

### Université Cadi Ayyad ÉCole Supérieure de Technologie -Safi-Département informatique



Filière :Licence professionnelle en ingénierie des systèmes d'information et réseaux

# Rapport : Mini Projet sous Power BI Étude multidimensionnelle de la vaccination contre la Covid-19.

Réalisé par : BOUAMIR Assia ELAKIL Hakima

 $\frac{Encadr\'{e}~par~:}{\text{Mme. MOUNIR}~~\text{Ilham}}$ 

### TABLE DES MATIÈRES

Resume	2
Cahier de charges	3
Introduction	3
Problématique	3
Idée du projet :         4.1 questions :	3 3 4 4
L'aspect multidimentionnel et power BI  5.1 L'ETL  5.1.1 Les Dimensions  5.1.2 la Collection des donnees  5.2 Le Traitement des Données  5.3 les requêtes DAX  5.4 la creation des mesures et l'exploration des requêtes DAX  5.5 Dashboard	7 8 8 8 9 15 15
Conclusion:	19
Références	20
Table des figures	
1       Le schéma de modélisation         2       l'architecture power BI         3       ETL         4       Extraire         5       Transform (json)         6       Transform (xlsx)         7       standarisation         8       normalisation	5 7 8 9 10 11 12 13
	Cahier de charges

1 Résumé

# RÉSUMÉ

Notre initiative visait à élaborer des rapports paginés, un tableau de bord et une application destinée aux parties prenantes, dans le contexte de la gestion de la crise du coronavirus. Ces outils ont été développés pour permettre le suivi en temps réel des indicateurs clés de performance liés à la propagation du virus, aux capacités hospitalières et aux décisions stratégiques nécessaires.

Nous avons opté pour l'utilisation de Power BI en raison de sa souplesse, de sa convivialité, et de ses multiples fonctionnalités de visualisation de données, essentielles pour comprendre rapidement l'évolution de la situation.

Dans le cadre de notre projet, nous avons configuré des rapports à l'aide de Power BI pour suivre de manière claire et précise les activités liées à la pandémie. Nous avons importé des ensembles de données provenant de diverses sources, effectué un nettoyage approfondi pour garantir l'intégrité et la qualité des informations. Des requêtes DAX ont été exécutées pour extraire et manipuler les données, les rendant ainsi exploitables dans Power BI.

Grâce aux fonctionnalités avancées de Power BI, telles que les graphiques interactifs, les tableaux dynamiques et les filtres en temps réel, nous avons créé un outil de visualisation performant et facile à utiliser. Ces rapports paginés sont conçus pour simplifier l'accès aux informations cruciales, même pour les utilisateurs non spécialisés en informatique. Ils sont destinés à faciliter la prise de décisions éclairées et à améliorer l'analyse des activités liées à la gestion de la crise sanitaire. Les rapports Power BI se sont révélés être des atouts inestimables pour une gestion efficace au quotidien, sur le plan opérationnel, mensuel et annuel.

4.1 QUESTIONS:

## Première partie Cahier de charges

### INTRODUCTION

La crise mondiale provoquée par la pandémie de COVID-19 a généré un besoin urgent de comprendre, de suivre et de réagir de manière agile aux multiples aspects de cette situation inédite. Notre rapport Power BI émerge comme une solution innovante, un outil stratégique essentiel pour analyser et gérer efficacement la dynamique de la pandémie. Ce rapport offre une immersion complète dans les indicateurs clés de performance liés à la propagation du virus, aux prises de décision cruciales. Grâce à la flexibilité et à la puissance de Power BI, notre objectif est de fournir aux décideurs une vision en temps réel, claire et détaillée de l'évolution de la pandémie. Plongeons ensemble dans cet outil interactif qui vise à catalyser des décisions informées et des actions proactives pour surmonter cette crise sanitaire mondiale.

# Problématique

La problématique centrale de cette étude réside dans la conception de stratégies de vaccination globales et adaptables, tenant compte des variations notables d'un pays à l'autre, des évolutions temporelles, des disparités démographiques et de la diversité des types de vaccins utilisés. Face à cette complexité, la question essentielle à résoudre est de savoir comment élaborer des approches résilientes qui s'ajustent aux spécificités de chaque région, évoluent au fil du temps en réponse aux changements de la pandémie, et prennent en considération les différences significatives dans la répartition des vaccinations selon les groupes d'âge. Cette problématique vise à explorer les dimensions multiples de la vaccination pour aboutir à des stratégies efficaces, adaptées et capables de répondre aux défis actuels et futurs de la santé publique à l'échelle mondiale.

### IDÉE DU PROJET:

Idée de projet Power BI : Développer un tableau de bord interactif pour l'analyse mondiale de la vaccination. Ce projet vise à créer une plateforme Power BI dynamique permettant l'exploration des variations de la vaccination entre les pays et au fil du temps. En utilisant des graphiques géographiques, des séries temporelles et des visualisations spécifiques aux groupes d'âge, le tableau de bord offrira une compréhension approfondie des tendances de vaccination. De plus, il intégrera des données sur les types de vaccins utilisés, mettant en lumière l'impact de cette diversité sur la couverture vaccinale. Cette approche holistique aidera les décideurs et les professionnels de la santé à formuler des stratégies de vaccination mieux informées, plus adaptables et efficaces.

4.1 QUESTIONS:

4.2 Analyse des besoins

1. Quelle est la répartition géographique des taux de vaccination à l'échelle mondiale, et quelles sont les variations significatives entre les pays?

2. Quelles différences observe-t-on dans la couverture vaccinale en fonction des groupes d'âge, et comment ces disparités peuvent-elles orienter les stratégies de vaccination?

#### 4.2 Analyse des besoins

#### 4.2.1 Besoins fonctionnels:

plusieurs besoins fonctionnels émergent pour répondre aux questions posées et élaborer des stratégies de vaccination adaptées. Voici quelques-uns de ces besoins fonctionnels :

Analyse Géographique : Permettre une analyse approfondie des variations de la vaccination entre les pays et les continents au fil du temps. Il serait nécessaire de pouvoir filtrer les données par lieu et continent pour visualiser les tendances spécifiques à chaque région.

Suivi Temporel : Fournir des outils pour suivre l'évolution quotidienne des campagnes de vaccination. Des fonctionnalités de filtrage par date et une représentation visuelle des données temporelles permettraient de saisir les dynamiques à différentes échelles temporelles.

Répartition par Groupes d'Âge : Offrir une compréhension détaillée des différences dans la répartition des vaccinations en fonction des groupes d'âge. La capacité de filtrer et d'analyser les données par groupe d'âge serait essentielle pour adapter les stratégies de vaccination.

Impact des Types de Vaccins : Explorer l'impact de la diversité des types de vaccins utilisés sur la couverture vaccinale. Des fonctionnalités permettant de filtrer par type de vaccin et d'analyser les performances relatives de chaque vaccin contribueraient à une prise de décision éclairée.

Comparaison de Performances : Faciliter la comparaison des performances de vaccination entre les pays. Des outils de comparaison visuelle et des indicateurs de performance pourraient aider à identifier les meilleures pratiques et les leçons à tirer.

Corrélations et Analyses Approfondies : Permettre des analyses avancées pour identifier des corrélations potentielles entre différents facteurs (géographiques, temporels, types de vaccins) et les taux de vaccination. Des fonctionnalités de visualisation et d'exploration approfondie des données seraient nécessaires.

Reporting et Tableaux de Bord Interactifs : Concevoir des rapports et des tableaux de bord interactifs, offrant aux utilisateurs la flexibilité nécessaire pour explorer les données de manière personnalisée et répondre à des questions spécifiques.

Documentation du Processus ETL : Inclure des fonctionnalités qui permettent aux utilisateurs de comprendre le processus ETL appliqué aux données, assurant ainsi la transparence et la reproductibilité des analyses.

En intégrant ces besoins fonctionnels dans la conception du projet Power BI, il sera possible de fournir une plateforme robuste pour explorer les dimensions multiples de la vaccination et formuler des stratégies efficaces.

4.2 Analyse des besoins

### Modélisation de données :

#### DIMENSIONS EN FLOCON:

La modélisation en flocon de neige décrit la structuration des données issues des sources "vaccination by manufacturer" et "vaccination by age group". Pour la première source, centrée sur la vaccination par fabricant, la modélisation est construite autour de la date et du lieu comme clés primaires, avec le continent et le type de vaccin comme dimensions associées. La deuxième source, axée sur la vaccination par groupe d'âge, suit une structure similaire avec la date et le lieu comme clés principales, et le continent et le groupe d'âge comme dimensions associées. L'approche en flocon, notamment en constellation, offre une flexibilité qui permet d'explorer les données sous divers angles, en maintenant l'intégrité des informations liées à la vaccination et en facilitant l'adaptabilité du modèle aux besoins changeants d'analyse des données.

c'est pourquoi nous avons décidées d'utiliser un mélange entre le schéma en flocon particulièrement en constellation, comme illustré ci-dessous après l'application du processus ETL, au moment de la Transformation des données :

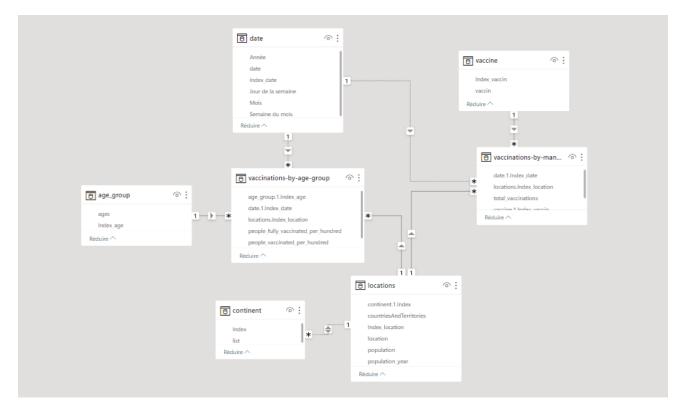


FIGURE 1 – Le schéma de modélisation

### **Conclusion:**

En conclusion de ce chapitre du cahier des charges, il est primordial de souligner l'importance cruciale de répondre de manière exhaustive aux besoins fonctionnels identifiés. La conception de ce projet Power BI pour l'analyse mondiale de la vaccination représente une opportunité significative pour les décideurs et les professionnels de la santé d'accéder à des informations pertinentes, dynamiques et stratégiques. En alignant les fonctionnalités du tableau de bord sur les exigences spécifiques liées à la variation géographique, temporelle, démographique et

4.2 Analyse des besoins

aux types de vaccins utilisés, nous visons à fournir une plateforme intégrée, souple et réactive. Cette conclusion souligne également l'engagement envers la transparence, la documentation du processus ETL, et l'accessibilité des données, garantissant ainsi la fiabilité et la compréhension approfondie des analyses. Ce projet aspire à être une ressource essentielle pour la prise de décisions informées et l'élaboration de stratégies de vaccination adaptées, résilientes et efficaces dans le contexte complexe de la santé publique mondiale.

## L'ASPECT MULTIDIMENTIONNEL ET POWER BI

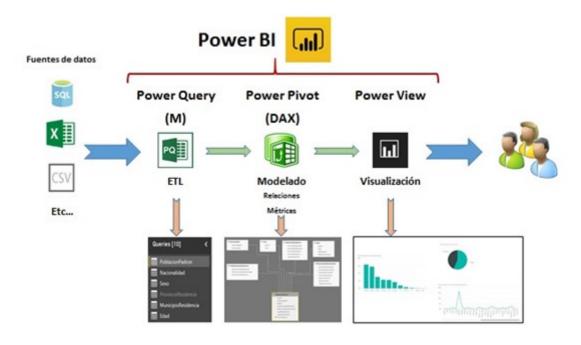


Figure 2 – l'architecture power BI

- 1. Qu'est-ce que Microsoft Power BI?
- 2. Quels sont ses avantages?

Microsoft Power BI est une solution de business intelligence (i.e Les outils Business Intelligence permettent de soutenir à la fois les processus décisionnels, stratégiques et tactiques.) cloud qui permet aux utilisateurs de collecter, analyser et visualiser des données de manière simple et intuitive. Power BI est un outil de tableau de bord interactif qui peut être utilisé pour créer :

- des graphiques
- des diagrammes
- des visualisations de données

pour aider les utilisateurs à prendre des décisions éclairées.

#### Les avantages :

Microsoft Power BI offre de nombreux avantages pour les entreprises, nous pouvons notamment citer :

- Visualisation de données : Power BI permet de créer des graphiques et des visualisations de données facilement et rapidement, ce qui vous aidera à mieux comprendre les tendances et les modèles.
- Analyse de données : des fonctionnalités avancées d'analyse de données pour découvrir des tendances, des relations et des modèles dans les données.
- Partage de données : Power BI vous permet de partager vos données et rapports avec vos collègues et même des partenaires externes, et ce de manière sécurisée et contrôlée.

5.1 L'ETL

• Interaction avec Excel : connexion de requêtes, de modèles de données et de rapports Excel avec vos tableaux de bord Power BI pour rapidement, collecter, analyser, publier et partager des données métier Excel de façon innovante.

• Compression de données : d'une performance exceptionnelle pour ceux qui travaillent avec des volumétries de données importantes.

#### 5.1 L'ETL

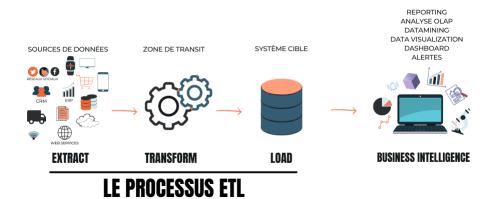


FIGURE 3 – ETL

- L'ETL (extraction, transformation et chargement) est le processus qui consiste à combiner les données provenant de plusieurs sources dans un grand référentiel central appelé entrepôt des données.
- processus ETL :
  - Extraire des données de l'entreprise brutes.
  - Qualifier et consolider ces données sources selon des règles métiers.
  - Ensuite il suffit de les charger dans un warehouse, entrepôt de données, qui deviendra la base de données des outils de Business Intelligence de l'entreprise.

#### 5.1.1 Les Dimensions

Table de fait pour le nombre total de personnes entièrement vaccinées en fonction du groupe d'âge.

Table de fait pour le nombre total de personnes vaccinées en fonction du type de vaccin. Les dimensions incluses dans ces tables de fait sont les suivantes :

- Date
- $\bullet \ \operatorname{Pays} \to \operatorname{Continent}$
- Groupe d'âge
- Type de vaccin

### 5.1.2 la Collection des donnees

• Extraction :

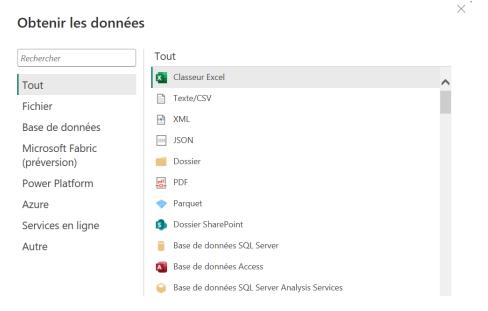


FIGURE 4 – Extraire

De la même manière, on peut extraire les diverses catégories de sources.

#### • Transformation :

Avant de passer au chargement des données en vue de la visualisation, il est nécessaire de les nettoyer et de les traiter.

• Après la transformation des données au format .csv, nous commençons la manipulation :

### 5.2 LE TRAITEMENT DES DONNÉES

Les différentes étapes et fonctionnalités fournies par Power BI que j'ai effectuées dans la phase de transformation des données :

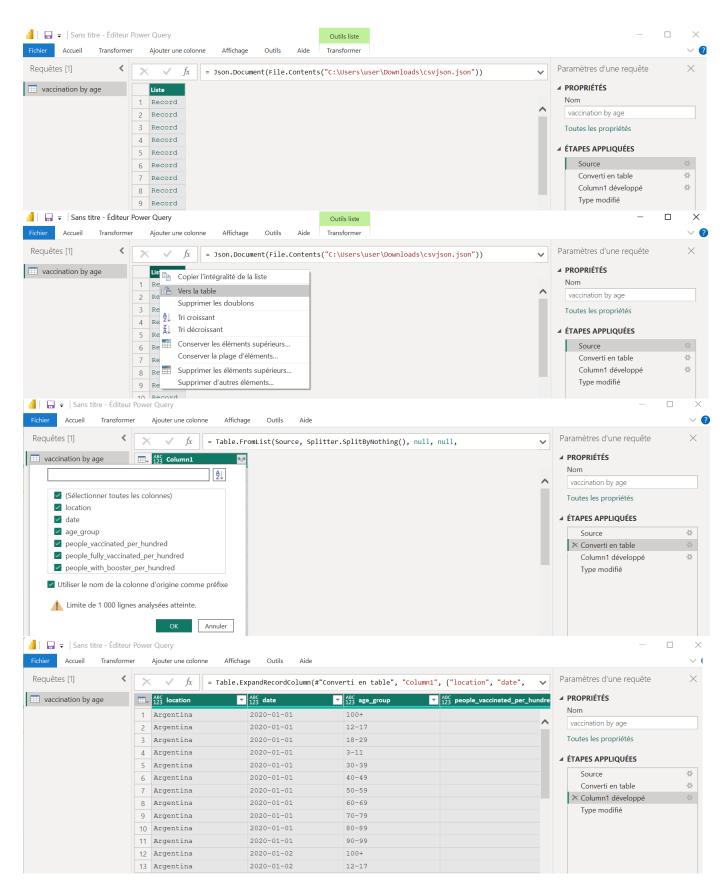


Figure 5 – Transform (json)

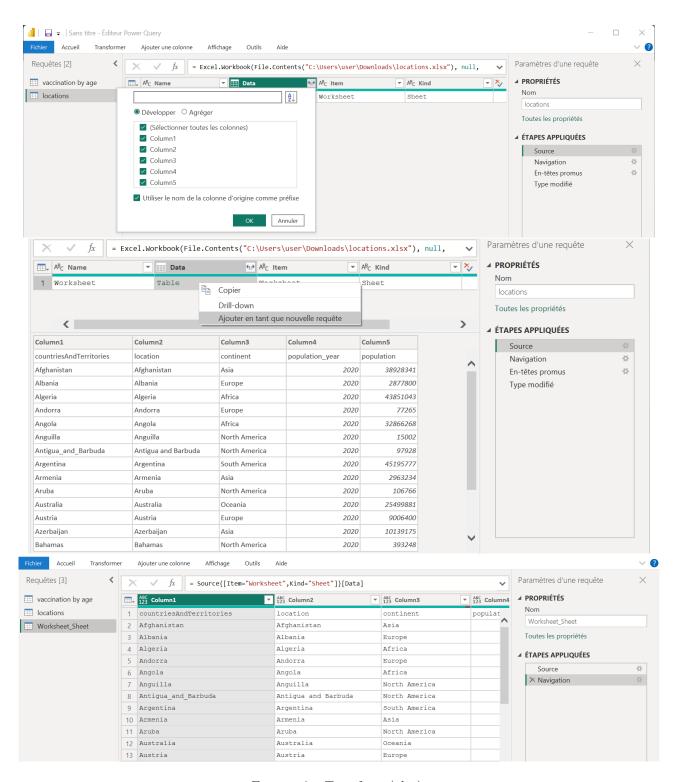


Figure 6 – Transform (xlsx)

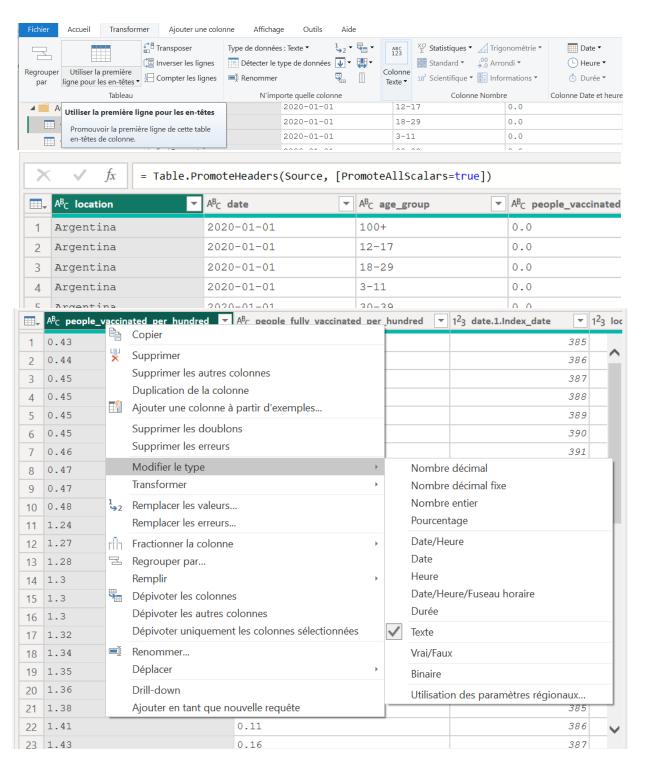


Figure 7 – standarisation

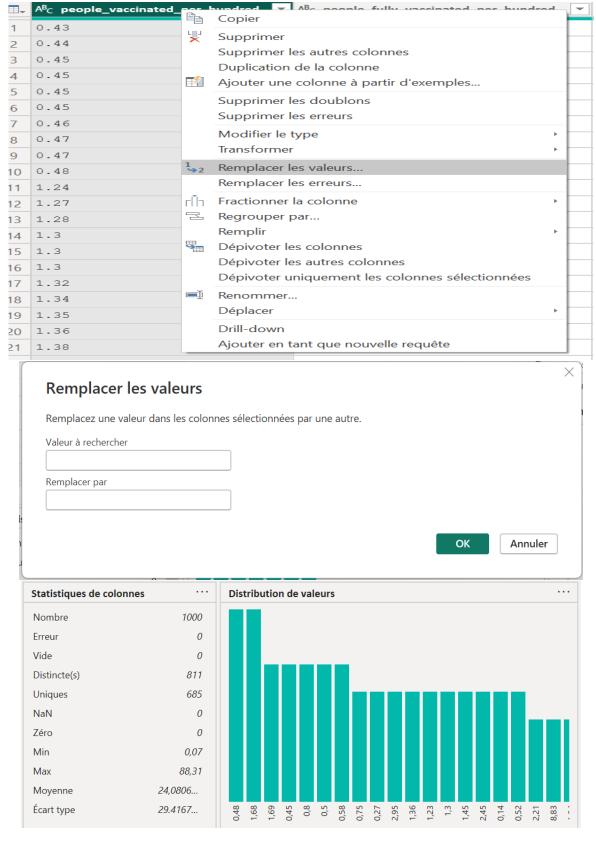


Figure 8 – normalisation

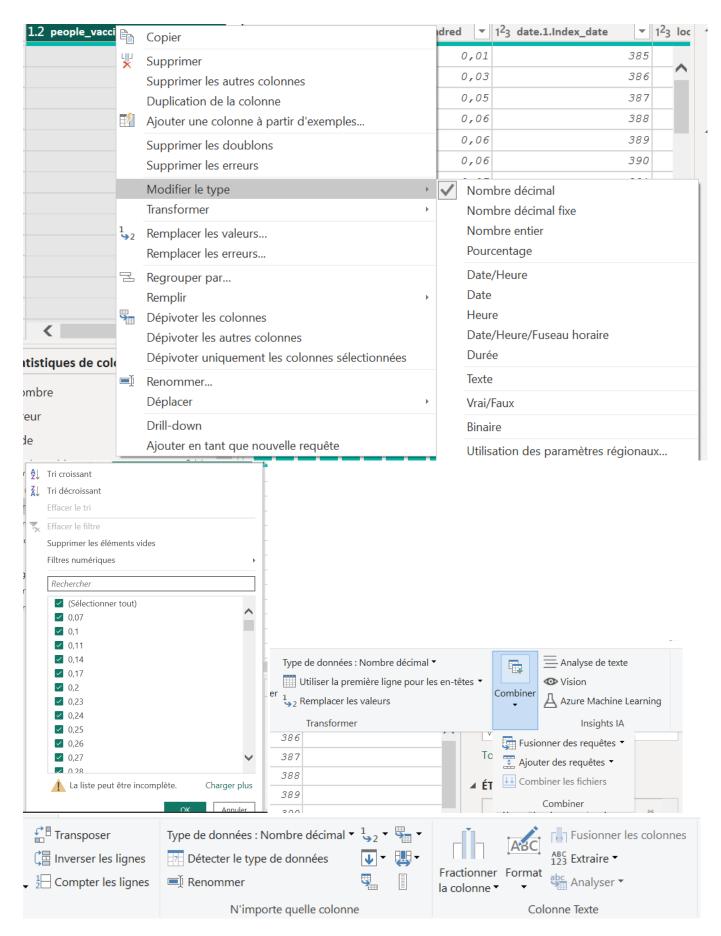


FIGURE 9 – normalisation, filtrage, Tri separation et combinaison

### 5.3 LES REQUÊTES DAX

Le langage **DAX**, acronyme anglais de « Data Analysis Expressions » est un langage de formules utilisé dans les vues Donnée et Rapport de Power BI pour le formatage, mais surtout pour la programmation de données. DAX se retrouve également dans d'autres produits Microsoft, tels que Power Pivot et le modèle tabulaire SSAS. Avec ce langage, les fonctions peuvent contenir d'autres fonctions imbriquées, des instructions conditionnelles et des références à des valeurs. L'exécution de DAX commence à partir de la fonction ou du paramètre le plus profond, et progresse vers l'extérieur. DAX est conçu pour fonctionner avec des tables. Il utilise la surcharge d'opérateur, ce qui signifie que des types de données peuvent être combinés dans les calculs, et que les résultats changent en fonction du type de données utilisé dans les entrées.

### $5.4\,$ La creation des mesures et l'exploration des requêtes DAX

Figure 10 – Mesure dax

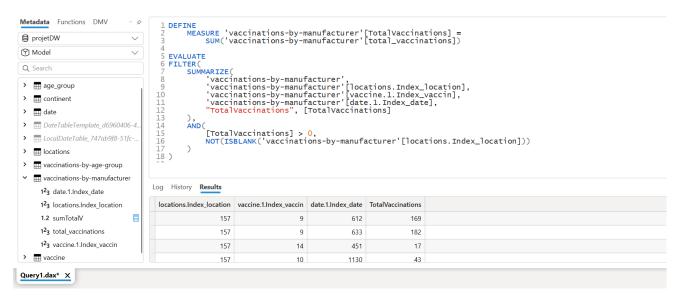


FIGURE 11 - requête DAX -réalisation des filtres sur les données par DAX

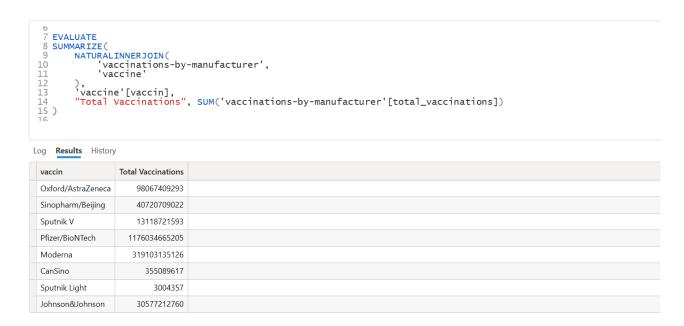


FIGURE 12 – nombre de vaccinees par rapport le type vaccin

FIGURE 13 – NATURALINNERJOIN pour générer une table du total de vaccinations par rapport le type de vaccin et le lieu

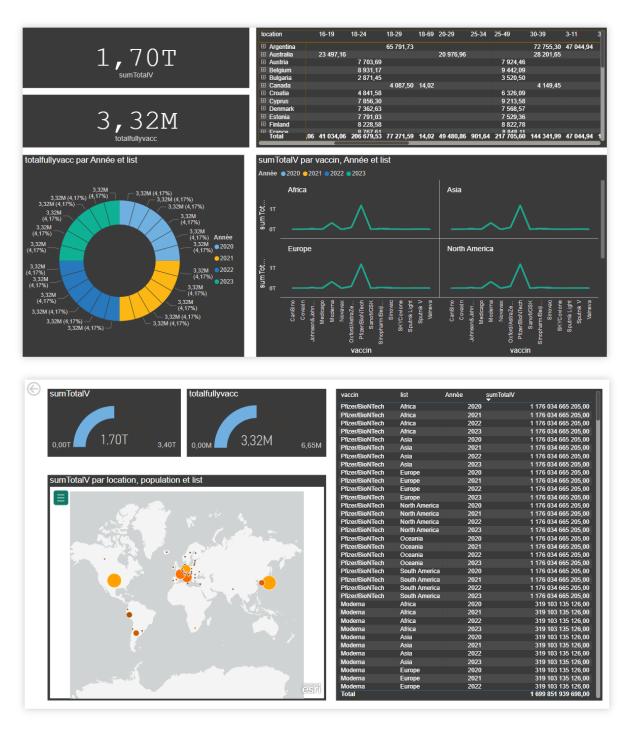
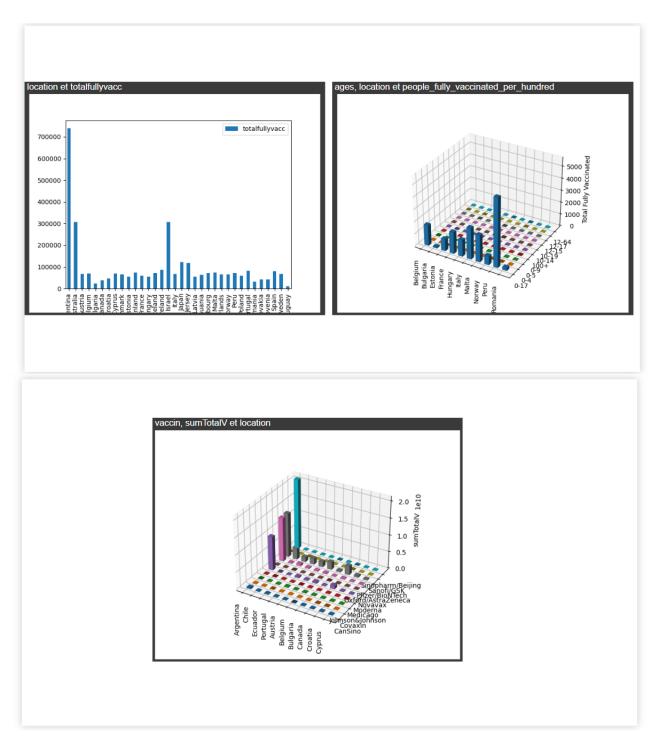


Figure 14 – Dashboard



 ${\tt Figure~15-visualisation~en~utilisant~python}$ 

CONCLUSION:

1	6
	CONCLUSION:
	Concession.

Grâce à sa large palette de visuels, Power BI garantie une meilleure compréhension des données permettant d'améliorer considérablement les processus décisionnels et opérationnels d'une entreprise. Désormais, il est possible pour l'ensemble de vos collaborateurs de prendre des décisions éclairées sur vos données.

7 Références

# RÉFÉRENCES

- $\bullet \ \texttt{https://learn.microsoft.com/fr-fr/power-bi/transform-model/desktop-query-overview}$
- $\bullet \ \texttt{https://learn.microsoft.com/fr-fr/power-query/power-query-what-is-power-query}$
- https://lesdieuxducode.com/blog/2019/3/presentation-de-power-bi
- https://learn.microsoft.com/fr-fr/dax/dax-queries