**Analyse des Données COVID-19**

**Contexte du Projet**

Ce projet consiste en une analyse approfondie des données COVID-19, en utilisant des requêtes SQL pour explorer les tendances d'infection, de vaccination et de mortalité à travers le monde. Le travail réalisé met en avant les compétences essentielles pour un analyste de données, telles que la manipulation de bases de données, l'exploration des données, le nettoyage des données et la création de visualisations prêtes à l'emploi. Les techniques SQL avancées utilisées démontrent la capacité à générer des insights critiques à partir de vastes ensembles de données, ce qui est essentiel pour la prise de décision éclairée en entreprise.

**Objectifs du Projet**

* **Évaluer l'impact du COVID-19 sur les populations mondiales** : Analyser les taux d'infection et de mortalité par pays et continent.
* **Étudier la progression des campagnes de vaccination** : Examiner les tendances de vaccination au fil du temps et leur impact sur la pandémie.
* **Fournir des insights clés pour les prises de décision** : Calculer et visualiser les pourcentages de mortalité et d'infection dans divers pays et continents.
* **Démontrer des compétences SQL avancées** : Utiliser des requêtes complexes, des mises à jour de bases de données, des vues, et des tables temporaires pour manipuler les données et faciliter l'analyse.

**Démarche Analytique**

1. **Nettoyage et Préparation des Données** : Les données brutes comportant des valeurs incorrectes (comme des zéros ou des valeurs manquantes) ont été nettoyées à l’aide de requêtes SQL UPDATE, garantissant ainsi une analyse plus précise et fiable.
2. **Exploration des Données** : Utilisation de requêtes pour comparer le taux d'infection par rapport à la population et le taux de mortalité par rapport aux cas confirmés dans plusieurs pays et continents.
3. **Calculs Statistiques et Analyse des Ratios** : Mise en place de calculs SQL pour déterminer les pourcentages d'infection et de mortalité, ainsi que l'évolution des campagnes de vaccination par rapport à la population.
4. **Utilisation de Tables Temporaires et de Vues** : Création de tables temporaires et de vues pour optimiser l'analyse des données et faciliter la réutilisation des requêtes complexes.
5. **Visualisation et Présentation** : Préparation des données à des fins de visualisation dans des outils BI pour présenter les résultats de manière claire et concise.

**Outils Utilisés**

* **SQL** : Manipulation des bases de données COVID-19 pour extraire, transformer, nettoyer et analyser les données.
* **Power BI/Tableau** : Visualisation des résultats pour une analyse plus approfondie et des présentations claires.
* **GitHub** : Partage du projet et gestion du code SQL.

**Résultats Clés**

**1. Taux de Mortalité par Cas Confirmé**

Les requêtes SQL ont permis de calculer le taux de mortalité par cas confirmé de COVID-19. Par exemple, en France, les résultats ont montré :

* **Taux de mortalité** : (total\_deaths/total\_cases)\*100 a permis de déterminer la proportion de personnes décédées parmi celles infectées. Ce ratio a mis en évidence les pays où le taux de mortalité était le plus élevé.

**2. Taux d'Infection par Rapport à la Population**

Les requêtes ont calculé la probabilité d'infection par rapport à la population totale d'un pays :

* **Taux d'infection** : (total\_cases/population)\*100 a permis de comparer l'impact du virus sur différentes populations et de montrer quels pays avaient les taux d'infection les plus élevés par rapport à leur population totale.

**3. Analyse de la Progression des Campagnes de Vaccination**

Grâce à l’utilisation des fonctions de fenêtre SQL (OVER), le projet a suivi l'évolution des vaccinations dans le monde, en comparant ces données à la population totale :

* **Pourcentage de personnes vaccinées** : Le calcul (EvolutionPeopleVaccinated/population)\*100 a été utilisé pour mesurer la couverture vaccinale dans différents pays, illustrant ainsi la rapidité avec laquelle les populations ont été vaccinées.

**4. Comparaison des Taux de Vaccination par Continents**

Une vue SQL a été créée pour suivre l’évolution des vaccinations par continent, permettant de comparer les stratégies de vaccination entre différentes régions du monde :

* **Résultat** : Les continents ayant les campagnes de vaccination les plus avancées ont été identifiés, permettant d’évaluer l'efficacité des réponses gouvernementales.

**Tableau de Bord Interactif**

Pour faciliter l’interprétation des résultats, un tableau de bord interactif a été créé dans **Power BI** ou **Tableau**. Ce tableau permet de visualiser les tendances globales de la pandémie et d'explorer les données par pays, continent, et période de temps. Il présente également des indicateurs clés tels que le **taux d’infection**, le **taux de mortalité**, et le **taux de vaccination**, permettant ainsi aux utilisateurs de générer des insights actionnables.

**Compétences SQL Démontrées**

Le projet a permis de mettre en pratique des compétences SQL essentielles, notamment :

* **Nettoyage des Données** : Utilisation des commandes UPDATE pour traiter les données incorrectes ou manquantes.
* **Jointures et Agrégations** : Maîtrise des jointures complexes et des fonctions d'agrégation pour croiser les données entre les tables CovidDeaths et CovidVaccinations.
* **Tables Temporaires** : Création de tables temporaires (#PercentPeopleVacc) pour organiser les résultats intermédiaires et optimiser les requêtes.
* **Vues SQL** : Mise en place de **vues SQL** (PeopleVaccinated) pour simplifier les requêtes répétitives et préparer les données pour les visualisations futures.
* **CTE (Common Table Expressions)** : Utilisation des CTE pour rendre les requêtes plus lisibles et faciliter la gestion des calculs complexes comme l’évolution des vaccinations.
* **Fonctions de Fenêtre** : Application de la fonction SUM() OVER() pour suivre la progression cumulative des vaccinations sur des périodes données.
* **Calculs de Pourcentage** : Calcul des pourcentages d'infection et de vaccination à l'aide d’opérateurs SQL, fournissant des mesures cruciales pour évaluer l'impact du COVID-19.

**Recommandations Business**

L'analyse SQL a permis de formuler plusieurs recommandations clés pour les décideurs et les responsables de santé publique :

* **Renforcer les efforts de vaccination dans les régions sous-développées** : Les pays et continents affichant un faible taux de vaccination devraient intensifier leurs campagnes pour atteindre l'immunité collective plus rapidement.
* **Focus sur la réduction du taux de mortalité** : Les pays ayant des taux de mortalité plus élevés devraient analyser les facteurs locaux contribuant à ce phénomène et ajuster leurs politiques de santé en conséquence.
* **Suivi continu des tendances d'infection** : Le suivi régulier des taux d'infection et de mortalité permet de réagir rapidement aux nouvelles vagues de la pandémie.