

Rapport projet

Nom : cheragui

Prenom : mayar

1. Introduction

Ce projet a pour objectif de concevoir une **solution Business Intelligence complète** à partir de la base de données **Northwind**.

Il couvre l'ensemble de la chaîne décisionnelle : **extraction des données, transformation, modélisation et visualisation**.

Le projet est réalisé **en monôme**, en utilisant **Python**, sans outils BI externes comme Power BI, conformément aux consignes pédagogiques.

.

2. Objectifs du projet

Les objectifs principaux sont :

- Extraire les données depuis la base Northwind
- Mettre en place un **ETL en Python**
- Construire une **table de faits** et des **dimensions**
- Créer des **indicateurs clés (KPI)**
- Produire un **dashboard analytique**
- Présenter les résultats sous forme de graphiques et de notebook

3. Source de données

La base utilisée est **Northwind**, une base relationnelle contenant notamment :

- Employees (employés)
- Customers (clients)
- Orders (commandes)
- Order Details
- Products
- Categories

Les données ont été extraites sous forme de fichiers CSV et stockées dans le dossier
data/raw

.

4. ETL en Python

L'ETL est divisé en plusieurs scripts afin de respecter une **bonne organisation du code**.

4.1 Extraction (Extract)

Les données sont chargées à partir des fichiers CSV de la base Northwind :

- Employees.csv
 - Customers.csv
 - Orders.csv
-

4.2 Transformation (Transform)

Les transformations principales sont :

- Conversion des dates (OrderDate, ShippedDate)
 - Création des indicateurs :
 - nb_commandes_livrees
 - nb_commandes_non_livrees
 - Construction des dimensions :
 - Dimension Employé
 - Dimension Client
 - Création de la table de faits `fact_orders`
-

4.3 chargement

Les données transformées sont sauvegardées dans `data/processed` :

- `dim_employee.csv`
- `dim_customer.csv`
- `fact_orders.csv`

5. Modélisation décisionnelle

Le modèle décisionnel adopté est un **schéma en étoile** :

- **Table de faits** : `fact_orders`
- **Dimensions** :
 - Employé
 - Client
 - Temps (via `OrderDate`)

Ce modèle facilite l'analyse des commandes selon différents axes.

6. Indicateurs clés (KPI)

Les indicateurs analysés sont :

- Nombre total de commandes
 - Nombre de commandes livrées
 - Nombre de commandes non livrées
 - Répartition par client
 - Répartition par employé
 - Évolution mensuelle des commandes
-

7. Visualisation et Dashboard

Les visualisations sont réalisées en Python avec :

- **Matplotlib** pour les graphiques classiques
- **Plotly** pour les graphiques interactifs 3D

Les graphiques produits sont :

- Répartition livrées / non livrées
 - Livrées vs non livrées par client
 - Livrées vs non livrées par employé
 - Évolution mensuelle des commandes
 - Graphe 3D (date – employé – client)
-

8. Notebook d'analyse

Un notebook `analysis.ipynb` est fourni afin de :

- Explorer les données
 - Justifier les choix analytiques
 - Présenter les résultats étape par étape
-

9. Conclusion

Ce projet démontre la mise en place complète d'une solution BI en Python, depuis l'extraction des données jusqu'à la visualisation finale.

Les résultats permettent d'analyser efficacement la performance des employés, des clients et le suivi des livraisons.