Projet Data Science :

Analyse des Ventes et Indicateurs de Performance en Python

Application : Analyse des ventes et visualisation dynamique

Réalisé par : Hajar Atyq

Encadré par : Pr. Mohammed QBADOU

# 1. Introduction

Dans le cadre du cours de Data Science avec Python, ce projet vise à explorer, analyser et visualiser des données commerciales provenant d’un fichier Excel. Les objectifs principaux incluent l’identification de tendances de ventes, le calcul de KPIs pertinents, et la comparaison des performances avec l’année précédente.

# 2. Objectifs du Projet

Ce projet vise à :  
- Nettoyer et préparer les données issues du fichier 'sales.xlsx'  
- Créer des indicateurs clés de performance (KPI)  
- Visualiser les données à l’aide de bibliothèques Python comme Matplotlib et Seaborn  
- Fournir une compréhension claire des tendances commerciales  
- Comparer les résultats actuels avec ceux de l’année précédente

# 3. Description des Données

Le fichier Excel contient 4 feuilles principales :  
- Sales Orders : données des commandes (quantité, prix, coût, dates...)  
- Customers : informations sur les clients (nom, taille, capital)  
- Regions : informations géographiques (ville, adresse, coordonnées)  
- Products : catalogue des produits

# 4. Méthodologie de traitement

## 4.1 Préparation des données

Les données ont été nettoyées : suppression des doublons, gestion des valeurs manquantes et conversions de types. Une jointure entre les feuilles a été faite pour créer une base complète exploitable.

## 4.2 Table de dates

Une table de dates a été générée afin de permettre une analyse temporelle précise (année, mois, jour, etc).

## 4.3 Création des KPIs

À partir des données, les indicateurs suivants ont été calculés :  
- Total Sales = somme des ventes  
- Total Profit = somme des bénéfices  
- Total Orders = nombre de commandes  
- Profit Margin % = profit / ventes  
- YOY % = variation annuelle des ventes et bénéfices

# 5. Analyse des Indicateurs

Les KPIs ont été utilisés pour évaluer la performance de l’entreprise. Les ventes totales et les bénéfices ont été calculés, puis comparés entre années. La marge bénéficiaire moyenne a été analysée pour chaque canal de vente et chaque produit.

# 6. Visualisation des Résultats

Des graphiques ont été générés avec Matplotlib et Seaborn :  
- Évolution mensuelle des ventes  
- Répartition des ventes par produit  
- Top 5 des clients par chiffre d’affaires  
- Répartition géographique des ventes  
- Comparaison YOY (année précédente)

# 7. Comparaison avec l’année précédente

L’analyse Year-over-Year (YOY) permet de voir la croissance ou la décroissance des ventes et bénéfices entre deux années. Cela a été fait pour chaque produit, client, canal et mois, afin de déterminer les tendances positives ou à surveiller.

# 8. Résultats clés et insights

Les résultats ont montré que certaines villes et certains clients génèrent une grande part des ventes. Le canal Wholesale est le plus performant en termes de profits. Quelques produits montrent une baisse des ventes d’une année à l’autre.

# 9. Conclusion

L’approche Data Science appliquée à l’analyse des ventes a permis d’extraire des informations utiles pour la stratégie commerciale. Les outils Python comme Pandas, Matplotlib, Seaborn ont facilité cette analyse. Des améliorations peuvent être apportées avec des outils de BI interactifs comme Power BI ou Tableau.