

Institut National des Sciences Appliquées et de Technologie



Rapport de Projet

Analyse BI et Visualisation des Ventes AdventureWorks

Réalisé par :

Iheb Alimi, Riadh Ibrahim, Mohamed Ben Jdira

Spécialité : Réseaux Informatiques et Télécommunications

Niveau : 4ème Année

Année Universitaire : 2024-2025

Table des matières

1	Introduction	2
1.1	Contexte du Projet	2
1.2	Présentation du Datawarehouse	2
1.3	Objectifs du Projet	2
1.4	Installation et préparation de l'environnement	3
1.4.1	Configuration nécessaire de visual studio	3
2	Création du Cube	5
2.1	Nettoyage des données	5
2.2	Configuration du Cube	6
2.3	Création des Hiérarchies	12
2.3.1	Hiérarchie Calendar (DimDate)	13
2.3.2	Hiérarchie Fiscal (DimDate)	13
2.3.3	Hiérarchie Territory (DimSalesTerritory)	13
3	Exploration du Cube sous Excel	14
3.1	Méthodologie	14
3.2	Résultats de l'Exploration	14
4	Création du Rapport Power BI	15
4.1	Chargement du Cube	15
4.2	Conception du Rapport	15
4.3	Analyse des Résultats	16
5	Conclusion	18
5.1	Synthèse des Résultats	18

Chapitre 1

Introduction

1.1 Contexte du Projet

Dans le cadre de l'ère du **Business Intelligence** et de la visualisation de données, l'analyse de grands volumes de données à travers des outils comme les datawarehouses et les cubes multidimensionnels est devenue essentielle pour extraire des insights exploitables. Ce projet s'inscrit dans cette dynamique en se concentrant sur l'analyse du datawarehouse **AdventureWorks**, en utilisant SQL Server Analysis Services (SSAS) pour créer un cube de données et Power BI pour visualiser les résultats.

Ce projet vise à appliquer les concepts de Business Intelligence appris en cours, en mettant en œuvre un pipeline complet allant de la vérification des données à la création de rapports interactifs, en passant par l'exploration des données avec Excel.

1.2 Présentation du Datawarehouse

Le datawarehouse **AdventureWorks** contient des données relatives aux ventes en ligne d'une entreprise fictive. La table de fait centrale est **FactInternetSales**, qui enregistre les ventes sur Internet, et les dimensions associées sont :

- **DimProduct** : Informations sur les produits
- **DimDate** : Données temporelles
- **DimCustomer** : Informations sur les clients
- **DimSalesTerritory** : Données géographiques des territoires de vente

Datawarehouse AdventureWorks

- **Table de Fait** : FactInternetSales
- **Dimensions** : DimProduct, DimDate, DimCustomer, DimSalesTerritory
- **Source** : Base de données AdventureWorks

1.3 Objectifs du Projet

Ce projet vise à accomplir les tâches suivantes :

1. Création d'un Cube :

- Construire un cube basé sur **FactInternetSales** avec les dimensions spécifiées
- Vérifier la santé des données
- Inclure des attributs spécifiques pour chaque dimension

2. Création de Hiérarchies :

- Définir des hiérarchies pour DimDate (Calendar, Fiscal) et DimSalesTerritory (Territory)
- Optimiser les hiérarchies pour des performances optimales

3. Exploration sous Excel :

- Utiliser des tableaux croisés dynamiques pour explorer le cube
- Exploiter les hiérarchies créées

4. Création de Rapport Power BI :

- Charger le cube dans Power BI
- Créer un rapport interactif avec graphiques et filtres

Approche Méthodologique :

1. Vérification et préparation des données
2. Création et configuration du cube SSAS
3. Définition des hiérarchies
4. Exploration des données via Excel
5. Développement et publication du rapport Power BI

1.4 Installation et préparation de l'environnement

1.4.1 Configuration nécessaire de visual studio

On doit vérifier que pour visual studio, les outils "Stockage et traitement des données" sont cochés. On doit également configurer avant une nouvelle règle firewall entrante (in-bound rule) permettant la connexion sur le port 2383 :

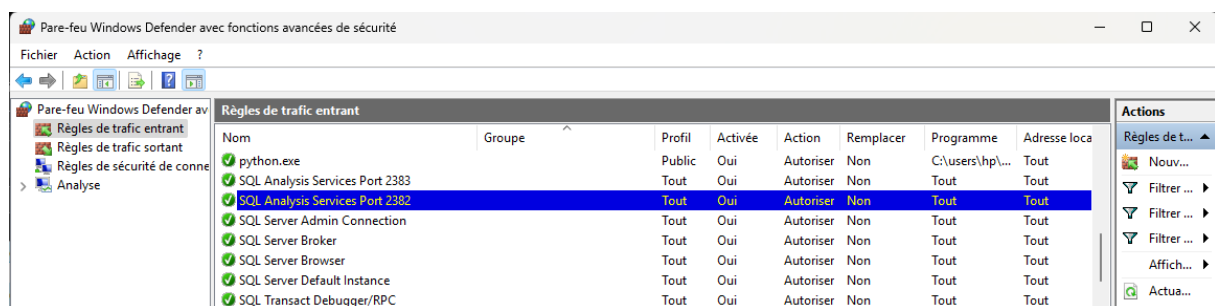


FIGURE 1.1 – Firewall configuration

On doit également préparer les serveurs dans microsoft SQL Server management, un serveur "Database engine" et un serveur portant le meme nom mais de type "analysis services".

On doit aussi vérifier dans le gestionnaire de configuration sql server que tous les services nécessaires sont en cours d'exécution :

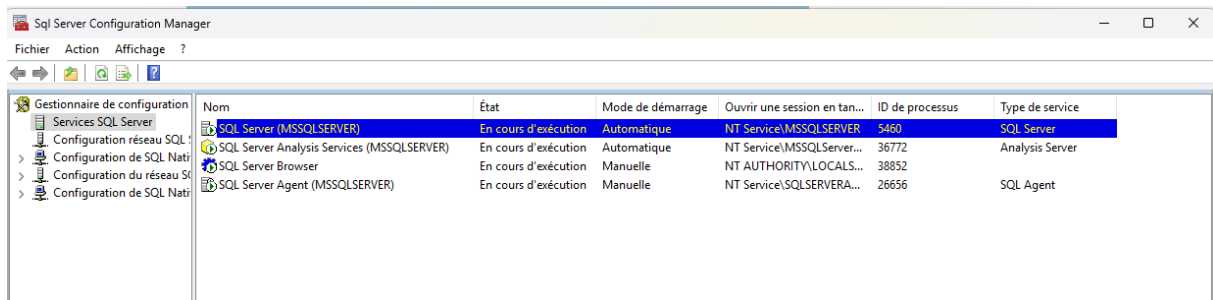


FIGURE 1.2 – SQL server configuration

Chapitre 2

Création du Cube

Un cube est une représentation des données sous forme multidimensionnelle. La création d'un cube donc nécessite une source de donnée qui va être la base de donnée Adventurework19 qu'on a restauré dans les TPs .

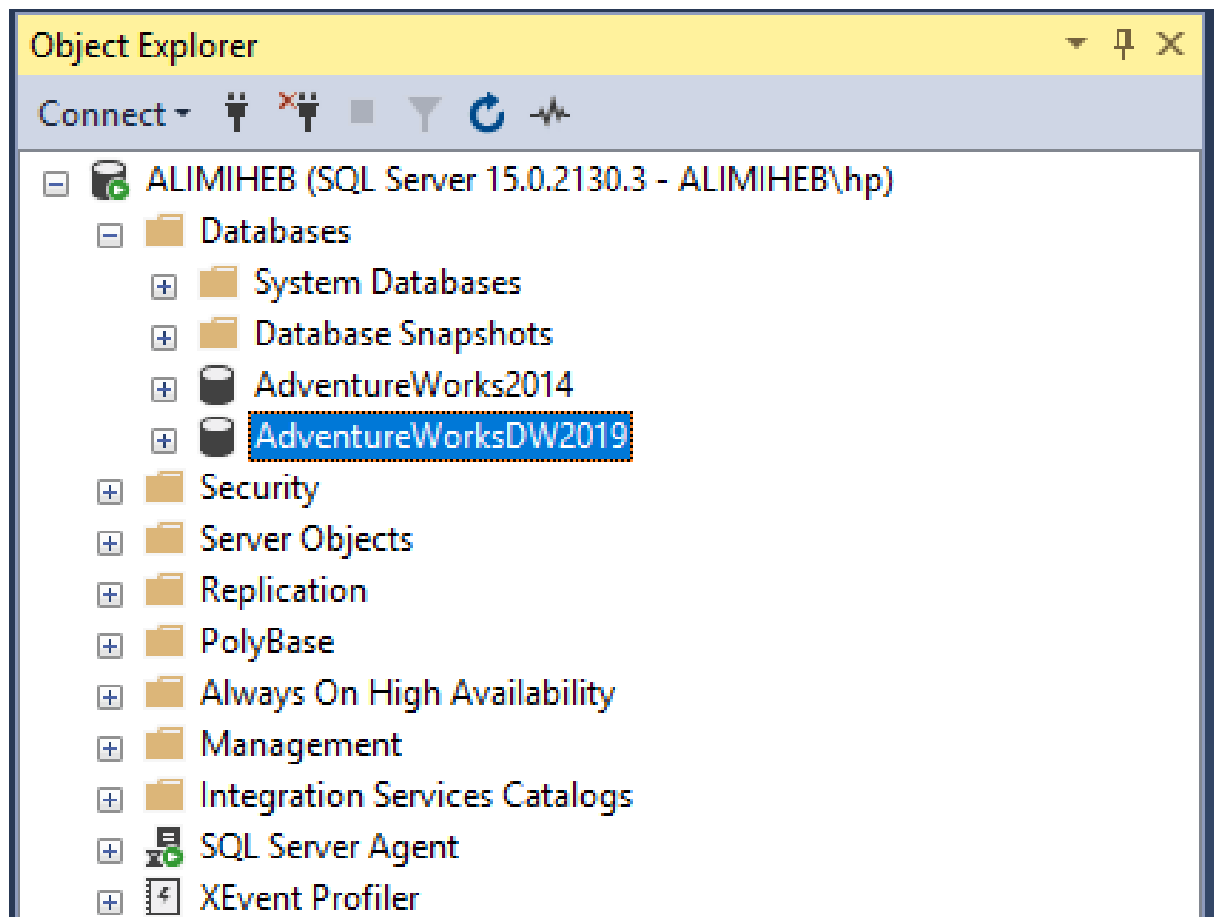


FIGURE 2.1 – database overview

2.1 Nettoyage des données

Le nettoyage des données est une étape fondamentale pour assurer la qualité des analyses générées par le cube SSAS AdvWokCube. Cette section se concentre sur le nettoyage de la colonne StandardCost dans la table DimProduct de la base de données AdventureWorksDW2019, utilisée pour évaluer les coûts des produits et calculer les marges bénéficiaires. Des anomalies dans cette colonne ont été identifiées et corrigées pour garantir des résultats fiables.

Certains produits dans DimProduct avaient un StandardCost non défini (NULL), ce qui pouvait entraîner des erreurs dans les calculs de marge bénéficiaire

Les valeurs négatives ont été corrigées en les remplaçant par la médiane du Standard-Cost de la sous-catégorie correspondante.



FIGURE 2.2 – database overview

2.2 Configuration du Cube

Le cube a été créé dans SQL Server Analysis Services (SSAS) avec la table de fait `FactInternetSales` et les dimensions suivantes, incluant les attributs spécifiés :

Attributs des Dimensions

Dimension	Attributs
DimProduct	Product Key, English Product Name, English Product Subcategory Name, Standard Cost
DimCustomer	Customer Key, Marital Status, Yearly Income
DimSalesTerritory	Sales Territory Key, Sales Territory Region, Sales Territory Country
DimDate	Date Key, English Month Name, Calendar Year

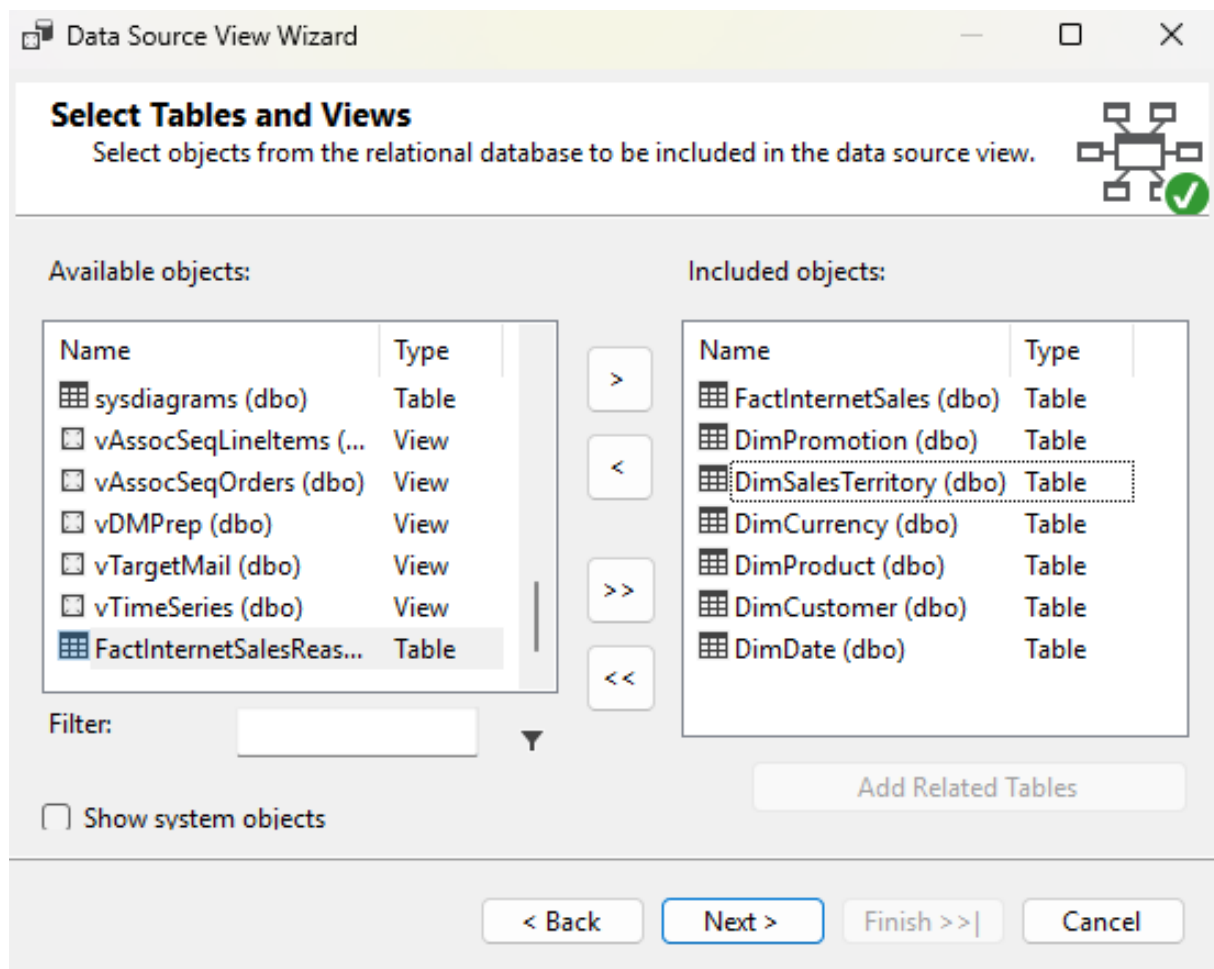


FIGURE 2.3 – Configuration des dimensions dans SSAS

La vue de ces dimensions donne :

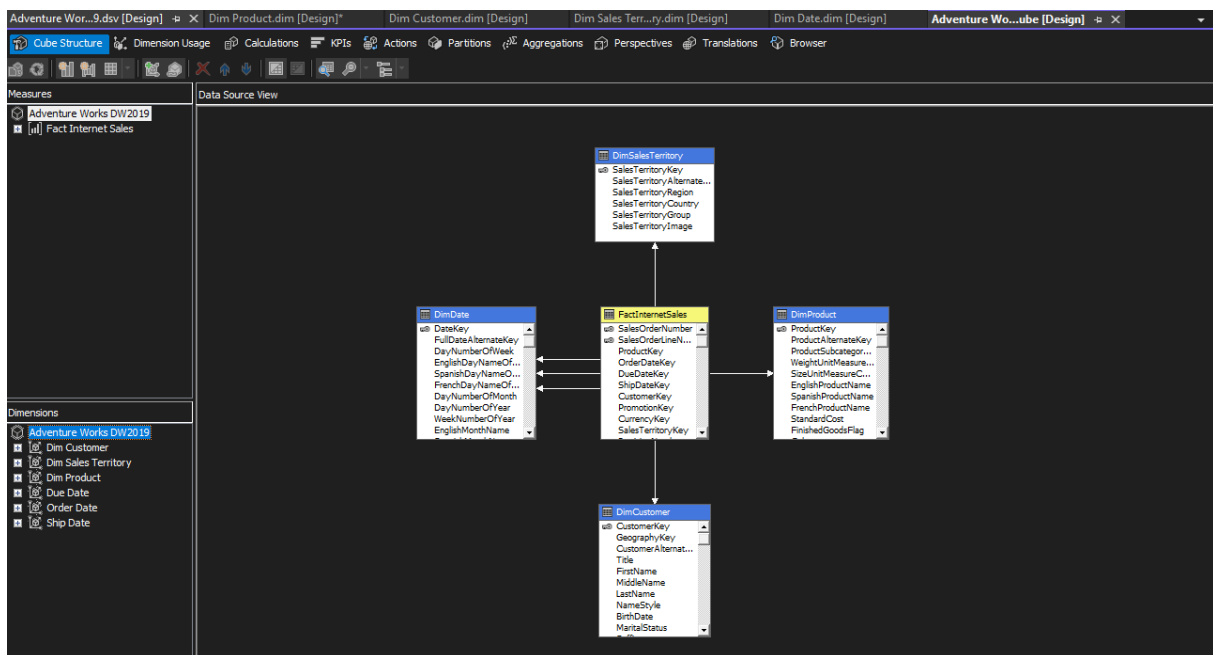


FIGURE 2.4 – Vue des dimensions dans la DSV

On sélectionne les tables de mesures et les dimensions :

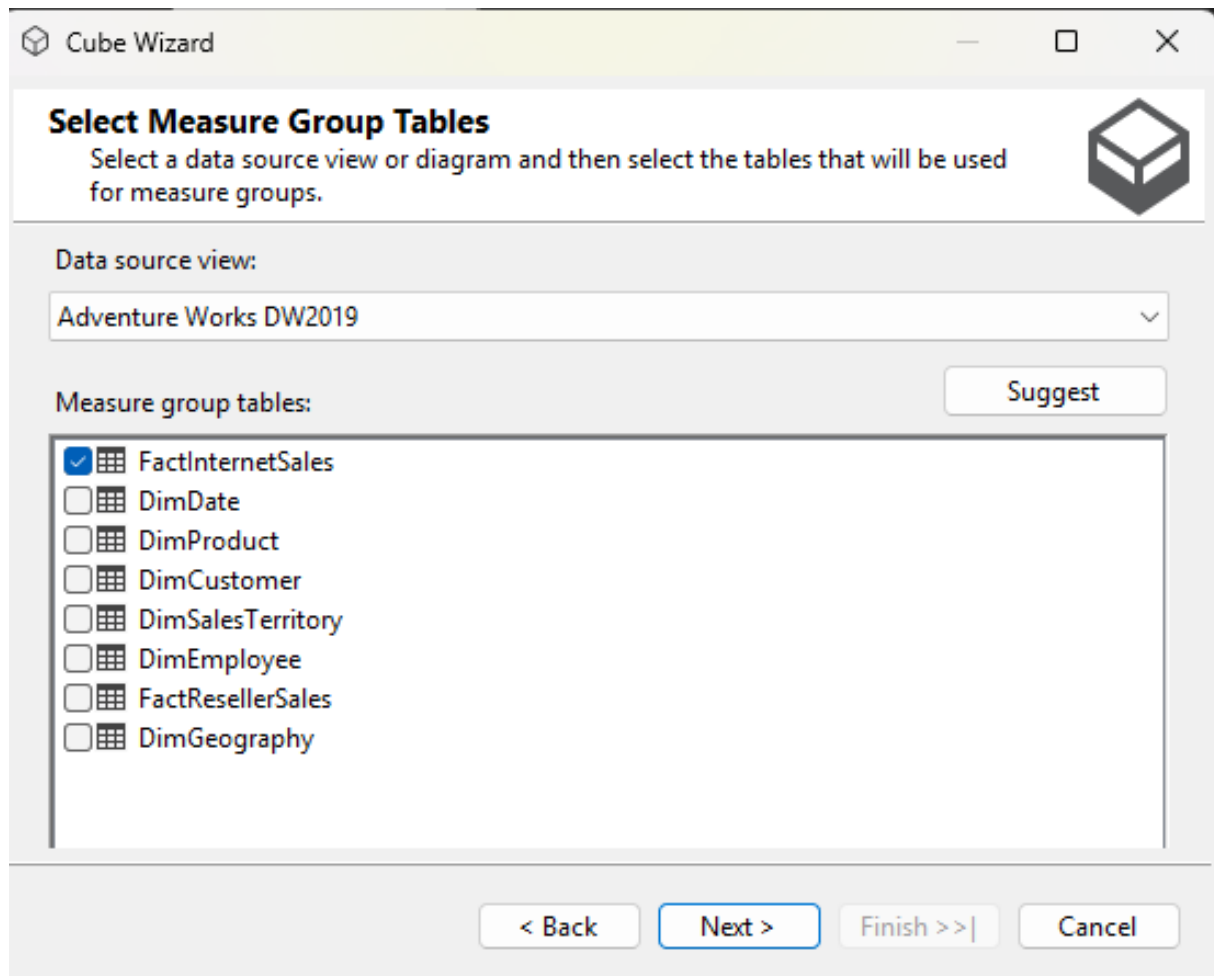


FIGURE 2.5 – Selection des tables de mesures

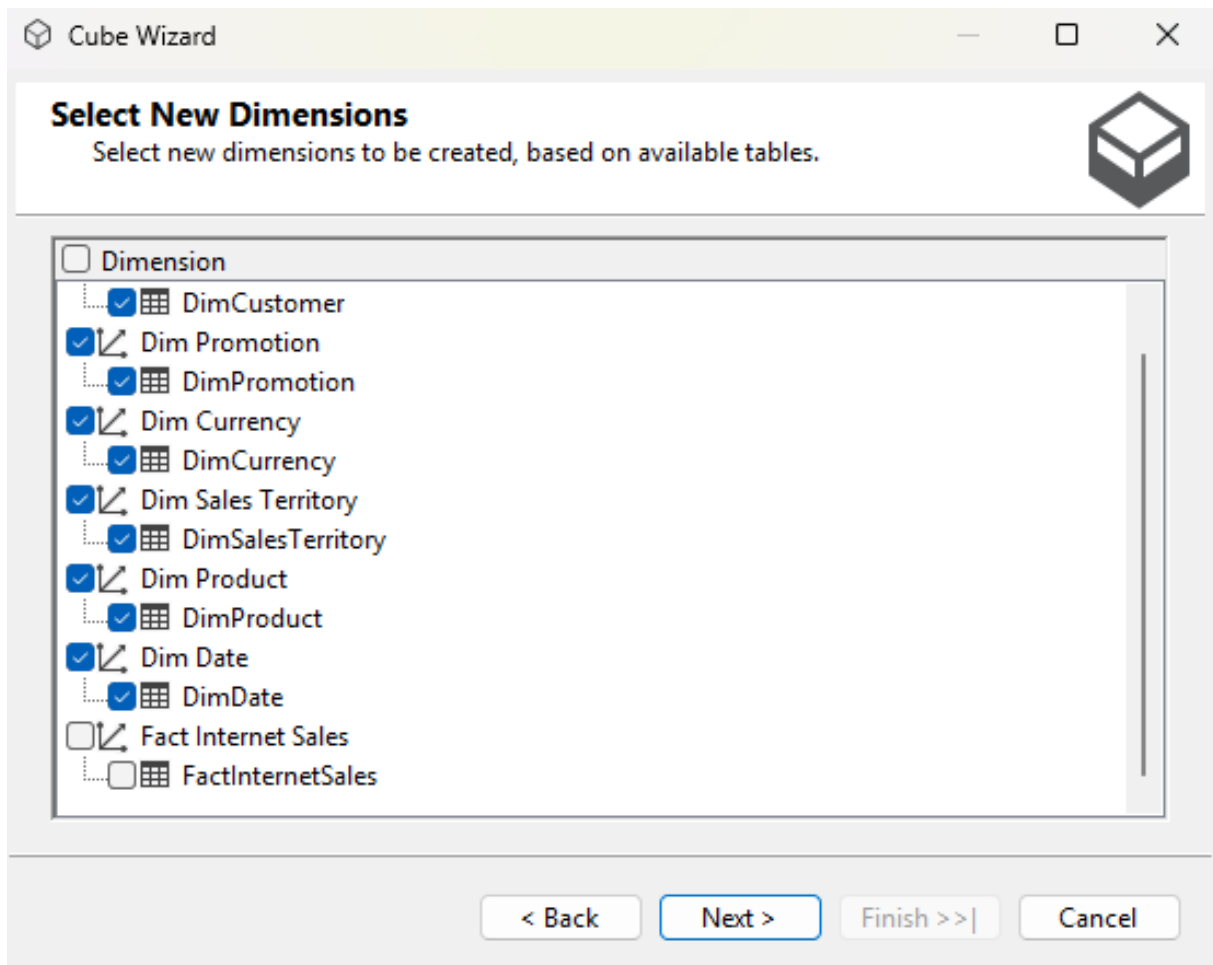


FIGURE 2.6 – Selection des dimensions

Maintenant il reste juste de déployer le cube sur le serveur :

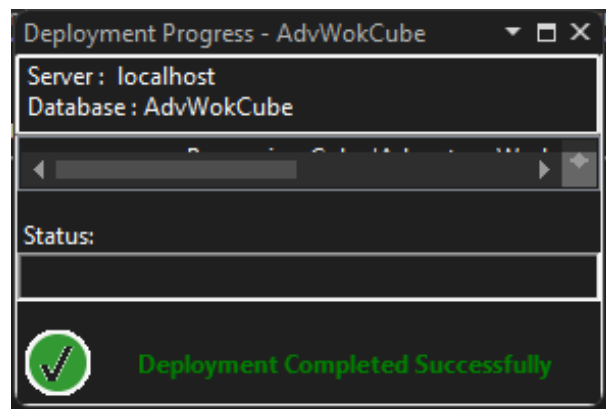


FIGURE 2.7 – Déploiement du cube sur le serveur SSAS

Maintenant on doit modifier les attributs de chaque dimensions.

Pour dimProduct on doit avoir English Product Subcategory Name pour cela on doit ajouter la table subcategory.

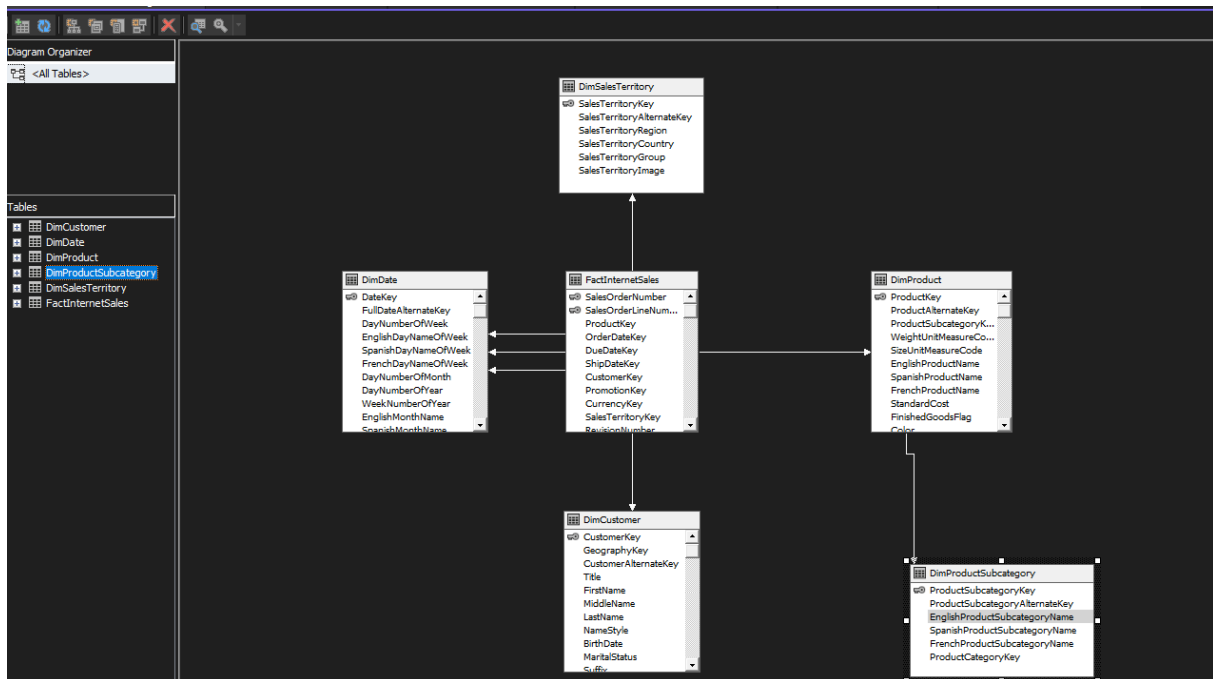


FIGURE 2.8 – Ajout de la table Subcategory pour DimProduct

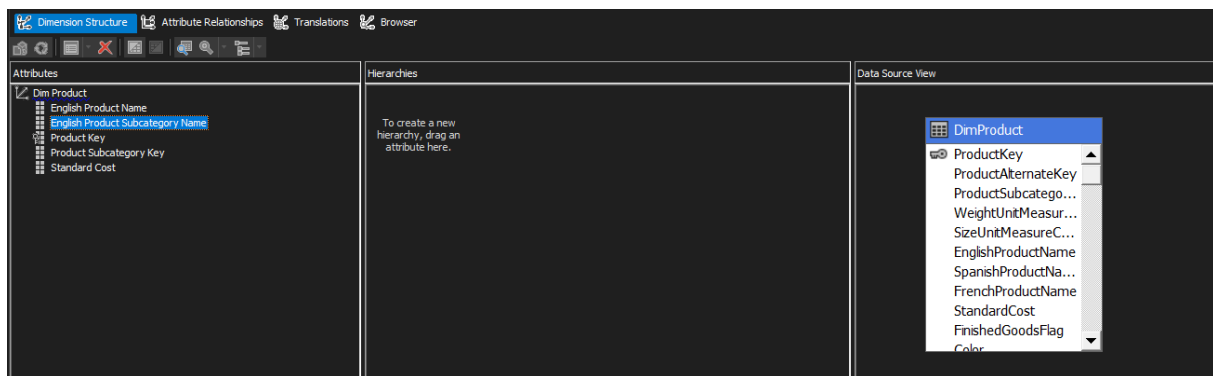


FIGURE 2.9 – Configuration de l'attribut Subcategory dans DimProduct

Pour le reste il suffit d'ajouter les attribut dans chaque dimension

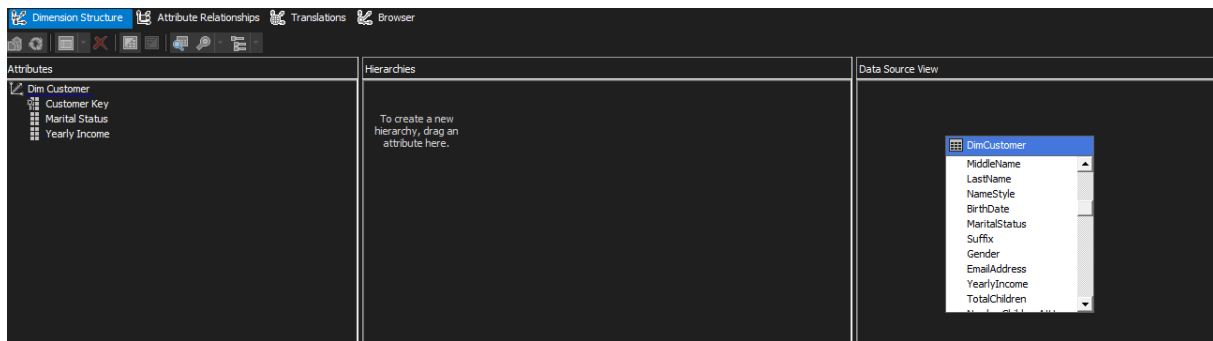


FIGURE 2.10 – Ajout d'attributs au dimension Customer

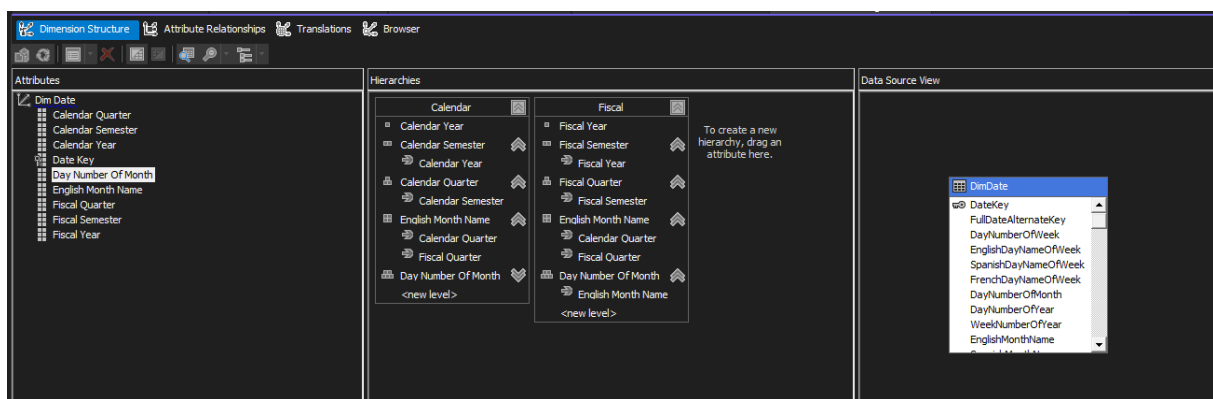


FIGURE 2.11 – Ajout d'attributs au dimension Date

2.3 Création des Hiérarchies

Une fois le cube est déployé on va maintenant s'intéresser à la création des hiérarchies qui vont nous aider à l'interprétation de nos résultats.

une hiérarchie est une structure de données organisée de manière hiérarchique, avec des niveaux qui représentent des catégories de données qui se succèdent les unes aux autres selon une relation parent-enfant. Les hiérarchies sont utilisées pour organiser et structurer les données de manière logique et compréhensible, afin de faciliter l'analyse et la prise de décision.

Trois hiérarchies ont été définies pour faciliter l'analyse multidimensionnelle :

2.3.1 Hiérarchie Calendar (DimDate)

- **Niveaux** : CalendarYear, CalendarSemester, CalendarQuarter, MonthCalendar, DayNumberOfMonth
- **Objectif** : Permettre une analyse temporelle basée sur le calendrier civil

2.3.2 Hiérarchie Fiscal (DimDate)

- **Niveaux** : FiscalYear, FiscalSemester, FiscalQuarter, MonthCalendar, DayNumberOfMonth
- **Objectif** : Analyser les ventes selon l'année fiscale

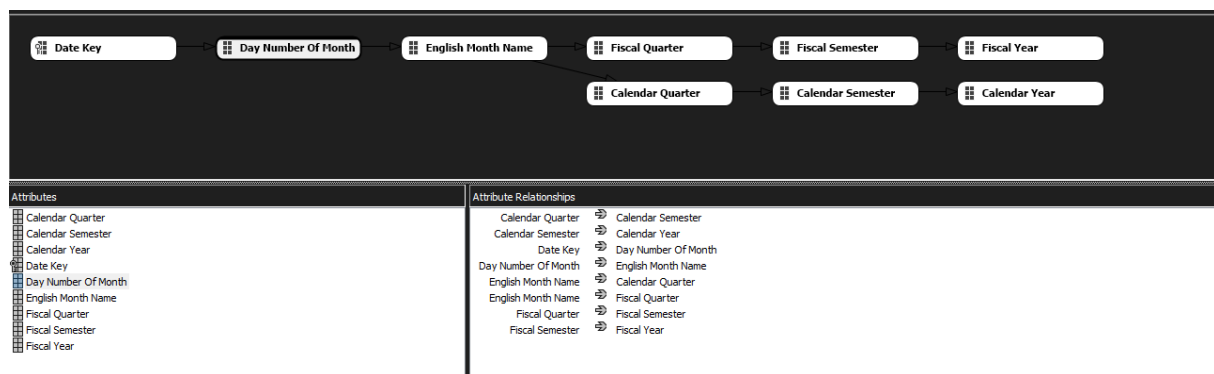


FIGURE 2.12 – Creation des hiérarchies Calendar et Fiscal

2.3.3 Hiérarchie Territory (DimSalesTerritory)

- **Niveaux** : Sales Territory Country, Sales Territory Region
- **Objectif** : Explorer les ventes par zone géographique

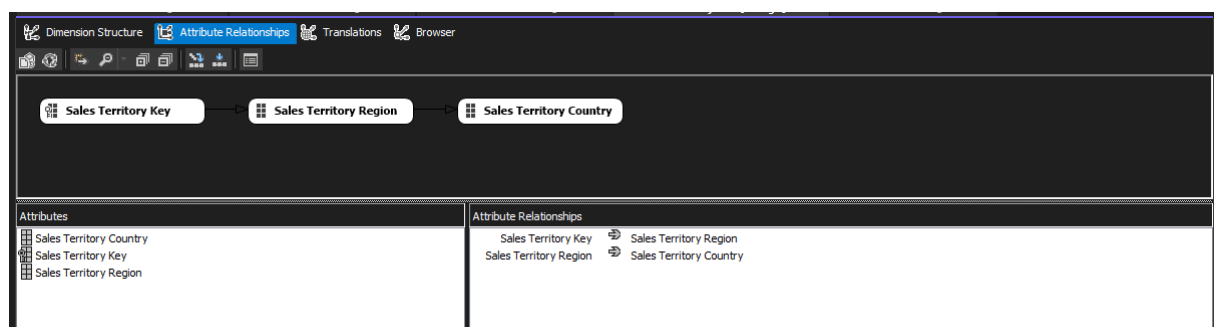


FIGURE 2.13 – Creation d'hiérarchies Territory

Chapitre 3

Exploration du Cube sous Excel

3.1 Méthodologie

Le cube a été exploré à l'aide de tableaux croisés dynamiques dans Excel, en se connectant directement au serveur SSAS. Les hiérarchies créées ont été utilisées pour structurer les analyses.

On va analyser notre cube au niveau de excel.

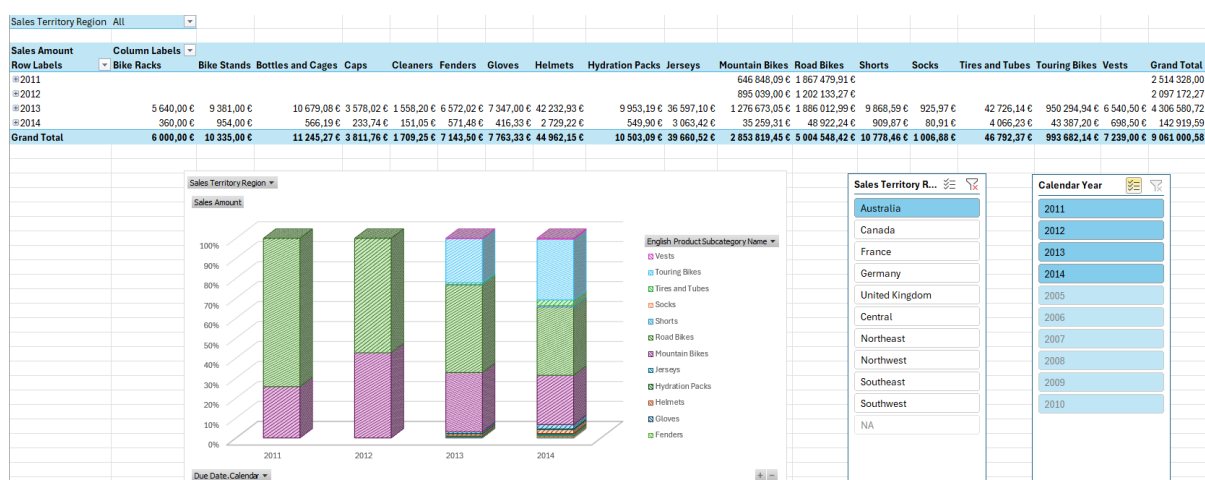


FIGURE 3.1 – PivotTable

3.2 Résultats de l'Exploration

Tendance des ventes : Les ventes totales augmentent légèrement de 2011 (2,138,298.00) à 2014 (2,729,002.14), avec un pic en 2013 (4,492,373.34) si toutes les régions sont incluses.

Distribution par sous-catégorie : Les sous-catégories comme Mountain Bikes et Road Bikes dominent les ventes (fortes proportions dans les barres), tandis que des catégories comme Gloves ou Fenders ont une contribution moindre.

Impact régional : Le filtre Sales Territory Region permet d'isoler les performances par région (par exemple, North America ou Europe), mais la capture montre "All", donc une vue globale. Filtre temporel : Le slicer Calendar Year permet une analyse dynamique (par exemple, sélectionner 2013 montre 4,492,373.34).

Chapitre 4

Création du Rapport Power BI

4.1 Chargement du Cube

Le cube SSAS a été importé dans Power BI via une connexion directe, sans charger l'ensemble du datawarehouse AdventureWorks. Cela garantit des performances optimales et une analyse ciblée.

4.2 Conception du Rapport

Le rapport Power BI a été conçu avec des visualisations interactives, incluant :

- **Graphique en barres** : Ventes par région (hiérarchie Territory)
- **Graphique en ligne** : Évolution des ventes par mois (hiérarchie Calendar)
- **Matrice** : Ventes par sous-catégorie de produit et année
- **Filtres** : Année, région, statut marital des clients, revenu annuel

Dans ce rapport on va afficher l'attribut sales ammount en fonction des autres attributs Comme suivant :

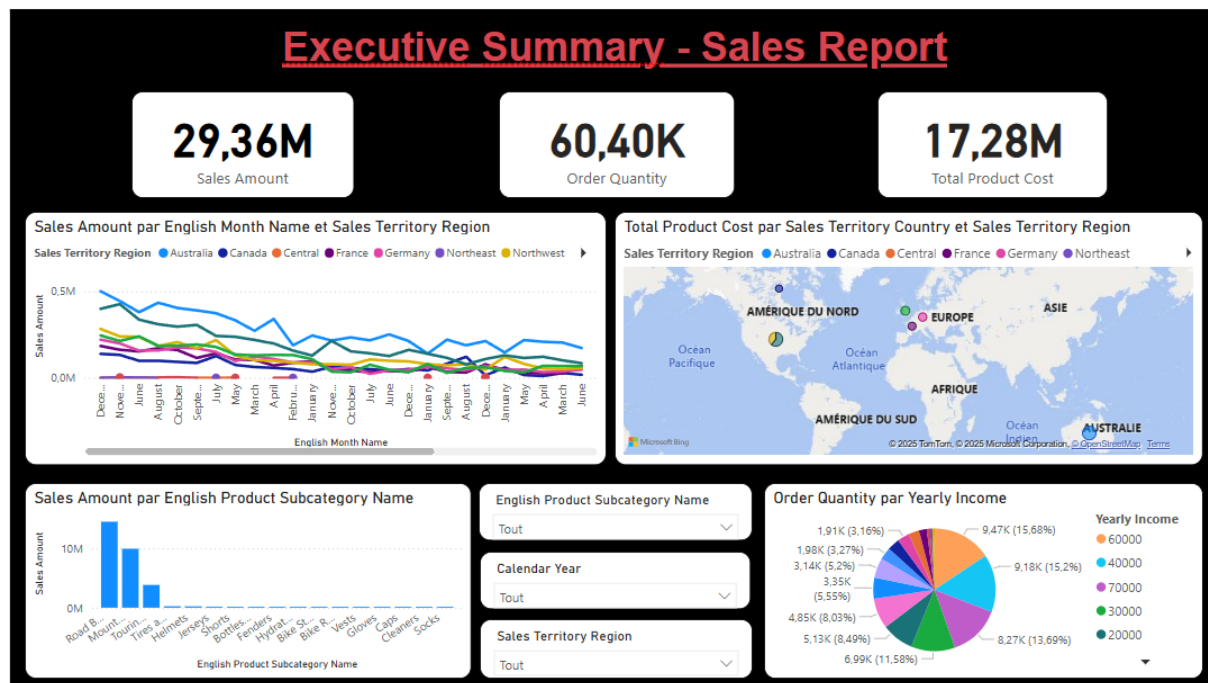
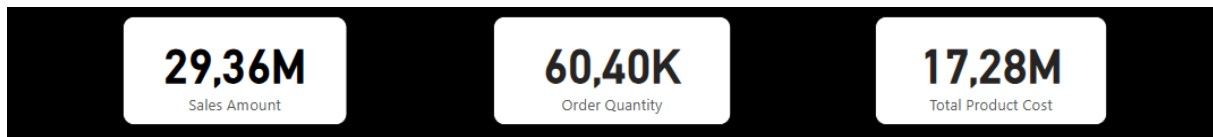


FIGURE 4.1 – Structure générale du tableau de bord

4.3 Analyse des Résultats

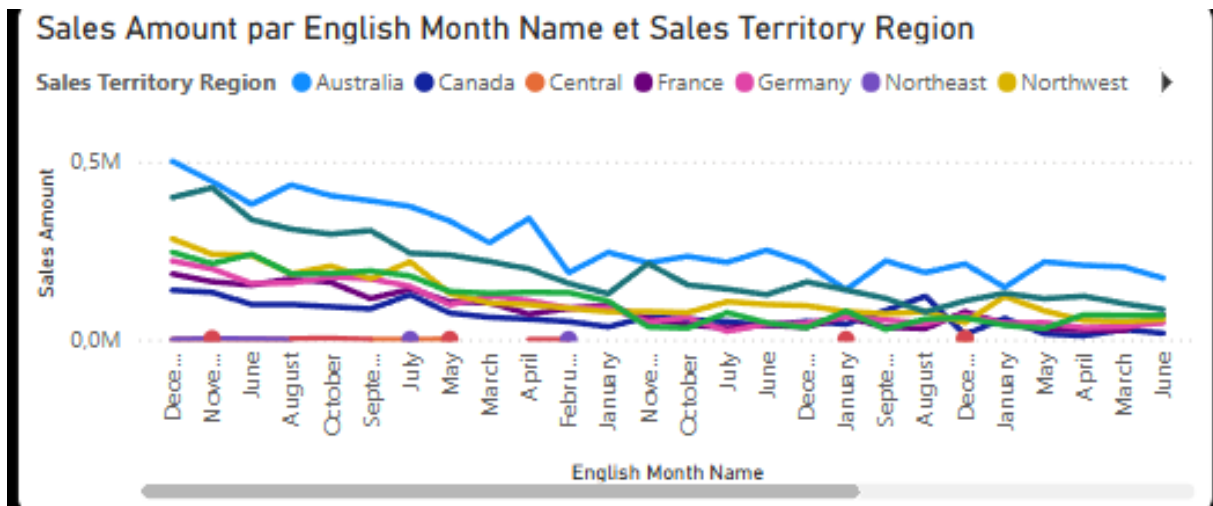
Le tableau de bord suit une mise en page professionnelle avec les éléments suivants :



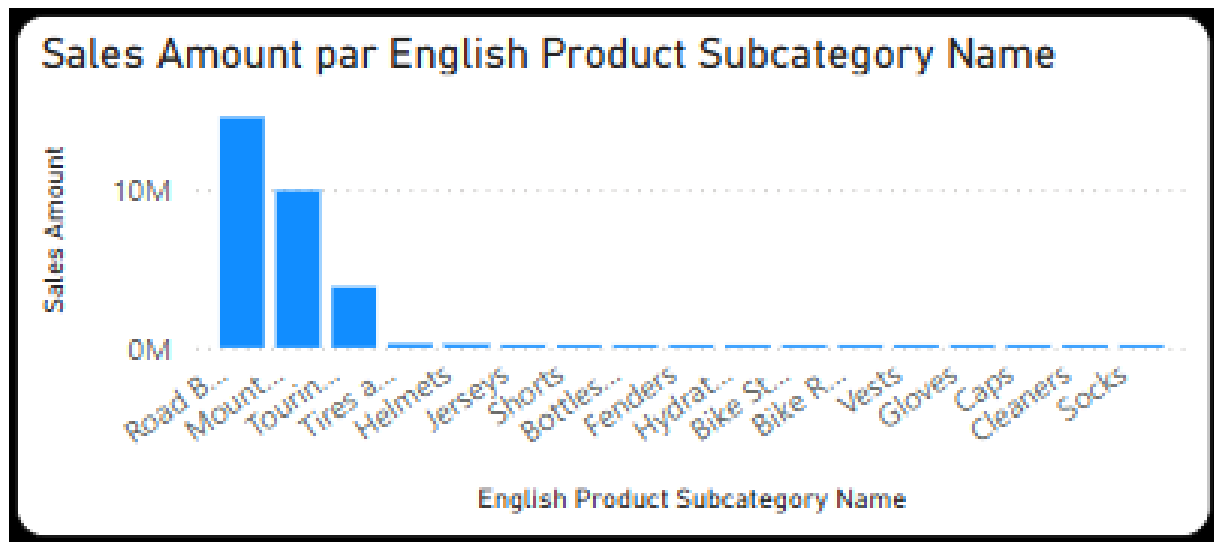
Total Sales Amount (12,568,473) : Montre le total des ventes pour les filtres actuels (toutes les années, toutes les régions).

Total Order Quantity (29,358) : Indique le volume des commandes, utile pour évaluer l'activité.

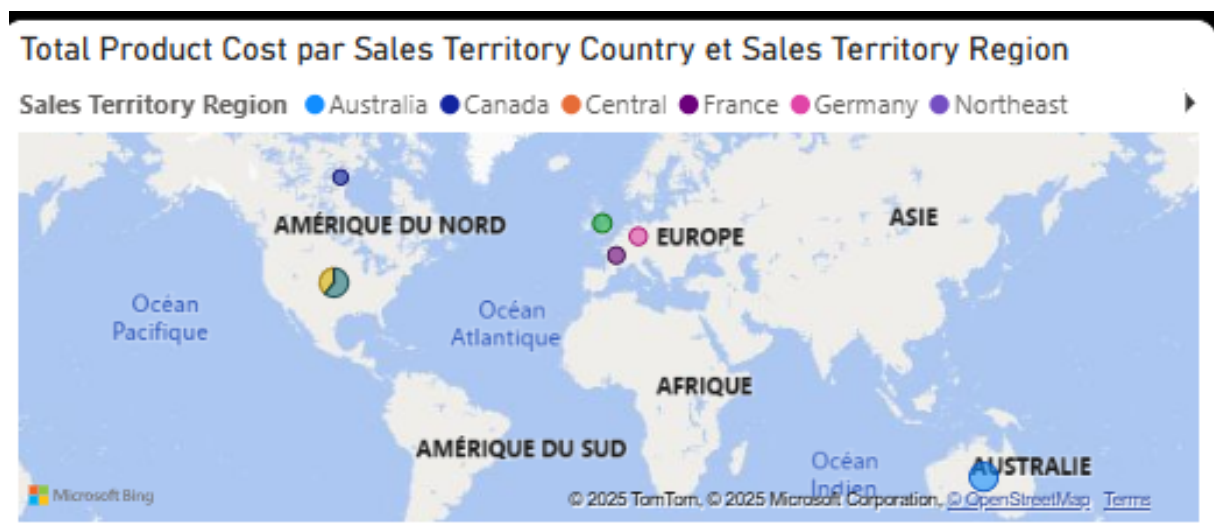
Insight : Ces métriques sont claires et bien formatées (en devise ou unités), mais pourraient bénéficier d'une comparaison temporelle (par exemple, variation par rapport à l'année précédente).



Les ventes globales augmentent de 2011 à 2013, avec un pic en 2013 (4,492,373), puis une légère baisse en 2014 (2,729,002). North America domine les ventes, représentant environ 50-60. Europe montre une croissance constante, tandis qu'Australie reste stable mais faible.



Mountain Bikes et Road Bikes dominent, représentant environ 70 % des ventes. Les sous-catégories comme Gloves ou Fenders ont des ventes négligeables (à 5 %).



North America est la région la plus performante (bleu foncé, environ 7M). Europe suit avec 3M, tandis qu'Australie est à 1M. La carte réagit bien au slicer Calendar Year, montrant une baisse en 2014 pour toutes les régions.

Chapitre 5

Conclusion

Ce projet démontre la puissance des outils de Business Intelligence pour analyser les données de ventes et fournir des insights exploitables. L'utilisation combinée de SSAS, Excel et Power BI a permis de créer un pipeline d'analyse robuste, ouvrant la voie à des applications pratiques et des améliorations futures dans la gestion des données d'entreprise.

5.1 Synthèse des Résultats

Principales Réalisations :

- **Création du Cube :**
 - Cube SSAS basé sur **FactInternetSales** avec quatre dimensions
 - Données vérifiées et attributs configurés selon les spécifications
- **Hiérarchies :**
 - Hiérarchies Calendar, Fiscal et Territory créées et optimisées
 - Performances améliorées grâce à la résolution des avertissements SSAS
- **Exploration Excel :**
 - Analyse des ventes par région, période et produit via tableaux croisés
- **Rapport Power BI :**
 - Visualisations interactives des ventes par région, période et profil client
 - Publication du rapport dans Power BI App