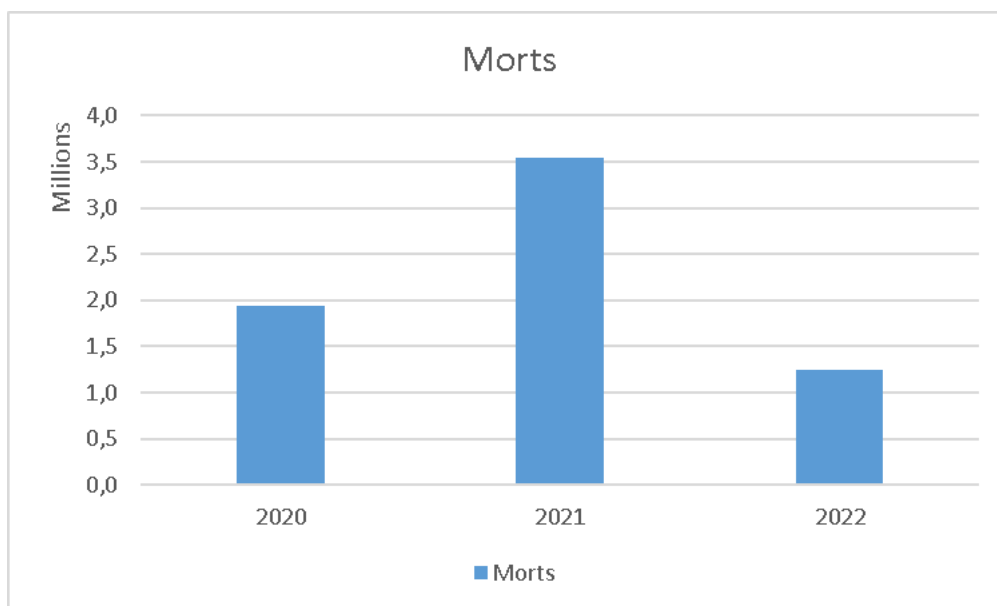
### Vaccination :

- Comment à évolué la vaccination en France et aux USA
- Quel est l'impact de la vaccination sur les décès de ces 2 pays

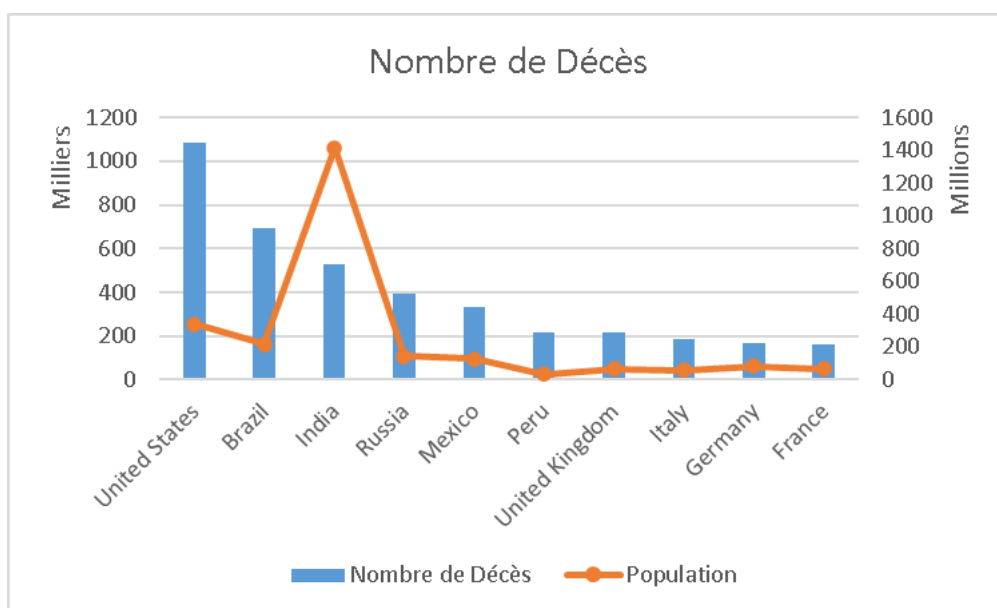
## Analyse

Il y avait dans la colonne des pays, des données n'étant pas des pays, comme Asia, Africa, Low Income, World... Ces données présentaient cependant une valeur NULL dans la colonne Continent, ce qui a permis de les exclure facilement

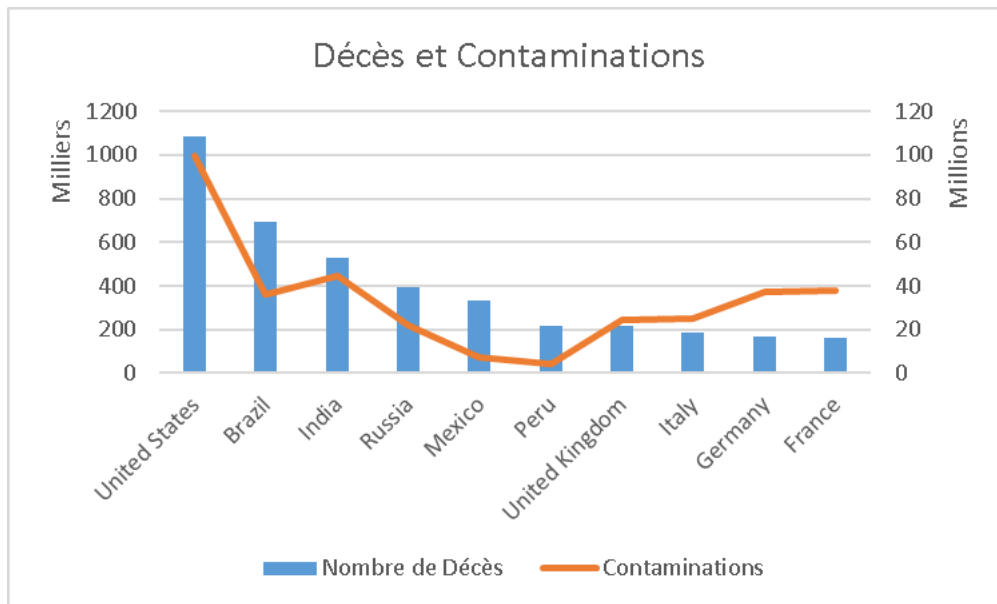
On a au total environ 6.72 millions de décès sur ces 3 années pour un total de 243 pays.



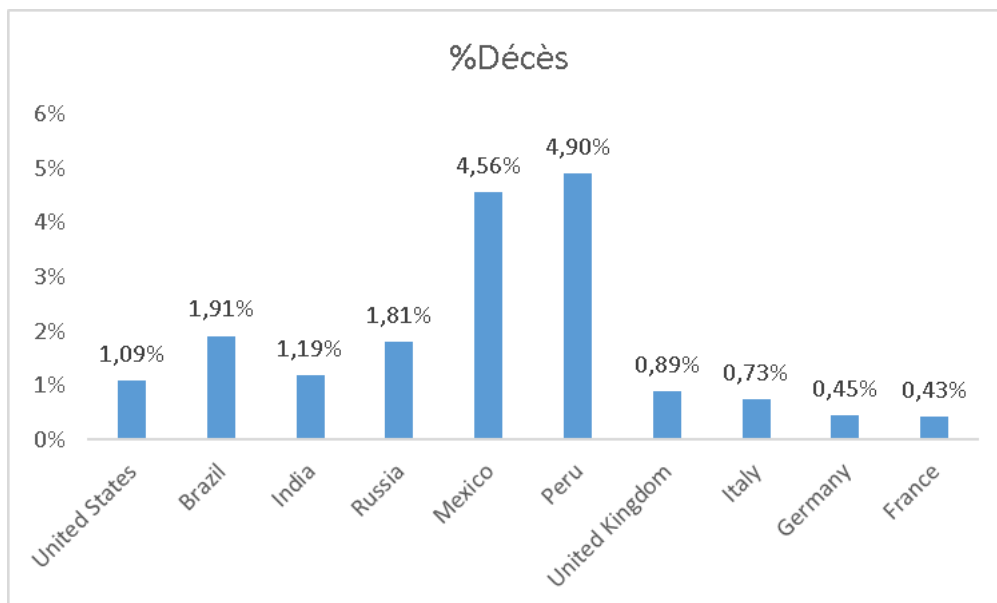
Parmi les 10 pays ayant le plus de décès on retrouve en première position les Etats-Unis et en dernière la France. On remarque aussi que malgré une population bien plus élevée que les Etats-Unis et le Brésil, l'Inde s'en sort avec un bilan moins lourd.



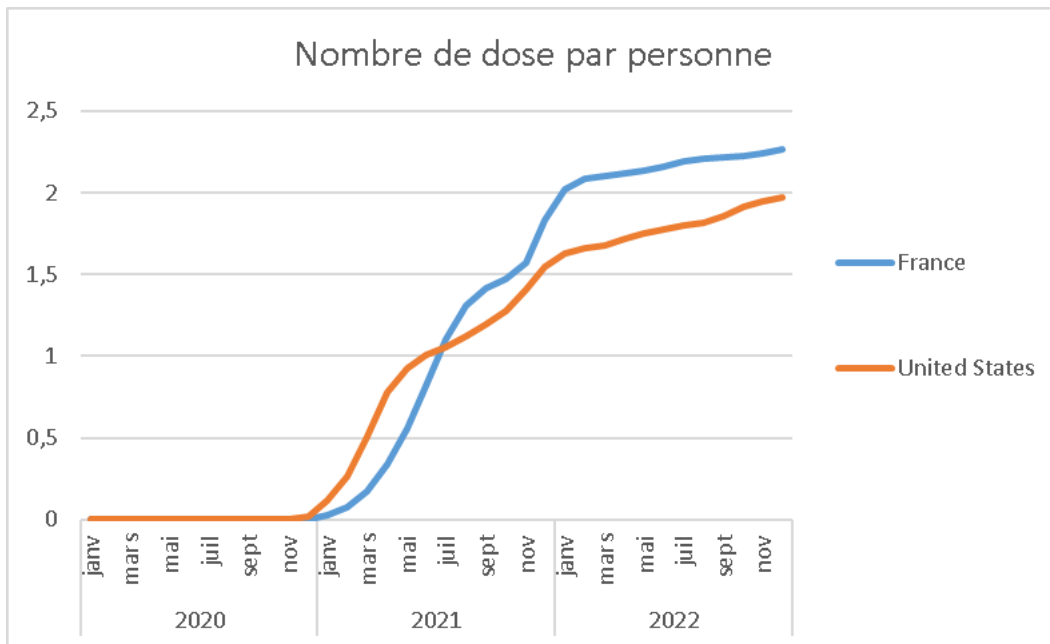
L'hypothèse la plus probable à ce constat est que les pays ayant eu moins de décès ont aussi eu moins de contamination.



Hypothèse qui se vérifie. Attention cependant à la fiabilité des données. En effet on peut supposer qu'en Inde et dans les pays ayant un système de santé moins performant, il est difficile de recenser une partie des cas de Covid-19. Cependant on va quand même considérer que les pays occidentaux ont globalement moins de décès et ont un taux de décès (par rapport aux contaminations) moins élevé.

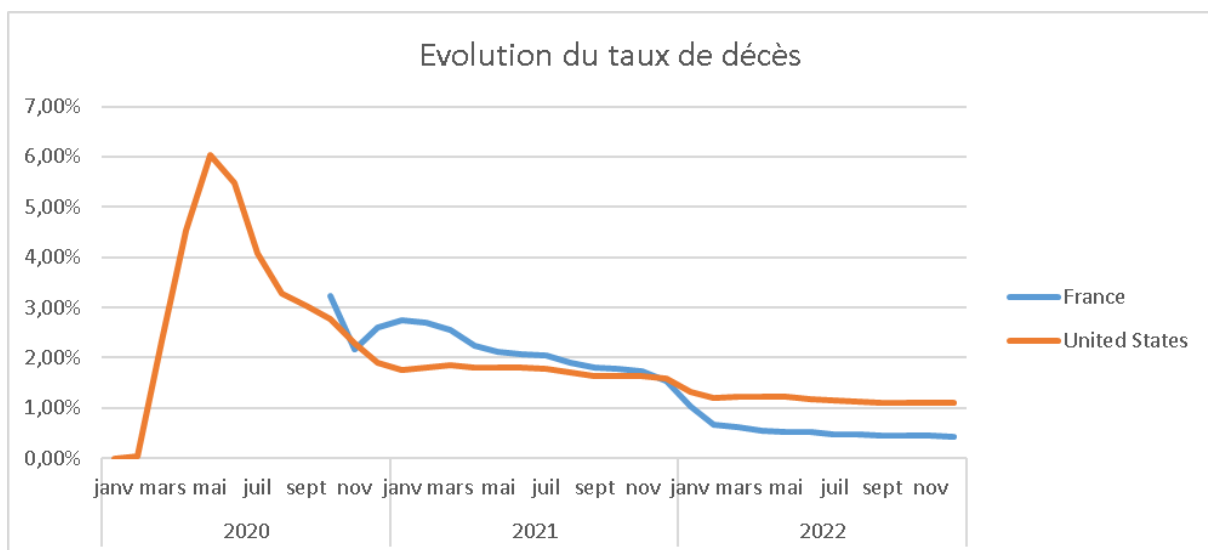


Maintenant, est-ce que la vaccination seule explique la différence de mortalité entre les Etats-Unis et la France ?



On voit bien que la vaccination en nombre de dose par personne est similaire entre les deux pays, ce qui n'explique pas la différence de mortalité des Etats-Unis, qui est de plus du double de celle de la France. Il y a donc d'autres facteurs qui expliquent un taux de mortalité plus élevé, mais on va rester dans un contexte général et étudier l'impact de la vaccination sur le taux de décès.

Les données concernant la France étant décalée (plus de mort que de cas jusqu'en Juillet 2020), je n'ai pris en compte que les données commençant le 11/10/2020.



On remarque aussi que l'on a des données dont on peut douter de la fiabilité pour les Etats-Unis en 2020, mais on peut cependant observer qu'à partir du début de la vaccination, le taux de décès tend à diminuer jusqu'à se stabiliser. Il y a donc une corrélation entre la vaccination et la diminution du taux de décès.

## CODE SQL

```
-- Nombres Total de Pays dans la base de donnée

SELECT count(DISTINCT(location)) Pays
FROM CovidDeaths$
```

```

WHERE continent IS NOT NULL -- Il y avait des valeurs n'étant pas des pays dans la colonne location, qui avait des valeurs vide dans la
ORDER BY Pays

-- Nombre de décès total

SELECT MAX(cast(total_deaths as int))
FROM CovidDeaths$

SELECT format(date,'yyyy') as Period_Year, sum(new_deaths) Deaths
FROM CovidDeaths$
WHERE continent IS NOT NULL
GROUP BY format(date,'yyyy')

SELECT date,location, sum(new_deaths) Deaths
FROM CovidDeaths$
WHERE continent IS NOT NULL AND location IN ('France','United States')
GROUP BY date, location

-- Pays comptant le plus de décès
SELECT location, MAX(cast(total_deaths as int)) Total_Décès, population, sum(new_cases) Contaminations
FROM CovidDeaths$
WHERE continent IS NOT NULL
GROUP BY location, population
ORDER BY Total_Décès DESC

-- Taux de décès des infectés

SELECT location, sum(new_deaths)/sum(new_cases)*100 Taux_Décès, sum(new_cases) Nouveaux_Cas, sum(new_deaths) Nouveaux_Décès
FROM CovidDeaths$
WHERE continent IS NOT NULL
GROUP BY location
HAVING sum(new_deaths)<>0
ORDER BY Taux_Décès DESC

-- Evolution de la vaccination
WITH CTE_Cumulé_Croissant as (
SELECT
    dea.location,
    dea.date,
    vac.new_vaccinations,
    sum(CONVERT(int,vac.new_vaccinations)) OVER (PARTITION BY dea.location ORDER BY dea.location, dea.date) as Cumulé_Croissant,
    dea.population,
    dea.total_deaths,
    dea.positive_rate

FROM [Portfolio Project | SQL Data Exploration | Project 1/4].dbo.CovidDeaths$ dea
JOIN [Portfolio Project | SQL Data Exploration | Project 1/4].dbo.CovidVaccinations$ vac
ON dea.location = vac.location
AND dea.date = vac.date
WHERE dea.location IN ('France','United States')
)

SELECT *,ROUND(Cumulé_Croissant/population,3) 'Nombre de dose par personne'
FROM CTE_Cumulé_Croissant

```