



Université Sultan Slimane
École Nationale des Sciences Appliquées (ENSA) de Béni-Mellal
Transformation Digitale Industrielle

Rapport dans le cadre du module : **Pilotage de la Performance et Tableaux de Bord**

Sous le thème :

Exploration et Analyse de AdventureWorks :
<<Achats et Fournisseurs >>

Encadré par : Pr. BE.ELBAGHAZAOU

Réalisé par :

HASNAOUI Siham

OUSSIDDAN Majda



Remerciements

Nous tenons à exprimer notre profonde gratitude à notre professeur encadrant, Pr. BE. ELBAGHAZAOUI, pour son précieux soutien tout au long de ce projet. Son expertise, sa disponibilité et son implication constante ont été des éléments clés pour mener à bien cette étude. Grâce à ses conseils avisés, nous avons pu surmonter les défis rencontrés et affiner nos compétences en matière d'analyse de données et d'utilisation des outils technologiques. Sa passion pour le partage des connaissances et son approche pédagogique ont été une source d'inspiration qui nous a motivés à donner le meilleur de nous-mêmes. Nous sommes également reconnaissants pour les discussions constructives et les retours détaillés qui ont permis d'améliorer la qualité de notre travail. Ce projet a été une opportunité exceptionnelle d'apprentissage et de développement personnel, et cela n'aurait pas été possible sans l'accompagnement bienveillant et rigoureux de notre encadrant.

Cette expérience restera gravée dans nos mémoires comme un moment fort de notre parcours académique, et nous sommes profondément reconnaissants envers notre professeur pour son accompagnement exemplaire. Nous espérons que les enseignements tirés de ce projet nous permettront d'exceller dans nos projets futurs et de contribuer efficacement à la mise en œuvre de solutions innovantes dans le monde professionnel. Merci infiniment, Pr. BE. ELBAGHAZAOUI, pour votre encadrement exceptionnel et pour avoir été une source d'inspiration tout au long de ce travail.



Table des matières

Introduction.....	6
Chapitre I. CONTEXT GENERAL DU PROJET	7
Introduction.....	8
Objectifs du Projet.....	9
Présentation de la Base de Données AdventureWorks	9
Découverte des Tables:.....	9
Achats (Purchasing)	9
Production (Production)	10
Ventes (Sales)	10
Ressources Humaines (HumanResources)	11
Personne (Person)	11
Choix du Module Purchasing.....	11
Rôle Stratégique du Module Purchasing	11
Tables Clés du Module Purchasing	12
Opportunités Analytiques Offertes par Purchasing.....	12
Importance du Module dans le Projet.....	13
Relations et Intégration des Données	13
Relations Clés dans le Module Purchasing	13
Conclusion	15
Chapitre 2. Outils et Méthodologie.....	17
Introduction.....	18
Environnement de travail :	18
SQL Server Management Studio (SSMS)	18
Power BI.....	18
KNIME.....	18
Méthodologie.....	19
Installation et Configuration de SQL Server :	19
Récupération des Données :.....	19
nettoyage des données avec KNIME	20
Nettoyage et Transformation des Données avec Power Query dans Power BI.....	22
Relation entre les tables	24
Conclusion	25
Chapitre 3.....	26
Résultats et visualisations	26



Introduction.....	27
Visualisation	28
Page1	28
Page2	30
Page3	31
Page 4	32
Page5	33
Conclusion	34
Conclusion générale	35



Table des Figure

Figure 1:AdventureWorks schéma	16
Figure 2:AdventureWorks Dans Explorateur d'Objets dans SSMS	19
Figure 3:Fenêtre de Connexion à SQL Server	19
Figure 4:Exportation des Données depuis SQL Server vers un Fichier CSV	20
Figure 5:Workflow de Nettoyage et Préparation des Données avec KNIME	22
Figure 6:Power Query pour le Nettoyage	23
Figure 7:vue modèle	24
Figure 8:Vision Stratégique Globale et Analyse des Performances.....	28
Figure 9:Aperçu des commandes	30
Figure 10:Détails des produits	31
Figure 11:fournisseurs	32
Figure 12:Analyse des coûts	33

Introduction

Dans un contexte où les entreprises s'appuient de plus en plus sur les données pour éclairer leurs décisions stratégiques, l'analyse et la visualisation des données jouent un rôle central dans l'identification des opportunités et l'optimisation des performances. Le projet présenté dans ce rapport s'inscrit dans cette dynamique et repose sur l'exploitation de la base de données AdventureWorks, un modèle de référence utilisé dans de nombreux contextes académiques et professionnels pour simuler des scénarios réels.

L'objectif principal de ce projet est de démontrer une approche complète allant de l'exploration initiale des données brutes jusqu'à leur transformation en informations visuelles pertinentes, exploitables par des décideurs. À travers ce travail, nous visons à :

Explorer et comprendre la structure de la base de données AdventureWorks, avec un focus particulier sur le module Purchasing.

Nettoyer et préparer les données afin d'assurer leur qualité et leur cohérence.

Créer des visualisations interactives pour extraire des insights clés et répondre à des questions stratégiques concernant les fournisseurs, les coûts, et les performances des produits.

Ce rapport s'articule autour de trois chapitres principaux :

- Présentation du projet, qui décrit les objectifs, le contexte et la méthodologie adoptée.
- Outils utilisés, qui explique les outils technologiques mobilisés (KNIME, SQL Server, Power BI) et leurs apports respectifs dans le projet.
- Visualisations et résultats, où nous présentons les différentes visualisations interactives réalisées, accompagnées d'analyses détaillées des résultats obtenus.

En synthétisant ces étapes, ce projet met en lumière l'importance de disposer d'un processus rigoureux pour extraire la valeur des données. Il constitue également une base solide pour des analyses futures, applicables à des cas réels et complexes.



CHAPITRE I. CONTEXT GENERAL DU PROJET



Description du chapitre :

Le chapitre Présentation du Projet explore les objectifs et le contexte de l'analyse basée sur la base de données AdventureWorks, axée sur les processus d'achats, de production et de gestion des fournisseurs. Il met en lumière les raisons du choix d'AdventureWorks, en détaillant la structure et les relations des données, avec un focus particulier sur le module Purchasing





Introduction

L'analyse et la gestion des données constituent aujourd'hui un enjeu majeur pour les organisations cherchant à optimiser leurs processus décisionnels. Le projet présenté dans ce rapport s'appuie sur la base de données AdventureWorks, une solution de simulation largement reconnue, qui permet de modéliser divers processus d'une entreprise, notamment ceux liés aux achats, à la production et à la gestion des fournisseurs.

Dans ce chapitre, nous exposerons les objectifs et le contexte général du projet, en mettant l'accent sur les problématiques adressées et les raisons ayant motivé le choix de travailler avec AdventureWorks. Nous aborderons également la structure et les relations des données utilisées, notamment celles liées au module Purchasing, qui constitue le cœur de notre analyse. Cette exploration initiale constitue la fondation de notre travail, en orientant les étapes de nettoyage, de transformation et de visualisation des données.

Ce chapitre vise donc à établir une compréhension claire du cadre du projet, en précisant les bases méthodologiques et en introduisant les principaux axes d'analyse abordés dans les sections suivantes.

Objectifs du Projet

Dans un monde où les décisions stratégiques reposent de plus en plus sur l'analyse des données, la capacité à collecter, nettoyer et visualiser ces données est devenue cruciale pour les organisations. Ce projet s'inscrit dans cette logique, en exploitant la base de données AdventureWorks pour analyser et comprendre les dynamiques internes d'une entreprise fictive spécialisée dans la fabrication et la vente de produits liés au cyclisme. AdventureWorks est largement reconnue comme une base de données de simulation idéale pour l'apprentissage et les projets analytiques, grâce à sa richesse de données et sa structure bien conçue.

L'objectif principal de ce projet est de démontrer comment une entreprise peut exploiter ses données pour optimiser ses processus d'achat, de production et de gestion des fournisseurs. Plus spécifiquement, les objectifs incluent :

Explorer et comprendre les données disponibles : identifier les relations entre les différents modules, en particulier ceux liés aux achats (Purchasing) et à la production (Production).

Nettoyer et structurer les données : éliminer les anomalies, corriger les incohérences et transformer les données pour les rendre exploitables.

Analyser les processus opérationnels clés : évaluer les performances des fournisseurs, identifier les produits les plus rentables, et examiner les coûts associés aux expéditions.

Créer des visualisations percutantes : présenter les résultats de manière claire et accessible pour aider à la prise de décision.

Ces objectifs orientent l'ensemble des étapes du projet, depuis la collecte initiale des données jusqu'à leur transformation en insights exploitables, en passant par le nettoyage et l'analyse.

Présentation de la Base de Données AdventureWorks

AdventureWorks est une base de données relationnelle développée par Microsoft, souvent utilisée comme modèle pour illustrer la gestion de données et les processus analytiques dans les contextes académiques et professionnels. Elle modélise une entreprise fictive spécialisée dans la fabrication, la distribution et la vente de vélos, accessoires, et vêtements associés. AdventureWorks est particulièrement utile pour simuler des scénarios métiers réels, offrant une plateforme robuste pour explorer les techniques d'analyse, de visualisation et de gestion des données.

Découverte des Tables:

Achats (Purchasing)

Ce module regroupe toutes les informations relatives aux fournisseurs, aux commandes passées, aux coûts d'expédition, ainsi qu'aux quantités reçues ou rejetées. Les principales tables comprennent :

Vendor : Fournit les détails sur les fournisseurs, y compris leur nom, statut et adresse.

PurchaseOrderHeader : Contient les en-têtes des commandes passées auprès des fournisseurs, y compris les informations sur les coûts, les taxes et les délais.

PurchaseOrderDetail : Décrit les articles commandés, leur quantité, leur coût unitaire, et les remises éventuelles.

ShipMethod : Fournit les méthodes d'expédition disponibles et les coûts associés.

Ce module est essentiel pour analyser la performance des fournisseurs et optimiser les processus d'approvisionnement.

Production (Production)

Le module Production se concentre sur les produits fabriqués, leur gestion et les opérations liées à la production. Les principales tables incluent :

Product : Contient des informations sur les produits, comme leurs noms, catégories, et coûts de fabrication.

ProductInventory : Suit les niveaux de stock des produits dans les différents entrepôts.

BillOfMaterials : Liste les composants nécessaires pour fabriquer un produit.

WorkOrder : Gère les ordres de fabrication, détaillant les étapes de production et les ressources nécessaires.

TransactionHistory : Enregistre les mouvements de stock, tels que les entrées et sorties d'inventaire.

Ce module permet de suivre les performances des processus de production et d'optimiser la gestion des stocks.

Ventes (Sales)

Ce module regroupe les informations sur les clients, les commandes de vente et les performances des commerciaux. Les principales tables comprennent :

Customer : Fournit des données sur les clients, comme leurs coordonnées et leurs préférences d'achat.

SalesOrderHeader et SalesOrderDetail : Détail des commandes de vente, incluant les montants, les dates et les produits commandés.

SalesPerson : Contient les données sur les commerciaux, leurs objectifs et performances.

SalesTerritory : Répartit les zones géographiques de vente pour une analyse des tendances régionales.

Ces tables permettent d'examiner les tendances des ventes, les produits les plus populaires et les performances des équipes commerciales.

Ressources Humaines (HumanResources)

Ce module est dédié à la gestion des employés et de leurs affectations. Les tables incluent :

Employee : Détaille les informations sur les employés, comme leurs rôles, salaires et dates d'embauche.

Department : Liste des départements au sein de l'entreprise.

EmployeeDepartmentHistory : Suivi des affectations des employés aux départements au fil du temps.

Ce module soutient les analyses liées à la gestion des ressources humaines et à la répartition des effectifs.

Personne (Person)

Ce module transversal regroupe les informations sur les contacts et les adresses. Les principales tables incluent :

Person : Contient les informations générales sur les personnes (clients, fournisseurs, employés).

Address : Gère les adresses associées aux contacts.

EmailAddress et **PhoneNumber** : Fournissent des détails complémentaires sur les contacts.

Ce schéma facilite la consolidation des informations relatives aux adresses et aux communications.

Relations et Intégration des Données

Choix du Module Purchasing

Le module Purchasing a été sélectionné comme cœur de l'analyse en raison de son rôle crucial dans les processus d'approvisionnement et de gestion des fournisseurs. Ce module offre une vue complète sur les transactions entre l'entreprise et ses partenaires fournisseurs, permettant d'évaluer leur performance, d'identifier des opportunités d'optimisation et de garantir l'efficacité des processus d'achat.

Rôle Stratégique du Module Purchasing

L'analyse des données d'achat est essentielle pour toute entreprise cherchant à réduire ses coûts, à améliorer ses relations avec les fournisseurs et à optimiser ses méthodes d'approvisionnement. Le module Purchasing fournit une structure détaillée et des informations riches qui couvrent plusieurs aspects clés :

Gestion des commandes : Suivi des quantités commandées, des coûts unitaires et des délais de livraison.

Évaluation des fournisseurs : Analyse des performances des fournisseurs à travers les taux de rejet, les coûts associés et la fiabilité des livraisons.

Optimisation des coûts : Compréhension des frais d'expédition, des taxes appliquées et des coûts totaux liés à l'approvisionnement.

Tables Clés du Module Purchasing

PurchaseOrderHeader : Cette table contient les informations globales des commandes passées auprès des fournisseurs. Elle inclut des données essentielles telles que :

Le coût total de chaque commande, incluant les taxes et les frais d'expédition.

Les dates de commande, d'approbation et de livraison prévues, permettant d'évaluer les délais et la planification.

La méthode d'expédition utilisée, cruciale pour analyser les coûts liés au transport.

PurchaseOrderDetail : Cette table fournit des informations détaillées sur chaque article commandé, en complément de la table PurchaseOrderHeader. Elle contient :

Les quantités commandées et reçues, permettant de mesurer l'efficacité des fournisseurs.

Les coûts unitaires et les sous-totaux pour chaque article, offrant une granularité dans l'analyse des coûts.

Les taux de rejet pour chaque article, qui permettent d'évaluer la qualité des produits livrés.

Vendor : La table des fournisseurs répertorie les informations clés sur les partenaires de l'entreprise, y compris :

Les coordonnées des fournisseurs pour les suivis et les négociations.

Le statut des fournisseurs, permettant de distinguer les partenaires actifs, potentiels ou inactifs.

Les informations sur leur fiabilité, basées sur les délais de livraison et les quantités acceptées/rejetées.

ShipMethod : Cette table recense les méthodes d'expédition disponibles et leurs coûts associés, offrant une base pour analyser et comparer les options de transport utilisées par l'entreprise.

Opportunités Analytiques Offertes par Purchasing

En exploitant les données du module Purchasing, plusieurs problématiques stratégiques peuvent être abordées, notamment :

Évaluation des fournisseurs : Identifier les fournisseurs les plus fiables et les plus performants, en fonction de leurs délais, de leurs taux de rejet et de leurs coûts.

Analyse des coûts : Repérer des opportunités de réduction des coûts grâce à une meilleure compréhension des frais d'expédition, des taxes et des coûts unitaires.

Optimisation des processus d'approvisionnement : Identifier les inefficacités dans les méthodes d'achat et d'expédition, et proposer des solutions pour améliorer la chaîne d'approvisionnement.

Importance du Module dans le Projet

Le choix du module Purchasing est justifié par son impact direct sur les performances opérationnelles et financières de l'entreprise. En se concentrant sur ce module, le projet vise à :

Améliorer les relations avec les fournisseurs en identifiant les partenaires stratégiques à long terme.

Réduire les coûts grâce à une analyse détaillée des frais associés à l'approvisionnement.

Renforcer la prise de décision en fournissant des insights exploitables sur les processus d'achat.

Relations et Intégration des Données

La force de la base de données AdventureWorks réside dans la structure relationnelle qui permet de connecter différents modules pour une analyse approfondie. Le module Purchasing joue un rôle central dans la gestion des fournisseurs et des commandes, mais son utilité est renforcée lorsqu'il est intégré avec d'autres tables comme Product du module Production. Ces relations permettent d'optimiser la gestion des achats, d'évaluer les performances des fournisseurs et de comprendre la dynamique entre produits et approvisionnement.

Relations Clés dans le Module Purchasing

PurchaseOrderHeader :

Cette table agit comme la base des informations globales des commandes passées auprès des fournisseurs.

Relation avec PurchaseOrderDetail : Une relation clé à clé primaire (PurchaseOrderID) connecte ces deux tables. Chaque commande dans PurchaseOrderHeader peut avoir plusieurs lignes détaillées dans PurchaseOrderDetail, contenant des informations sur les articles commandés, leurs quantités, et leurs coûts unitaires.

Relation avec ShipMethod : La colonne ShipMethodID dans PurchaseOrderHeader est liée à la table ShipMethod, qui fournit des détails sur les modes de transport, leurs coûts et leur efficacité.

PurchaseOrderDetail :

Cette table offre des informations détaillées pour chaque article inclus dans une commande.

Relation avec Product : Une connexion via ProductID relie les articles commandés dans PurchaseOrderDetail à leurs descriptions détaillées dans la table Product du module Production. Cela permet d'identifier les caractéristiques du produit, comme sa catégorie, son coût standard et sa disponibilité.

Relation avec Vendor : Grâce à PurchaseOrderID, cette table est indirectement reliée à Vendor via PurchaseOrderHeader, permettant d'associer chaque article à un fournisseur précis.

Vendor :

La table Vendor contient les informations des fournisseurs, notamment leur fiabilité (CreditRating), leur activité (ActiveFlag), et leurs coordonnées.

Relation avec VendorAddress et VendorContact : Ces relations permettent de lier les fournisseurs à leurs adresses physiques et aux personnes à contacter pour les négociations et le suivi.

Relation avec ProductVendor : Une relation clé à clé primaire (VendorID) connecte Vendor à ProductVendor. Cela permet d'associer les produits aux fournisseurs spécifiques qui les proposent, incluant les coûts associés, les quantités minimales et les délais moyens de livraison.

ShipMethod :

Cette table répertorie les différentes méthodes d'expédition disponibles et leurs coûts.

Relation avec PurchaseOrderHeader : Grâce à ShipMethodID, chaque commande peut être associée à un mode de transport, facilitant l'analyse des coûts et délais d'expédition.

VendorContact

Description : Cette table contient les informations sur les contacts des fournisseurs, telles que les noms, les numéros de téléphone et les e-mails.

Relations :

Avec Vendor : Reliée via VendorID, elle associe les contacts aux fournisseurs.

Avec ContactType : La clé ContactTypeID permet de spécifier le type de contact (par exemple : service commercial, logistique, etc.).

VendorAddress

Description : Cette table contient les informations d'adresse des fournisseurs, comme la ville (City), la région (StateProvinceID), et le pays (CountryRegionCode).

Relations :

Avec Vendor : Liée par VendorID, elle permet d'associer plusieurs adresses à un fournisseur.

Avec AddressType : Utilise AddressTypeID pour catégoriser les adresses (par exemple : adresse principale, adresse d'entrepôt, etc.).

Intégration avec le Module Production (Table Product)

La table Product du module Production est une pièce maîtresse de l'intégration avec Purchasing. Elle permet de relier les articles commandés aux informations détaillées sur les produits, renforçant ainsi les analyses et la prise de décision.

Relation avec PurchaseOrderDetail :

La table PurchaseOrderDetail utilise la clé ProductID pour se connecter à Product. Cela permet d'intégrer des données telles que :

Le coût standard (StandardCost) pour évaluer les marges bénéficiaires.

La catégorie et sous-catégorie des produits (ProductSubcategoryID) pour une analyse segmentée.

Les dates de mise en vente ou de retrait (SellStartDate, DiscontinuedDate) pour comprendre la disponibilité des produits.

Relation avec ProductVendor :

La table ProductVendor joue un rôle de pont entre les modules Purchasing et Production. Elle relie les produits aux fournisseurs spécifiques, en fournissant des informations supplémentaires comme :

Le coût d'approvisionnement (LastReceiptCost).

Les délais moyens de livraison (AverageLeadTime).

Les quantités minimales requises pour les commandes (MinOrderQty).

Analyse des Relations pour une Intégration Optimale

Ces relations permettent d'explorer des dynamiques cruciales pour la gestion des achats et de la production :

Performance des fournisseurs : En combinant les données de Vendor, PurchaseOrderHeader et PurchaseOrderDetail, il est possible d'évaluer les délais de livraison, les coûts totaux, et les taux de rejet par fournisseur.

Analyse des coûts des produits : La connexion entre PurchaseOrderDetail et Product permet de comprendre les coûts d'approvisionnement en fonction des produits et des fournisseurs.

Optimisation des expéditions : La relation entre ShipMethod et PurchaseOrderHeader permet d'évaluer les modes d'expédition les plus efficaces en termes de coût et de délai.

Synthèse

Le module Purchasing, grâce à ses relations avec des tables clés comme Product, ShipMethod et Vendor, offre une vue holistique des processus d'approvisionnement. L'intégration avec le module Production, en particulier la table Product, enrichit l'analyse en ajoutant une perspective produit, essentielle pour évaluer la rentabilité et l'efficacité des opérations. Cette interconnexion entre les modules permet une prise de décision éclairée, orientée vers l'optimisation des coûts et des performances.

Conclusion

En conclusion, ce chapitre a permis de comprendre l'importance stratégique du module

Purchasing dans l'optimisation des processus d'approvisionnement et de gestion des fournisseurs, tout en mettant en lumière ses interactions avec les autres modules, notamment Production. Ces relations offrent une base solide pour analyser les coûts, les performances des fournisseurs et l'efficacité des méthodes d'expédition.

Dans le chapitre suivant, nous explorerons les outils et techniques utilisés pour nettoyer et analyser les données, ouvrant ainsi la voie à une exploitation optimale des informations disponibles.

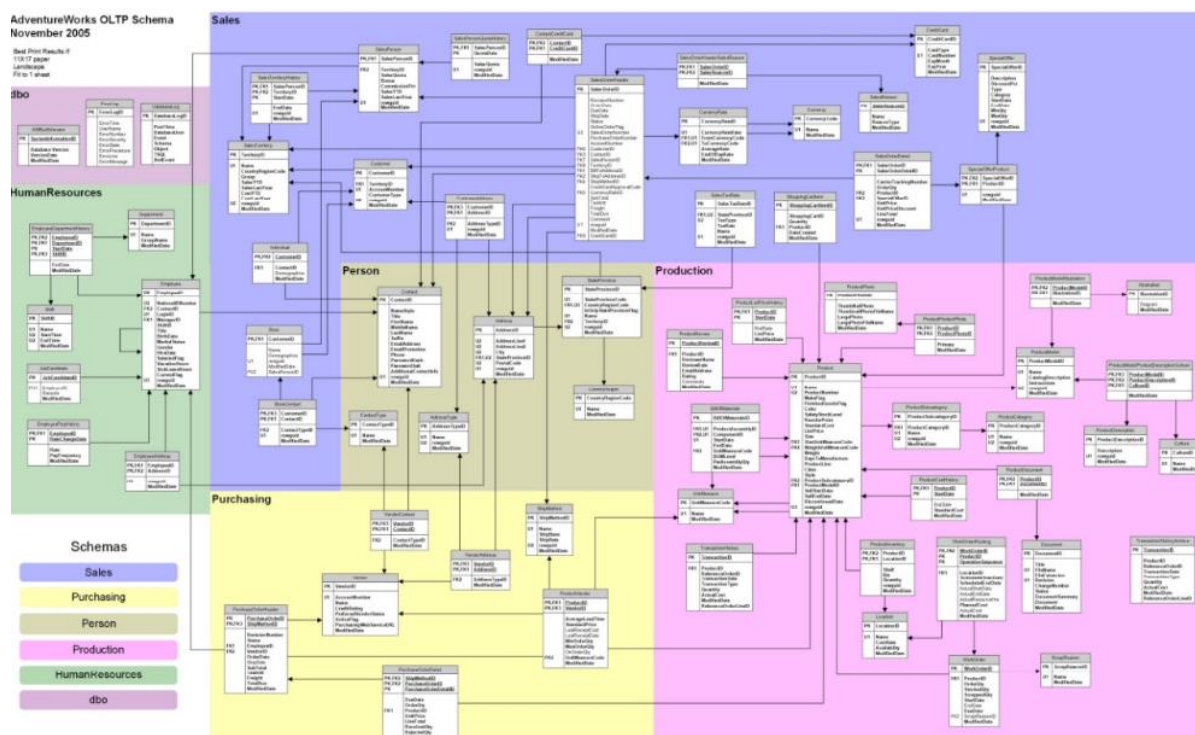


Figure 1:AdventureWorks schéma



CHAPITRE 2.

OUTILS ET MÉTHODOLOGIE



Description du chapitre :

le chapitre "Outils et Méthodologie" détaille les outils utilisés pour connecter, nettoyer et transformer les données issues de la base AdventureWorks. Il décrit le processus de connexion à SQL Server, les techniques de nettoyage et de préparation des données, ainsi que la création de relations et de mesures analytiques dans Power BI.



Introduction

Dans tout projet d'analyse de données, la méthodologie adoptée et les outils choisis jouent un rôle déterminant dans la qualité et la pertinence des résultats obtenus. Ce chapitre détaille les étapes clés qui ont permis de transformer les données brutes de la base AdventureWorks en insights exploitables, en s'appuyant sur une combinaison de technologies robustes et de pratiques méthodologiques rigoureuses.

L'accent est mis sur le processus de connexion aux données, la préparation et le nettoyage de ces dernières. Grâce à SQL Server une plateforme puissante pour la gestion des bases relationnelles, knime Analytics Platform est un outil open-source puissant pour le nettoyage, la transformation et l'analyse de données plateforme, et Power BI, un outil de visualisation et d'analyse de données de référence, nous avons pu extraire, structurer et analyser les informations critiques issues des modules clés de la base AdventureWorks. Ces outils, associés à une démarche structurée, garantissent une approche cohérente et reproductible pour l'analyse des données.

Environnement de travail :

SQL Server Management Studio (SSMS)

SQL Server Management Studio est un environnement intégré fourni par Microsoft pour gérer et interagir avec des bases de données relationnelles. SSMS est particulièrement adapté pour manipuler et explorer des bases comme AdventureWorks.



Power BI

Power BI est une solution de visualisation et d'analyse de données de Microsoft. Il permet de connecter des données provenant de SQL Server et d'autres sources pour les transformer en visualisations interactives.



KNIME

KNIME Analytics Platform est un outil open-source puissant pour le nettoyage, la transformation et l'analyse de données. Dans ce projet, KNIME a été utilisé en complément de SQL Server Management Studio (SSMS) et Power BI pour automatiser le processus de traitement des données et optimiser les étapes de préparation des données issues d'AdventureWorks.



Méthodologie

Installation et Configuration de SQL Server :

Télécharger et installer SQL Server et SSMS depuis le site officiel de Microsoft.

Configurer une instance de serveur local ou se connecter à une instance existante.

Connexion au Serveur :

Lancer SSMS.

Fournir les informations de connexion : nom du serveur, méthode d'authentification, nom d'utilisateur et mot de passe.

Chargement de la Base AdventureWorks :

Importer la base AdventureWorks si elle n'est pas déjà disponible. Ce processus implique l'attachement d'un fichier .bak (fichier de sauvegarde).

Une fois importée, la base apparaît dans le volet Object Explorer.

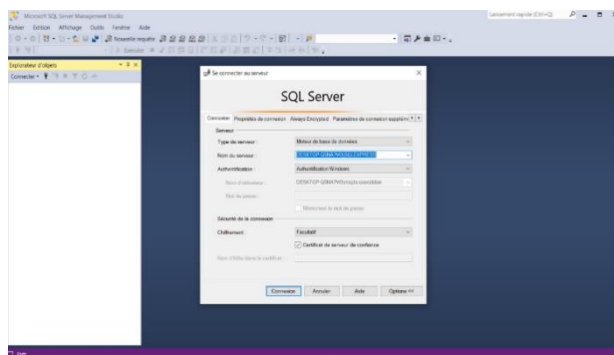


Figure 3:Fenêtre de Connexion à SQL Server

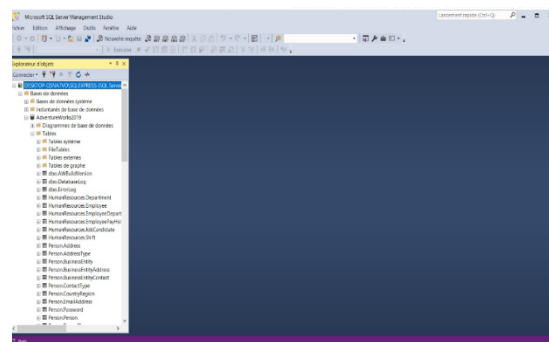


Figure 2:AdventureWorks Dans Explorateur d'Objets dans SSMS

Récupération des Données :

Sélection de la base de données dans SQL Server Management Studio (SSMS)

Dans l'explorateur d'objets de SSMS, sélectionnez la base de données que vous souhaitez exporter (ici, AdventureWorks2019).

Faites un clic droit sur la base de données, puis accédez à l'option Tâches et choisissez Exporter les données

Lancement de l'assistant d'importation et d'exportation SQL Server

Une fenêtre s'ouvre pour démarrer l'assistant.

Cet assistant guide l'utilisateur à travers plusieurs étapes pour configurer l'exportation ou l'importation des données de SQL Server.

Spécifier l'action à réaliser

Dans l'assistant, deux choix sont proposés :

Copier les données à partir d'une ou plusieurs tables ou vues : Cela permet d'exporter toutes les données d'une table ou d'une vue directement.

Écrire une requête pour spécifier les données à transférer : Cette option permet de sélectionner uniquement certaines données à exporter via une requête SQL personnalisée.

Ici, l'option "Copier les données à partir d'une ou plusieurs tables ou vues" est sélectionnée.

Définir la destination

Sélectionnez Flat File Destination comme type de destination.

Définissez le chemin du fichier où les données seront exportées, comme un fichier CSV.

Configurez les paramètres régionaux, le format (délimité, ici avec des virgules ou points-virgules), et assurez-vous que l'option "Noms de colonnes dans la première ligne de données" est cochée. Cela garantit que la première ligne du fichier contient les en-têtes des colonnes.

Préparer le fichier CSV

Une fois les configurations validées, cliquez sur "Suivant" pour lancer le processus.

Le fichier Csv est généré avec les données exportées

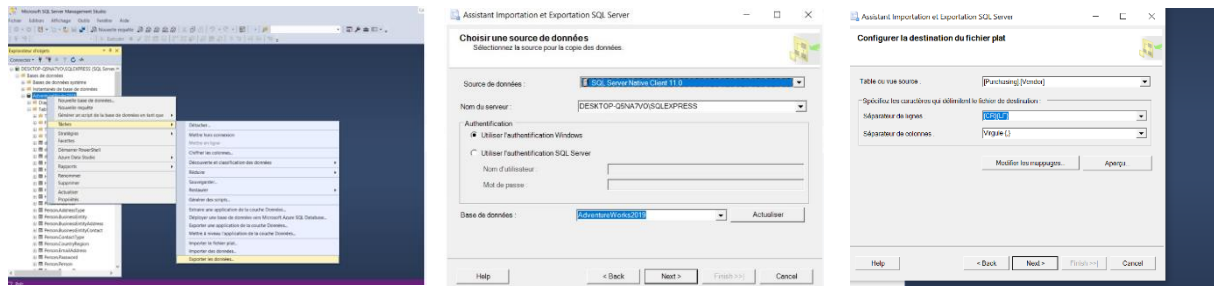


Figure 4: Exportation des Données depuis SQL Server vers un Fichier CSV

nettoyage des données avec KNIME

Lecture des données (CSV Reader)

Fonction : Ce nœud est utilisé pour importer les fichiers CSV contenant les données brutes à traiter.

Détails :

Il prend en charge les formats de fichiers délimités (par exemple, par des virgules ou des tabulations).

Permet de configurer la lecture des en-têtes et d'autres propriétés telles que l'encodage et les types de colonnes.

Importance : C'est la première étape pour rendre les données accessibles au reste du pipeline.

Filtrage des doublons (Duplicate Row Filter)

Fonction : Identifie et supprime les lignes dupliquées dans les données pour éviter les redondances.

Détails :

Configure des colonnes clés pour déterminer quelles lignes considérer comme doublons.

Offre une option pour conserver uniquement la première ou la dernière occurrence des doublons.

Importance : Garantit l'unicité des enregistrements et améliore la qualité des données.

Manipulation des chaînes (String Manipulation)

Fonction : Modifie ou nettoie les données textuelles dans les colonnes.

Détails :

Permet de formater les textes (par exemple, passer des chaînes en majuscules ou supprimer des caractères spéciaux).

Facilite l'extraction de parties spécifiques des chaînes ou la concaténation de colonnes.

Importance : Standardise les formats textuels pour une meilleure cohérence des données.

Gestion des valeurs manquantes (Missing Value)

Fonction : Remplit ou gère les valeurs manquantes dans les colonnes de données.

Détails :

Offre des options pour remplacer les valeurs manquantes par des valeurs par défaut (moyenne, médiane, zéro, etc.) ou pour les ignorer.

Peut aussi interpoler des données manquantes dans certains cas.

Importance : Améliore la complétude des données et évite les erreurs lors des analyses.

Transformation des formats de date (String to Date/Time)

Fonction : Convertit des chaînes textuelles représentant des dates en un format de date/heure utilisable.

Détails :

Requiert la spécification du format de la date d'entrée (par exemple, "yyyy-MM-dd").

Valide et standardise les dates pour permettre des calculs ou filtrages basés sur des plages de dates.

Importance : Normalise les dates pour une manipulation efficace dans les analyses.

Filtrage des colonnes (Column Filter)

Fonction : Permet de sélectionner les colonnes nécessaires à l'analyse ou d'exclure les colonnes inutiles.

Détails :

Configure les colonnes à conserver ou à ignorer dans le dataset.

Importance : Réduit la taille des données pour améliorer les performances et la lisibilité.

Détection des valeurs aberrantes (Numeric Outliers)

Fonction : Identifie les valeurs qui sortent des plages acceptables dans les colonnes numériques.

Détails :

Permet de spécifier des seuils ou des règles statistiques (écart-type, quartiles) pour détecter les anomalies.

Importance : Nettoie les données des valeurs extrêmes qui pourraient biaiser les analyses.

Écriture des données nettoyées (CSV Writer)

Fonction : Exporte les données traitées et nettoyées dans un fichier CSV pour les étapes ultérieures.

Détails :

Permet de configurer les options de sortie, telles que l'encodage, les séparateurs, et l'inclusion des en-têtes.

Importance : Rend les données prêtes à être importées dans d'autres outils d'analyse comme Power BI.

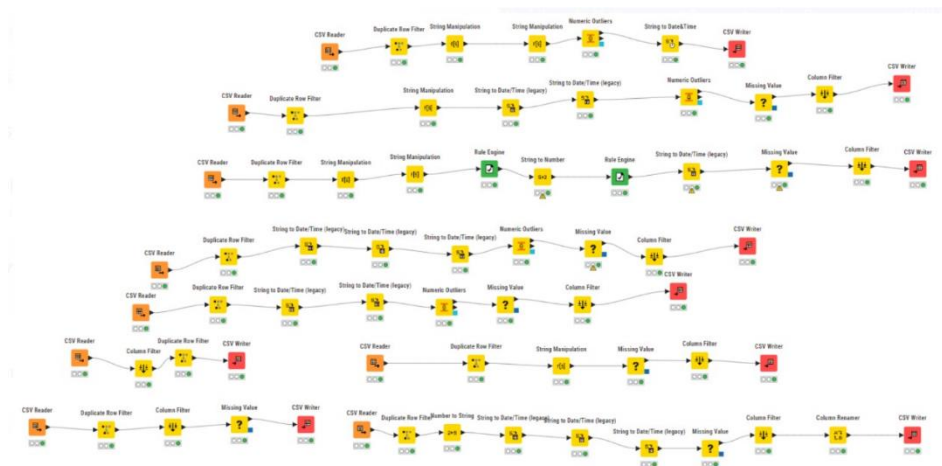


Figure 5:Workflow de Nettoyage et Préparation des Données avec KNIME

Nettoyage et Transformation des Données avec Power Query dans Power BI

Dans ce projet, Power Query a été utilisé comme une étape finale de nettoyage et de transformation des données, après leur traitement initial dans KNIME. Bien que KNIME ait permis d'effectuer un nettoyage approfondi, comme l'élimination des doublons, le traitement

des valeurs manquantes et la transformation des colonnes, Power Query a joué un rôle complémentaire essentiel pour affiner et préparer les données à des fins d'analyse dans Power BI.

Cette approche combinée a permis de tirer parti des points forts de chaque outil. KNIME a géré les traitements complexes et massifs sur des ensembles de données volumineux, tandis que Power Query a assuré une dernière vérification et des ajustements spécifiques. Parmi les actions réalisées dans Power Query, on retrouve :

Validation des Transformations :

Les données provenant de KNIME ont été vérifiées pour s'assurer qu'elles respectaient les exigences du modèle Power BI. Les types de données ont été revérifiés et ajustés si nécessaire.

Adaptation aux Visualisations :

Des colonnes supplémentaires ont été créées ou modifiées directement dans Power Query pour répondre à des besoins spécifiques des visualisations, comme la création de groupes, de catégories ou de colonnes calculées.

Enrichissement des Données :

Power Query a été utilisé pour enrichir les données avec des calculs spécifiques ou des formats requis, par exemple l'ajout de colonnes basées sur des expressions conditionnelles.

Harmonisation des Formats :

L'étape finale de Power Query a permis d'uniformiser les formats des données, comme les dates et les nombres, pour garantir leur cohérence dans les graphiques et les KPI de Power BI.

Élimination des Dernières Anomalies :

Si des incohérences résiduelles étaient identifiées, elles ont été corrigées rapidement dans Power Query, grâce à sa simplicité et son interface intuitive.

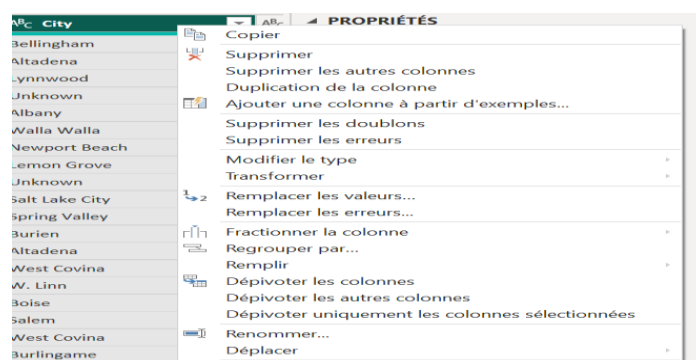


Figure 6: Power Query pour le Nettoyage

Relation entre les tables

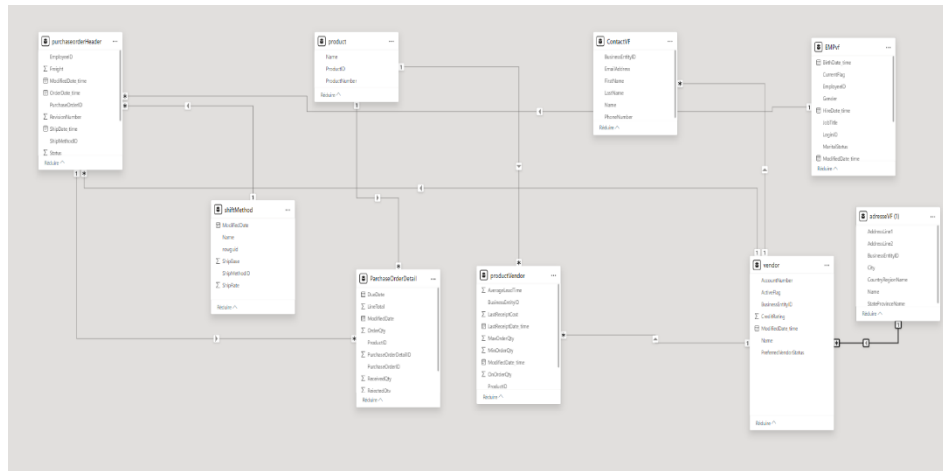


Figure 7:vue modèle

Relation	Table Source	Table Cible	Cardinalité	Direction de Filtrage	Explication
Fournisseur - Commande d'Achat	Vendor	PurchaseOrderHeader	1:*	De Vendor vers PurchaseOrderHeader	Un fournisseur peut avoir plusieurs commandes associées.
Commande d'Achat - Détail de Commande	PurchaseOrderHeader	PurchaseOrderDetail	1:*	De PurchaseOrderHeader vers PurchaseOrderDetail	Une commande peut contenir plusieurs articles détaillés.
Détail de Commande - Produit	PurchaseOrderDetail	Product	*:1	De Product vers PurchaseOrderDetail	Chaque ligne de commande correspond à un seul produit.
Méthode d'Expédition - Commande d'Achat	ShipMethod	PurchaseOrderHeader	1:*	De ShipMethod vers PurchaseOrderHeader	Une méthode d'expédition peut être utilisée pour plusieurs commandes.
Fournisseur - Produits Fournis	Vendor	ProductVendor	1:*	De Vendor vers ProductVendor	Un fournisseur peut fournir plusieurs produits.
Produit - Fournisseurs	Product	ProductVendor	1:*	De Product vers ProductVendor	Un produit peut être fourni par plusieurs fournisseurs.
Adresse Fournisseur	AddressVF	Vendor	1:1	De AddressVF vers Vendor	Chaque fournisseur est associé à une seule adresse.



Conclusion

Ce chapitre a présenté une méthodologie claire et structurée pour exploiter efficacement les données de la base AdventureWorks. En combinant des outils robustes comme SQL Server, KNIME Analytics Platform et Power BI avec des pratiques méthodologiques rigoureuses, nous avons assuré un processus de préparation, de nettoyage et de structuration des données de haute qualité. Cette approche a permis de transformer des données brutes en une base solide pour des analyses approfondies.

Dans le chapitre suivant, nous nous concentrerons sur les résultats obtenus et les visualisations réalisées. Nous analyserons les insights générés pour mieux comprendre les tendances et les performances des données, tout en mettant en lumière les opportunités d'amélioration et d'optimisation.

CHAPITRE 3.

RÉSULTATS ET VISUALISATIONS



Description du chapitre :

Ce chapitre se concentre sur l'interprétation des données analysées à travers des visualisations claires et informatives. Les résultats obtenus mettent en évidence les performances des fournisseurs, les coûts d'expédition, et les produits les plus rentables. Grâce aux tableaux de bord interactifs, des recommandations stratégiques peuvent être formulées pour optimiser les processus d'approvisionnement et de gestion.





Introduction

Ce chapitre s'inscrit dans la continuité des travaux précédents en mettant l'accent sur l'analyse et l'interprétation des données issues de la base AdventureWorks. Après avoir préparé, nettoyé et structuré les données, l'étape suivante consiste à extraire des insights significatifs et à les traduire en visualisations claires et informatives. Cette démarche permet de rendre les résultats plus accessibles et exploitables pour une prise de décision éclairée.

L'objectif principal de ce chapitre est de fournir une analyse approfondie des performances des fournisseurs, des coûts liés à l'expédition, et des produits les plus rentables. Ces éléments jouent un rôle central dans la gestion des processus d'approvisionnement et de logistique. Grâce à l'utilisation de tableaux de bord interactifs créés avec Power BI, nous présentons les données sous une forme intuitive, mettant en évidence des indicateurs clés et des tendances pertinentes.

Les résultats obtenus ne se limitent pas à une simple présentation des données. Ils permettent également de poser les bases d'une réflexion stratégique sur l'optimisation des processus opérationnels. Par exemple, l'analyse des performances des fournisseurs offre des pistes d'amélioration pour renforcer la chaîne d'approvisionnement, tandis que l'examen des coûts d'expédition identifie des opportunités pour réduire les dépenses logistiques. Par ailleurs, l'évaluation des produits les plus rentables guide les choix stratégiques en matière d'offre et de gestion des stocks.

En définitive, ce chapitre vise à transformer les données brutes en connaissances exploitables, en mettant en lumière les principaux leviers d'optimisation. Il constitue une étape essentielle dans l'exploitation des données pour répondre aux objectifs globaux du projet et propose des recommandations stratégiques basées sur des faits concrets et mesurables.

Visualisation

Page1

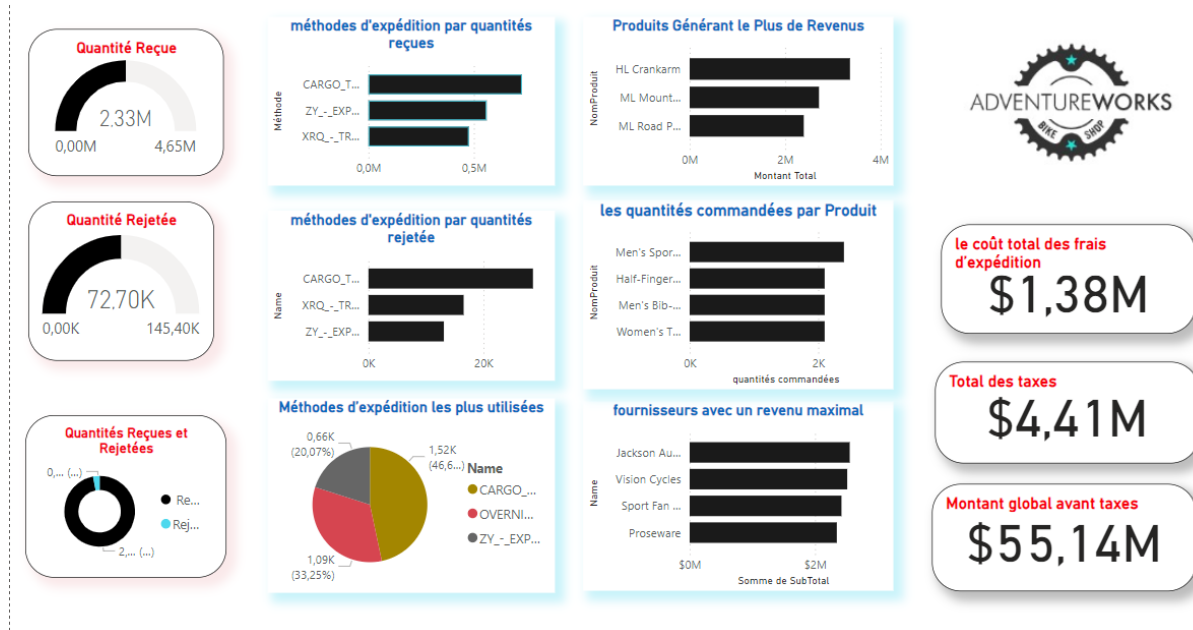


Figure 8: Vision Stratégique Globale et Analyse des Performances

Quantité Reçue (Gauge) :

Ce graphique indique le total des quantités reçues auprès des fournisseurs. Cela reflète le volume total des articles acceptés et reçus dans l'inventaire. Cette mesure est essentielle pour évaluer la fiabilité des fournisseurs et la gestion des stocks.

Quantité Rejetée (Gauge) :

Ce graphique montre le total des quantités rejetées, c'est-à-dire les articles qui n'ont pas satisfait les normes de qualité ou les spécifications attendues. Il permet de suivre les performances des fournisseurs en termes de qualité.

Quantités Reçues et Rejetées (Donut Chart) :

Ce graphique circulaire compare les proportions entre les quantités reçues et rejetées. Cela donne une vue d'ensemble sur la qualité des produits livrés et la proportion de produits acceptés par rapport à ceux rejetés.

Méthodes d'expédition par Quantités Reçues (Bar Chart) :

Ce graphique en barres indique les quantités reçues selon les différentes méthodes d'expédition utilisées. Il aide à identifier les méthodes de transport les plus fiables pour les livraisons réussies.

Méthodes d'expédition par Quantités Rejetées (Bar Chart) :

Ce graphique montre les quantités rejetées pour chaque méthode d'expédition. Il permet de détecter si certaines méthodes sont moins efficaces ou liées à des produits de qualité inférieure.

Méthodes d'expédition les plus utilisées (Pie Chart) :

Ce graphique circulaire met en évidence la répartition des méthodes d'expédition les plus fréquemment employées, en pourcentage. Cela aide à comprendre les préférences ou habitudes en matière de transport.

Produits générant le plus de revenus (Bar Chart) :

Ce graphique met en avant les produits ayant rapporté le plus de revenus, en fonction du montant total généré. Il permet d'identifier les articles les plus performants en termes de rentabilité.

Les Quantités Commandées par Produit (Bar Chart) :

Ce graphique montre les quantités commandées pour chaque produit, ce qui permet d'analyser les tendances de demande et les produits les plus populaires.

Fournisseurs avec un Revenu Maximal (Bar Chart) :

Ce graphique liste les fournisseurs ayant généré le plus de revenus pour l'entreprise, en fonction des montants totaux des commandes effectuées.

Coût Total des Frais d'Expédition (KPI Card) :

Ce KPI met en avant le coût cumulé des frais d'expédition pour toutes les commandes, un indicateur clé pour analyser les dépenses logistiques.

Total des Taxes (KPI Card) :

Ce KPI indique le montant total des taxes appliquées sur les commandes. C'est une mesure importante pour suivre les charges fiscales.

Montant Global Avant Taxes (KPI Card) :

Ce KPI résume le montant total des revenus générés avant l'application des taxes. Il reflète la performance brute des ventes et des achats.

Page2

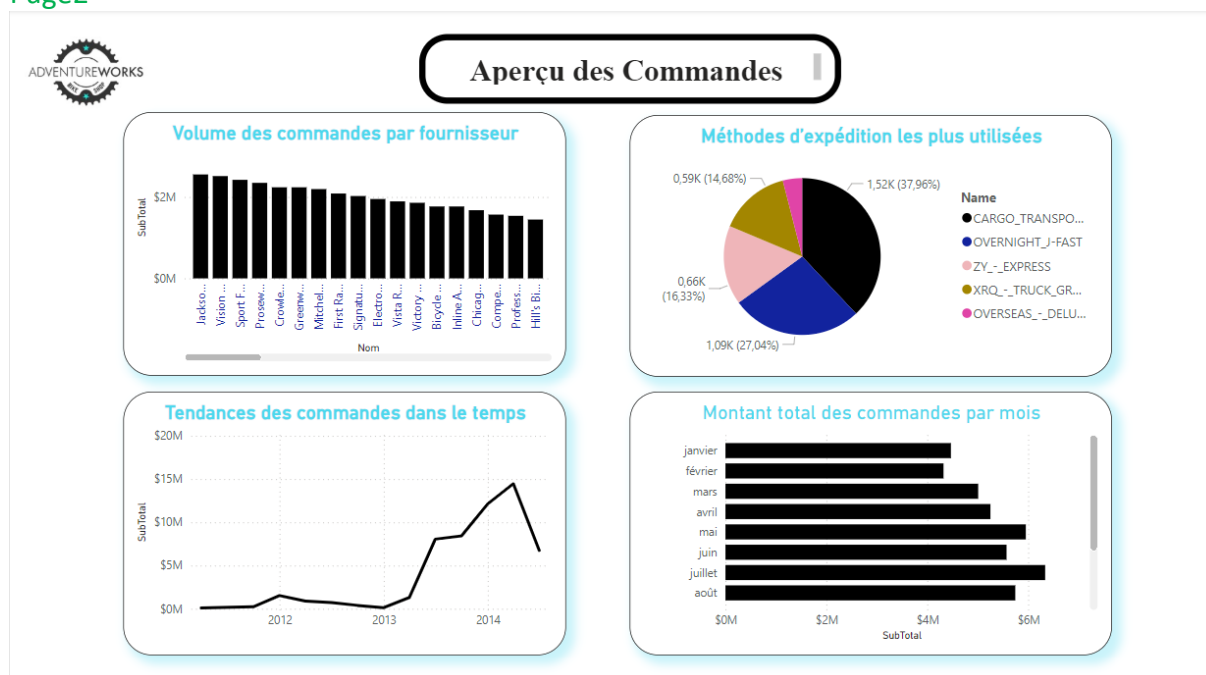


Figure 9:Aperçu des commandes

Explications des graphes de la page "Aperçu des Commandes"

Volume des commandes par fournisseur

Ce graphe montre le volume total des commandes passées à chaque fournisseur, exprimé en sous-totaux (SubTotal). Il permet d'identifier les principaux fournisseurs avec lesquels l'entreprise collabore le plus, en fonction du montant des commandes. Cela peut aider à évaluer les partenariats stratégiques et à prioriser les négociations.

Méthodes d'expédition les plus utilisées

Ce diagramme circulaire présente la répartition des différentes méthodes d'expédition utilisées pour les commandes. Chaque segment représente un mode d'expédition, accompagné de son pourcentage d'utilisation. Ce graphe permet de visualiser les méthodes les plus populaires et leur impact logistique.

Tendances des commandes dans le temps

Ce graphe linéaire illustre l'évolution du volume des commandes au fil des années. Il offre un aperçu des tendances, en mettant en évidence les périodes de croissance ou de baisse dans les commandes. Cette visualisation aide à repérer les périodes saisonnières ou les fluctuations dans les activités commerciales.

Montant total des commandes par mois

Ce graphe en barres montre le montant total des commandes réparti par mois. Il permet d'identifier les mois où les commandes atteignent leur pic ou leur creux, donnant un aperçu des cycles d'activités saisonniers ou des variations dans la demande.



Détails des produits

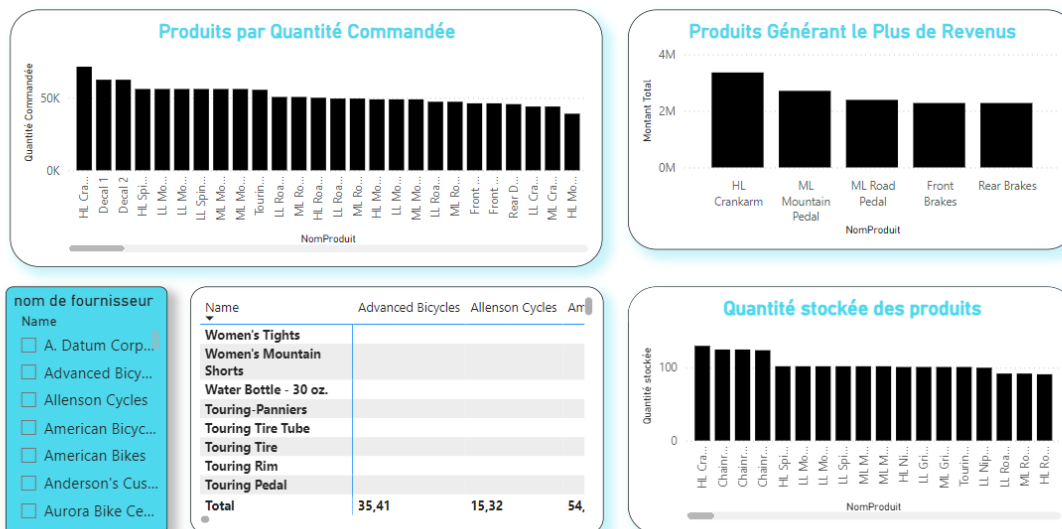


Figure 10:Détails des produits

Explications des graphes de la page "Détails des produits"

Produits par Quantité Commandée

Ce graphique en barres présente les produits les plus commandés en termes de quantité. Il permet d'identifier les articles les plus demandés et peut guider les décisions sur le réapprovisionnement et la production.

Produits Générant le Plus de Revenus

Ce graphique en barres met en évidence les produits qui génèrent les revenus les plus élevés. Il aide à repérer les articles stratégiques pour la rentabilité de l'entreprise et à orienter les efforts marketing ou de production.

Nom de Fournisseur et Liste des Produits

Cette table interactive associe les fournisseurs aux produits qu'ils fournissent, avec les quantités correspondantes. Elle offre une vue détaillée des relations entre fournisseurs et produits, facilitant l'analyse des performances de chaque fournisseur.

Quantité Stockée des Produits

Ce graphique en barres montre la quantité actuelle en stock pour chaque produit. Il aide à évaluer les niveaux de stock, à repérer les surplus ou les produits nécessitant un réapprovisionnement imminent.

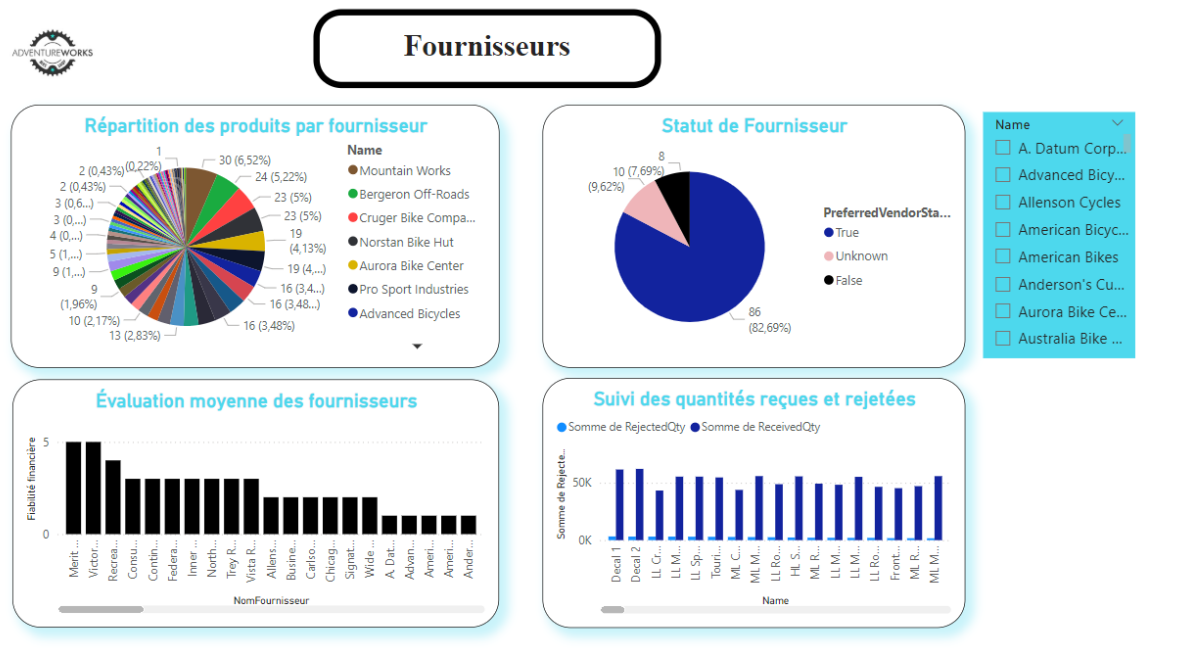


Figure 11:fournisseurs

Explications des graphes de la page "Fournisseurs"

Répartition des Produits par Fournisseur

Ce graphique circulaire illustre la proportion de produits fournis par chaque fournisseur. Il permet de visualiser la contribution relative de chaque fournisseur à l'approvisionnement global, facilitant l'identification des principaux partenaires commerciaux.

Statut de Fournisseur

Ce graphique circulaire indique le statut des fournisseurs :

PreferredVendorStatus (True) : Fournisseurs préférés ou fiables.

False : Fournisseurs non préférés.

Unknown : Fournisseurs dont le statut n'est pas défini.

Cette répartition aide à distinguer les partenaires de confiance des autres.

Évaluation Moyenne des Fournisseurs

Ce graphique en barres présente la fiabilité financière ou d'autres critères d'évaluation moyenne des fournisseurs. Il permet d'identifier les fournisseurs les plus fiables et de hiérarchiser les partenariats stratégiques.

Suivi des Quantités Reçues et Rejetées

Ce graphique en barres compare les quantités de produits reçues et rejetées pour chaque fournisseur. Il aide à analyser la qualité des produits livrés et à évaluer la performance des fournisseurs en termes de conformité aux exigences.

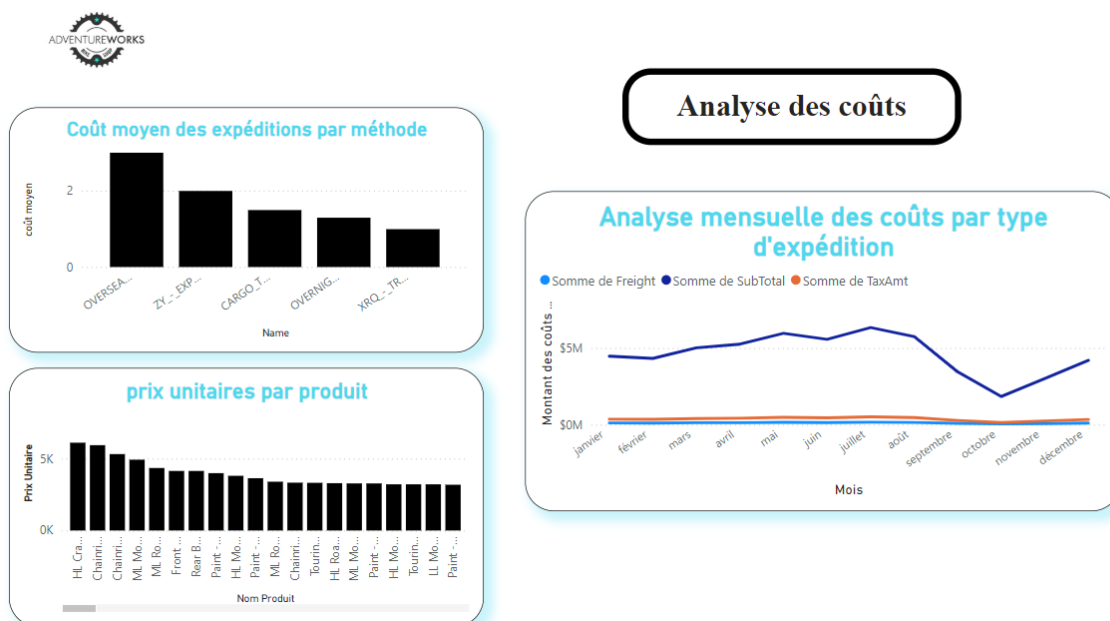


Figure 12:Analyse des coûts

Explications des graphes de la page "Analyse des Coûts"

Coût Moyen des Expéditions par Méthode

Ce graphique en barres montre le coût moyen des expéditions pour chaque méthode de transport (par exemple : OVERSEA, ZY_EXPRESS, etc.). Il permet d'identifier les méthodes les plus économiques et celles qui engendrent les coûts les plus élevés, offrant ainsi une base pour optimiser les dépenses liées à l'expédition.

Analyse Mensuelle des Coûts par Type d'Expédition

Ce graphique linéaire suit l'évolution mensuelle de trois types de coûts :

Somme de Freight (Frais d'expédition)

Somme de SubTotal (Sous-total des commandes)

Somme de TaxAmt (Taxes)

Ce suivi permet de comprendre les tendances saisonnières ou mensuelles des dépenses et d'identifier les périodes où les coûts sont les plus élevés.

Prix Unitaires par Produit

Ce graphique en barres répertorie les produits en fonction de leur prix unitaire. Il met en lumière les produits les plus coûteux et ceux qui sont plus abordables, facilitant ainsi des décisions stratégiques sur la gestion des prix et des marges bénéficiaires.



Conclusion

Résultats et Visualisations

Ce chapitre a permis d'explorer les principaux résultats obtenus grâce à l'analyse des données issues de la base AdventureWorks, en mettant en lumière des tendances stratégiques et des points d'amélioration dans les domaines des commandes, des produits, des fournisseurs et des coûts. Les visualisations présentées ont offert une vue d'ensemble claire et précise, facilitant l'interprétation des données et orientant la prise de décision.

Grâce à ces analyses, des opportunités significatives d'optimisation ont été identifiées, notamment dans la gestion des coûts d'expédition, l'évaluation des fournisseurs, et la rentabilité des produits. Ces insights serviront de base pour des recommandations opérationnelles et stratégiques à l'échelle de l'entreprise.

Conclusion générale

Ce rapport marque l'aboutissement d'un projet complexe mais enrichissant, qui a permis de démontrer l'importance de l'analyse de données dans la gestion stratégique des processus d'une entreprise. En partant de la base de données AdventureWorks, nous avons exploré différentes étapes clés allant de la connexion et du nettoyage des données à leur visualisation et à l'extraction d'insights exploitables. L'objectif principal était de simuler les opérations d'une entreprise fictive spécialisée dans les produits liés au cyclisme, tout en identifiant des opportunités d'optimisation dans les processus d'achats, de production et de gestion des fournisseurs.

À travers ce travail, nous avons mis en œuvre une méthodologie rigoureuse en combinant plusieurs outils technologiques, notamment SQL Server pour l'extraction des données, KNIME pour leur transformation et nettoyage, et Power BI pour la visualisation et l'interprétation des résultats. Ces outils nous ont permis de transformer des données brutes en informations stratégiques, apportant ainsi des réponses concrètes aux problématiques posées. Par exemple, nous avons analysé la performance des fournisseurs, identifié les produits les plus rentables, et évalué les coûts d'expédition. Ces analyses ne se sont pas limitées à une simple présentation des chiffres, mais ont permis de formuler des recommandations pratiques pour optimiser les performances opérationnelles et réduire les coûts.

Ce projet illustre également l'interconnexion et la complémentarité des différentes étapes d'un cycle analytique. Chaque phase a apporté sa valeur ajoutée, depuis la préparation des données jusqu'à leur interprétation finale, démontrant que l'analyse de données est bien plus qu'une démarche technique ; elle est au cœur des décisions stratégiques modernes. Enfin, ce travail met en avant l'importance de la collaboration entre les outils et les technologies pour générer des résultats cohérents et significatifs.

En conclusion, ce rapport constitue une contribution modeste mais significative à l'application des techniques analytiques dans des scénarios professionnels. Il offre une base solide pour les organisations cherchant à maximiser l'efficacité de leurs processus grâce à l'analyse de données. Nous espérons que ce projet servira d'exemple et d'inspiration pour des initiatives similaires dans le futur, où la rigueur méthodologique et l'innovation technologique joueront un rôle clé dans la réussite.