

Mise en situation:

Vous êtes un data analyst embauché par une entreprise de vente de meubles et fournitures de bureau. Votre mission consiste à analyser les données de ventes pour identifier les tendances, les opportunités d'amélioration et fournir des recommandations stratégiques pour augmenter la rentabilité.

Voici les étapes à suivre durant la semaine:

- **1- Lundi** : Collecte et Préparation des données (nettoyage, homogénéisation) :
    - \* Prise en main du jeu de données
    - \* Détection des anomalies, valeurs manquantes, doublons, valeurs aberrantes..
    - \* Uniformiser les formats des données (dates, chiffres, texte).
    - \* Faire une analyse univariée, multi-variée
  - **2- Mardi/Mercredi** : Stockage des données (conception et création de la BDD).
    - \* Modéliser les données propres avec un MCD
    - \* Construire un MLD cohérent
    - \* Préparer le script de création de la BDD conformément au schéma conçu
  - **3 - Lundi/mardi** : Mise à disposition des données et analyses.
    - \* Identifier des KPIs pertinents, Déterminer les KPIs en fonction des objectifs du projet
    - \* maquetter et mettre en place un dashboard d'analyse du projet
- => Vendredi : soutenance du projet

Pour mieux comprendre le dataset voici un dictionnaire des données:

### Métadonnées du Dataset

Nom de la colonne	Description	Type de données	Exemples de valeurs	Valeurs possibles/Plage de valeurs
order_id	Identifiant unique de la commande	Texte	CA-2016-152156	Alphanumérique unique
order_date	Date de la commande	Date	08/11/2016	Format JJ/MM/AAAA
ship_date	Date d'expédition de la commande	Date	11/11/2016	Format JJ/MM/AAAA
ship_mode	Mode d'expédition de la commande	Texte	Second Class	First Class, Second Class, Standard Class, Same Day
customer_id	Identifiant unique du client	Texte	CG-12520	Alphanumérique unique

Nom de la colonne	Description	Type de données	Exemples de valeurs	Valeurs possibles/Plage de valeurs
sales_rep	Représentant commercial associé à la vente	Texte	Organic	Noms des représentants
location_id	Identifiant unique de l'emplacement (ville et état)	Texte	42420,Henderson	Code postal, Ville
product_id	Identifiant unique du produit	Texte	FUR-BO-10001798	Alphanumérique unique
sales	Montant des ventes	Numérique	261.96	Valeurs monétaires positives
quantity	Quantité de produits vendus	Numérique	2	Nombres entiers positifs
discount	Remise appliquée sur la vente (en pourcentage)	Numérique	0.0	0.0 à 1.0 (0% à 100%)
profit	Profit réalisé sur la vente	Numérique	41.91	Valeurs monétaires (positives ou négatives)
customer_name	Nom du client	Texte	Claire Gute	Noms des clients
segment	Segment de marché du client	Texte	Consumer	Consumer, Corporate, Home Office
product_name	Nom du produit	Texte	Bush Somerset Collection Bookcase	Noms des produits
category	Catégorie du produit	Texte	Furniture	Furniture, Office Supplies, Technology
sub_category	Sous-catégorie du produit	Texte	Bookcases	Noms des sous-catégories (Bookcases, Chairs, Tables, etc.)
sales_team	Équipe de vente	Texte	Organic	Organic, Bravo
sales_team_manager	Responsable de l'équipe de vente	Texte	Organic	Noms des responsables
city	Ville où la commande a été expédiée	Texte	Henderson	Noms des villes
state	État où la commande a été expédiée	Texte	Kentucky	Noms des états

Nom de la colonne	Description	Type de données	Exemples de valeurs	Valeurs possibles/Plage de valeurs
postal_code	Code postal de la ville où la commande a été expédiée	Numérique	42420	Nombres entiers positifs
region	Région où la commande a été expédiée	Texte	South	South, West, Central, East