

# Rapport de l'Atelier finale\_projet

## Phase 1 : Analyse préliminaire et préparation des données

### 1. Présentation du jeu de données

Le travail s'appuie sur le dataset **Global Health Statistics**, un fichier au format CSV d'environ **134 Mo**, contenant **22 variables**. Il regroupe des informations issues de plusieurs pays et couvrant différentes périodes temporelles, permettant ainsi une analyse comparative et évolutive des indicateurs de santé mondiaux (prévalence, mortalité, traitements et impact socio-économique des maladies).

### 2. Structuration des variables

Afin de faciliter l'analyse, les variables ont été organisées en plusieurs catégories :

- **Variables qualitatives :**

Pays, nom de la maladie, catégorie de maladie, groupe d'âge, sexe, type de traitement, disponibilité des vaccins.

- **Variables quantitatives :**

Taux de prévalence, taux de mortalité, taux d'incidence, nombre de médecins pour 1000 habitants, DALYs, revenu par habitant.

- **Variables d'intérêt (cibles) :**

Taux de mortalité, taux de guérison et DALYs, utilisés pour mesurer la gravité et l'impact des maladies.

### 3. Nettoyage et enrichissement des données

Une phase de préparation a été réalisée pour garantir la qualité des données :

- Traitement des valeurs manquantes à l'aide de la **médiane**.
- Suppression des doublons sur la combinaison (Pays, Année, Maladie).
- Création de nouvelles variables analytiques :
  - **Indice de charge de la maladie** : basé sur les DALYs et la population touchée.
  - **Niveau de risque** : classification des maladies en faible, modéré ou élevé selon la mortalité et l'accès aux soins.

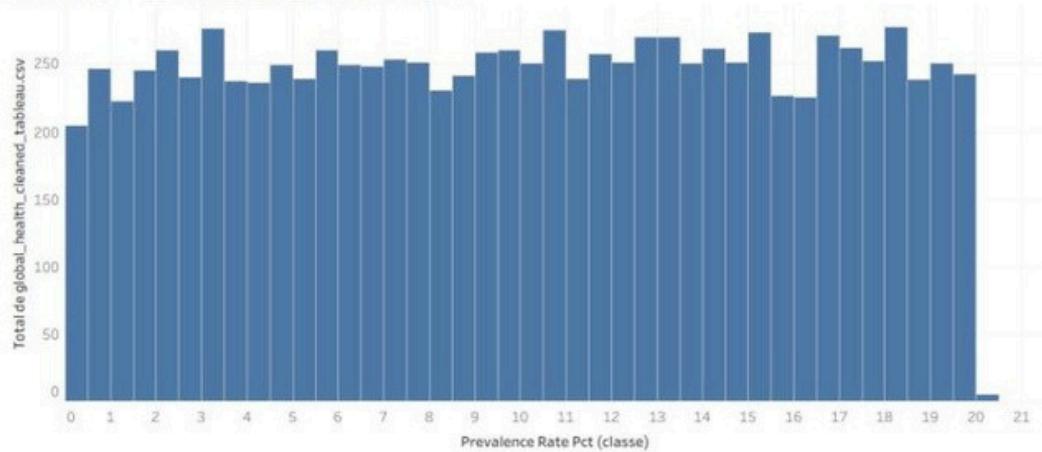
## Phase 2 : Analyse exploratoire des données (EDA)

L'analyse exploratoire a permis de mieux comprendre les tendances globales avant la conception des tableaux de bord.

- **Prévalence mondiale :**

Le taux moyen observé est d'environ 10 %, avec une distribution relativement homogène entre les populations.

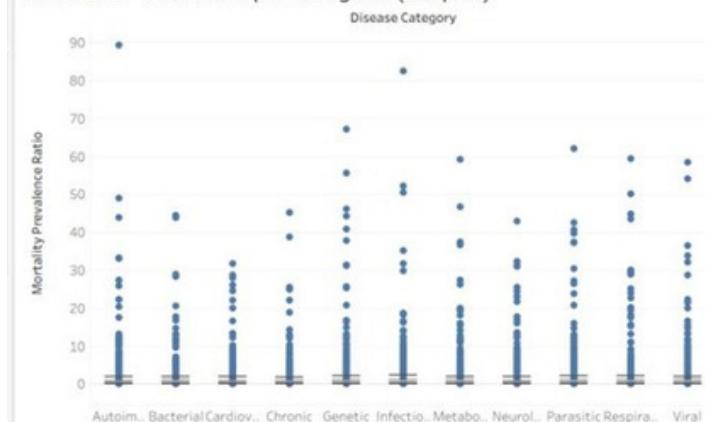
**FEUILLE 1 — Distribution de la Prévalence**



- **Mortalité :**

Bien que la moyenne soit proche de 5 %, certaines maladies présentent des valeurs extrêmes, atteignant des niveaux très élevés.

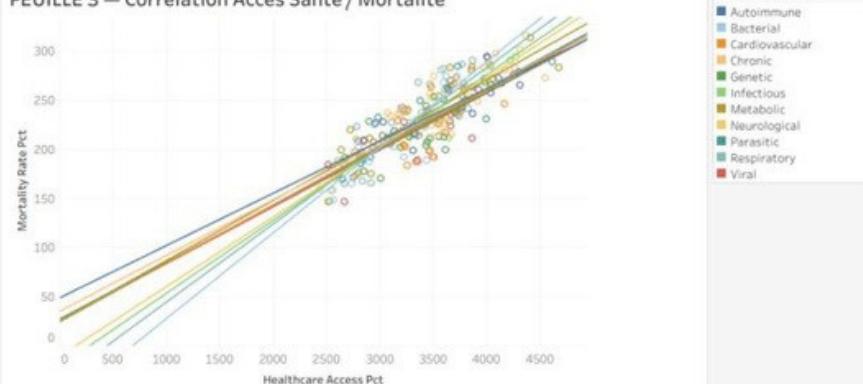
**FEUILLE 2 — Mortalité par Catégorie (Boxplot)**



- **Lien entre accès aux soins et mortalité :**

Une relation inattendue a été constatée : les pays disposant d'un meilleur système de santé déclarent parfois davantage de décès, ce qui s'explique principalement par une meilleure collecte et transparence des données.

**FEUILLE 3 — Corrélation Accès Santé / Mortalité**



## Phase 3 : Axes d'analyse et conception des dashboards

### Dashboard 1 : Analyse globale des maladies

**Question analytique :** Quelles maladies ont l'impact le plus important en termes de propagation et de mortalité ?

#### Indicateurs clés

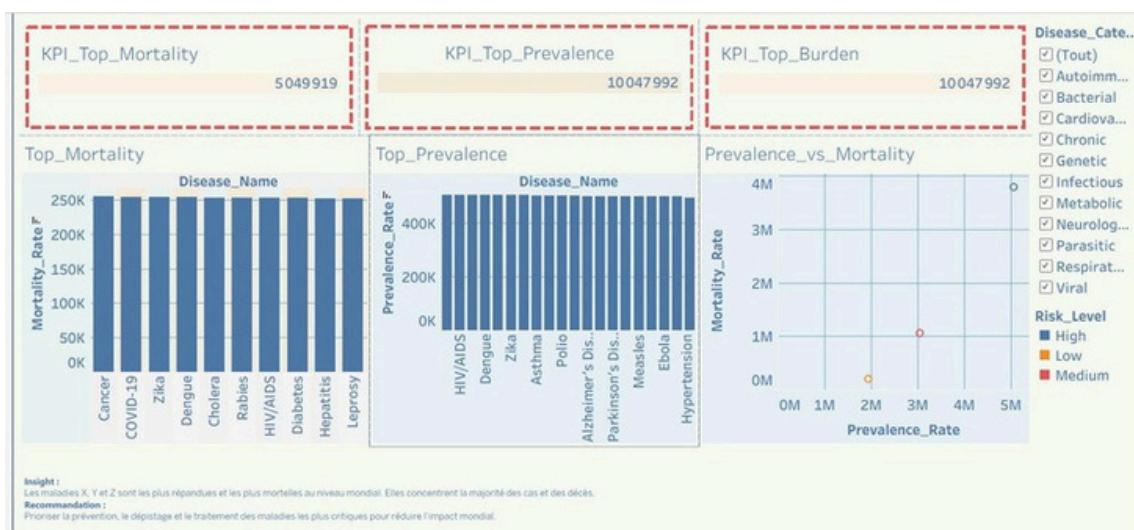
- Taux de mortalité des maladies les plus dangereuses.
- Taux de prévalence global des maladies dominantes.
- Charge globale de la maladie (DALYs).

#### Visualisations

- Classement des maladies selon la mortalité.
- Classement selon la prévalence.
- Analyse croisée prévalence-mortalité avec segmentation par niveau de risque.

#### Conclusion :

Les maladies classées à haut risque concentrent une part importante des décès et nécessitent une priorisation des actions de prévention et de dépistage.



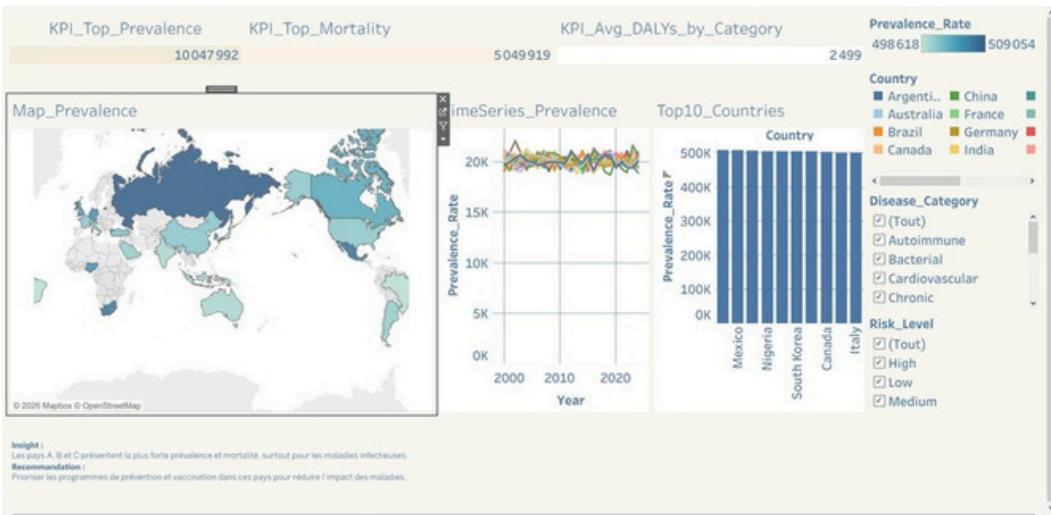
### Dashboard 2 : Analyse géographique

**Objectif :** Identifier les pays les plus touchés et analyser l'évolution temporelle des indicateurs de santé.

- Carte mondiale illustrant la prévalence par pays.
- Évolution des taux sur plusieurs années.
- Classement des pays les plus impactés.

#### Recommandation :

Renforcer les politiques de santé et l'aide internationale dans les régions présentant une charge sanitaire élevée malgré des ressources économiques limitées.



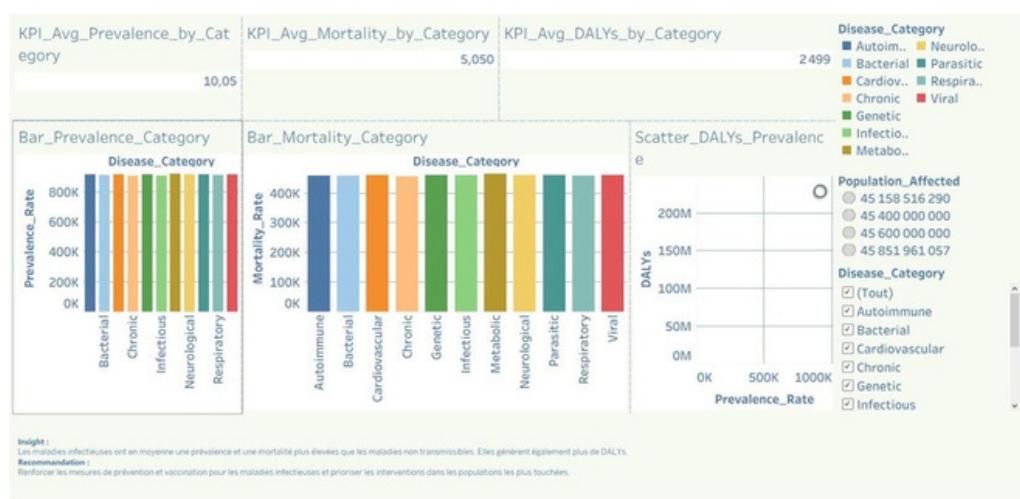
### Dashboard 3 : Analyse par catégorie de maladie

**Objectif**: Comparer les maladies transmissibles et non transmissibles.

- Comparaison de la prévalence par catégorie.
- Comparaison de la mortalité.
- Corrélation entre propagation et charge de la maladie.

**Résultat clé**:

Les maladies infectieuses restent les plus dangereuses en termes de mortalité, justifiant des stratégies de prévention ciblées.



**Conclusion générale** Cet atelier met en évidence l'intérêt de l'analyse de données et de la visualisation interactive pour comprendre les enjeux sanitaires mondiaux. Les dashboards développés constituent un outil d'aide à la décision permettant d'orienter efficacement les politiques de santé publique.