

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITÉ DES SCIENCES ET DE LA TECHNOLOGIE HOUARI
BOUMEDIENE



RAPPORT DE PROJET
Module : Business Intelligence

3^{ème} année Ingénieur Cyber Security

Conception et Implémentation d'une Solution Décisionnelle

End-to-End

Réalisé par :

Dr :MEKAHLIA FatmaZohra

BOUDERBA TadjEddine

222231244012

GROUPE : 03

1. Introduction

Avec l'évolution des systèmes d'information, les entreprises disposent aujourd'hui de grandes quantités de données issues de leurs activités quotidiennes. Cependant, ces données sont souvent stockées dans des bases transactionnelles (OLTP) qui ne sont pas adaptées à l'analyse décisionnelle.

Le **Business Intelligence (BI)** vise à transformer ces données opérationnelles en informations stratégiques afin d'aider les décideurs à mieux comprendre l'activité de l'entreprise et à prendre des décisions éclairées.

Dans ce contexte, ce projet a pour objectif la **mise en place complète d'une solution BI**, depuis l'extraction des données jusqu'à la visualisation analytique à l'aide de tableaux de bord interactifs, en utilisant la base de données **Northwind** comme cas d'étude.

2. Objectifs du projet

Les principaux objectifs de ce projet sont :

- Mettre en place un **processus ETL (Extract – Transform – Load)** en Python
- Concevoir un **Data Warehouse** basé sur un **schéma en étoile (Star Schema)**
- Centraliser et structurer les données pour l'analyse
- Calculer des **indicateurs clés de performance (KPI)**
- Réaliser des analyses **OLAP multidimensionnelles**
- Créer un **tableau de bord décisionnel interactif**
- Vérifier la qualité et la cohérence des données

3. Présentation de la base de données Northwind

La base de données **Northwind** représente une entreprise fictive spécialisée dans la vente et la distribution de produits.

Elle contient des informations relatives à :

- Clients
- Commandes
- Employés
- Livraisons
- Zones géographiques

Cette base est de type **transactionnel (OLTP)** et n'est pas directement exploitable pour l'analyse décisionnelle, ce qui justifie la construction d'un **Data Warehouse**.

4. Architecture globale du projet

L'architecture du projet BI suit une approche classique en plusieurs couches :

1. Sources de données

- Fichiers issus de la base Northwind

2. ETL (scripts Python)

- Extraction des données
- Nettoyage et transformation
- Chargement dans le Data Warehouse

3. Data Warehouse

- Données structurées selon un modèle multidimensionnel

4. Restitution et analyse

- Dashboards
- KPI
- Graphiques analytiques

5. Organisation du projet

Northwind_project_BI/

```
|__ data/
|   |__ raw/      # Données brutes extraites
|   |__ staging/  # Données nettoyées et transformées
|   |__ warehouse/ # Tables finales du Data Warehouse
|
|__ scripts/     # Scripts ETL
|
|__ notebooks/   # Analyses et dashboard
|
|__ figures/     # Captures des visualisations
```

```

|
└── reports/      # Rapport final
|
└── videos/      # Vidéo de démonstration
|
└── README.md

```

Cette organisation permet une séparation claire entre les différentes étapes du processus BI.

6. Processus ETL

6.1 Extraction

Les données sont extraites à partir des sources et stockées sous forme de fichiers CSV dans le dossier data/raw.

Cette étape permet de conserver une copie brute des données originales.

6.2 Transformation

Les transformations réalisées incluent :

- Normalisation des noms de colonnes
- Suppression des doublons
- Gestion des valeurs nulles
- Conversion des types de données
- Calcul du chiffre d'affaires (total_amount)
- Harmonisation des clés

Les données transformées sont stockées dans le dossier data/staging.

6.3 Chargement

Les données finales sont chargées dans le dossier data/warehouse sous forme de tables analytiques au format CSV et Parquet.

7. Modélisation du Data Warehouse

Le Data Warehouse est modélisé selon un schéma en étoile composé de :

7.1 Table de faits

- FactSales

- Mesures : chiffre d'affaires, nombre de commandes
- Statut de livraison (livrée / non livrée)

7.2 Tables de dimensions

- DimDate : temps (année, mois)
- DimClient : informations clients
- DimEmployee : informations employés

Ce modèle est optimisé pour les requêtes analytiques et les analyses OLAP.

8. Indicateurs clés de performance (KPI)

Les KPI calculés dans le dashboard sont :

- Chiffre d'affaires total
- Nombre total de commandes
- Nombre de commandes livrées
- Nombre de commandes non livrées
- Taux de livraison

Ces indicateurs permettent une vue synthétique et rapide de la performance globale.

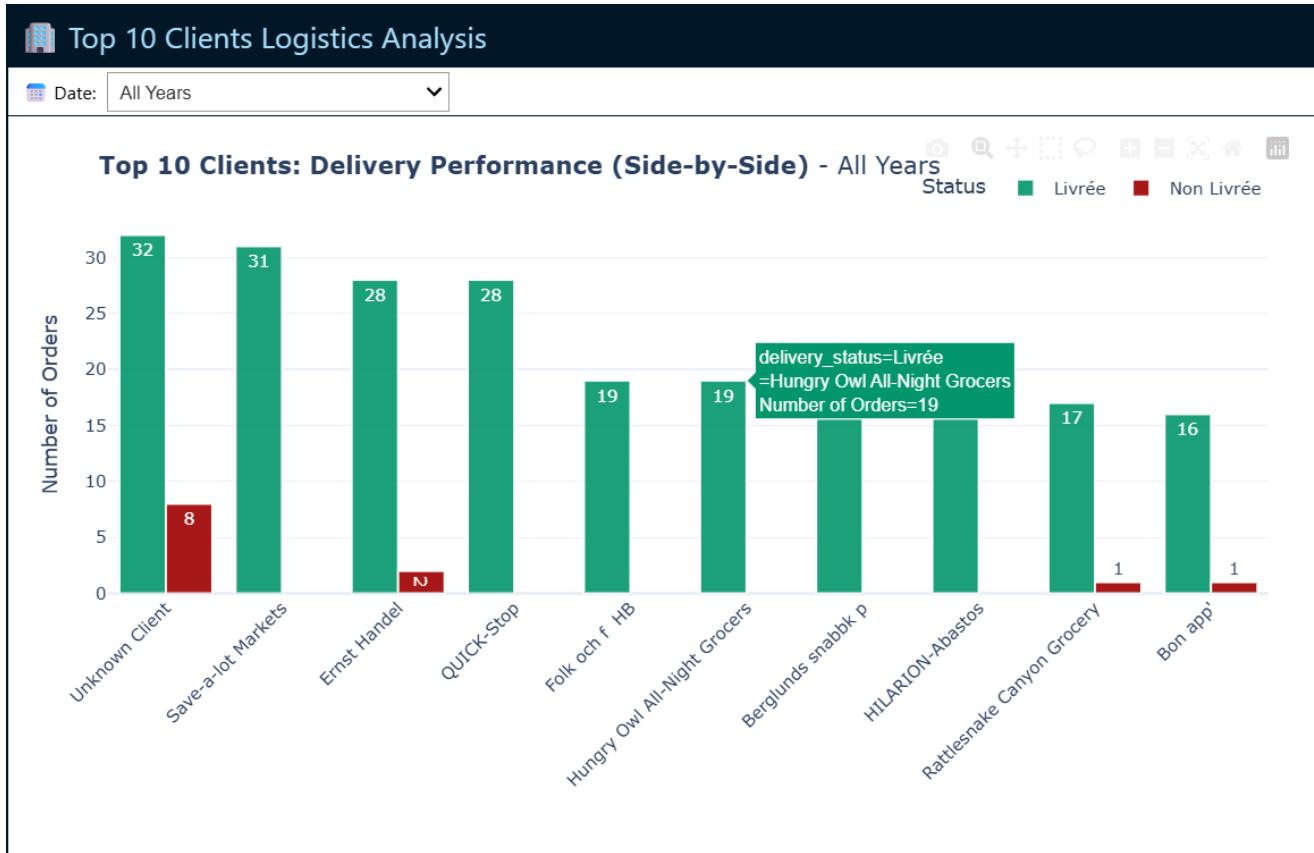
9. Tableau de bord et visualisations

Le tableau de bord interactif (dashboard.ipynb) comprend :

- KPI numériques



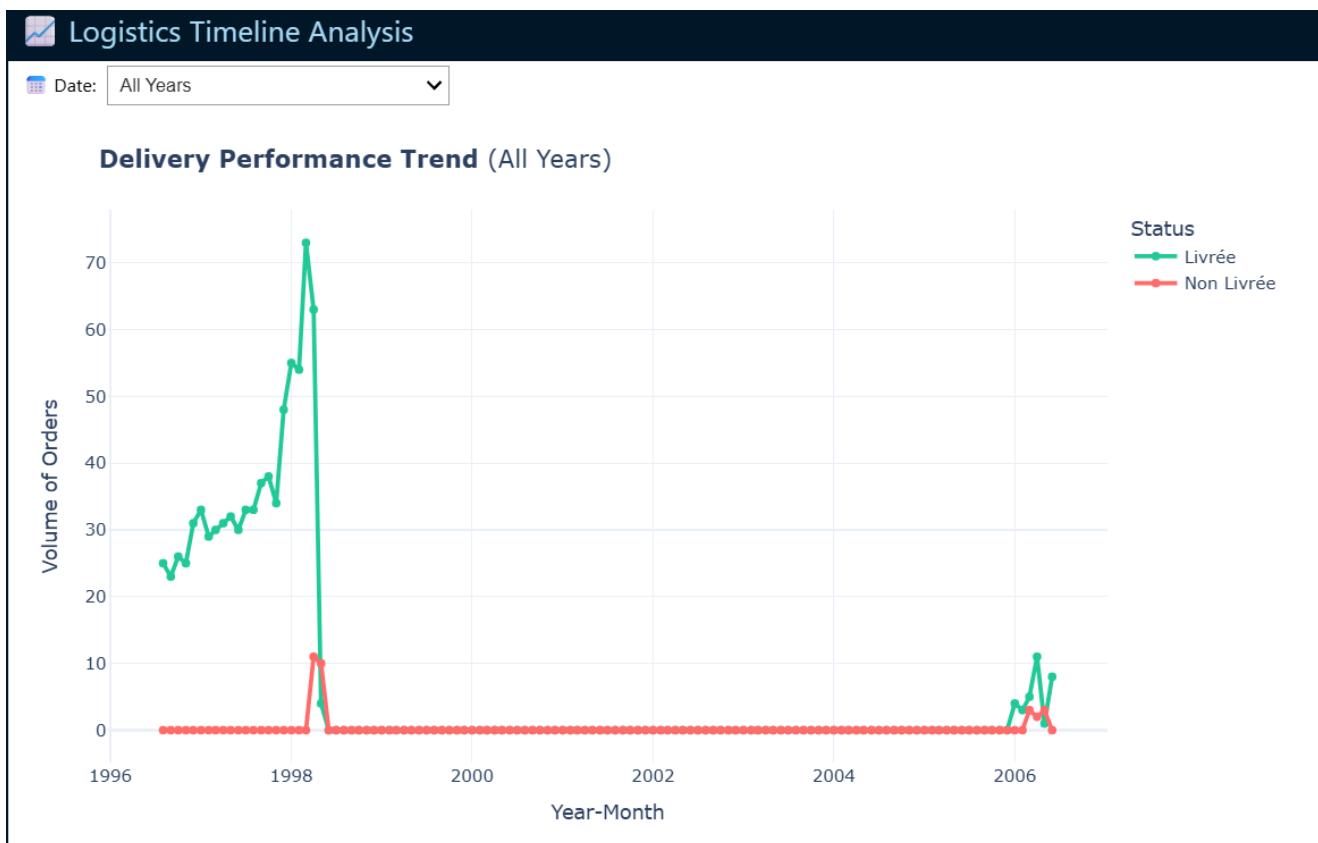
- Graphique livrée / non livrée (vert / rouge)



- Analyse des livraisons par client

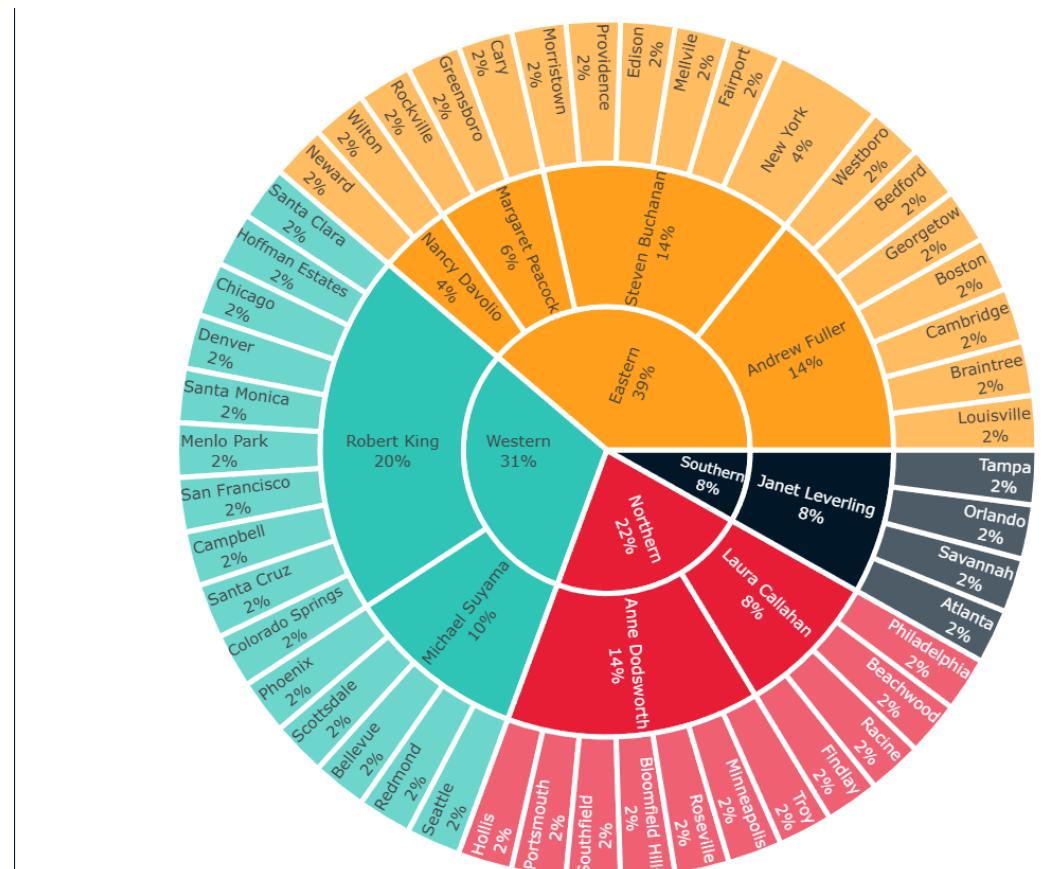
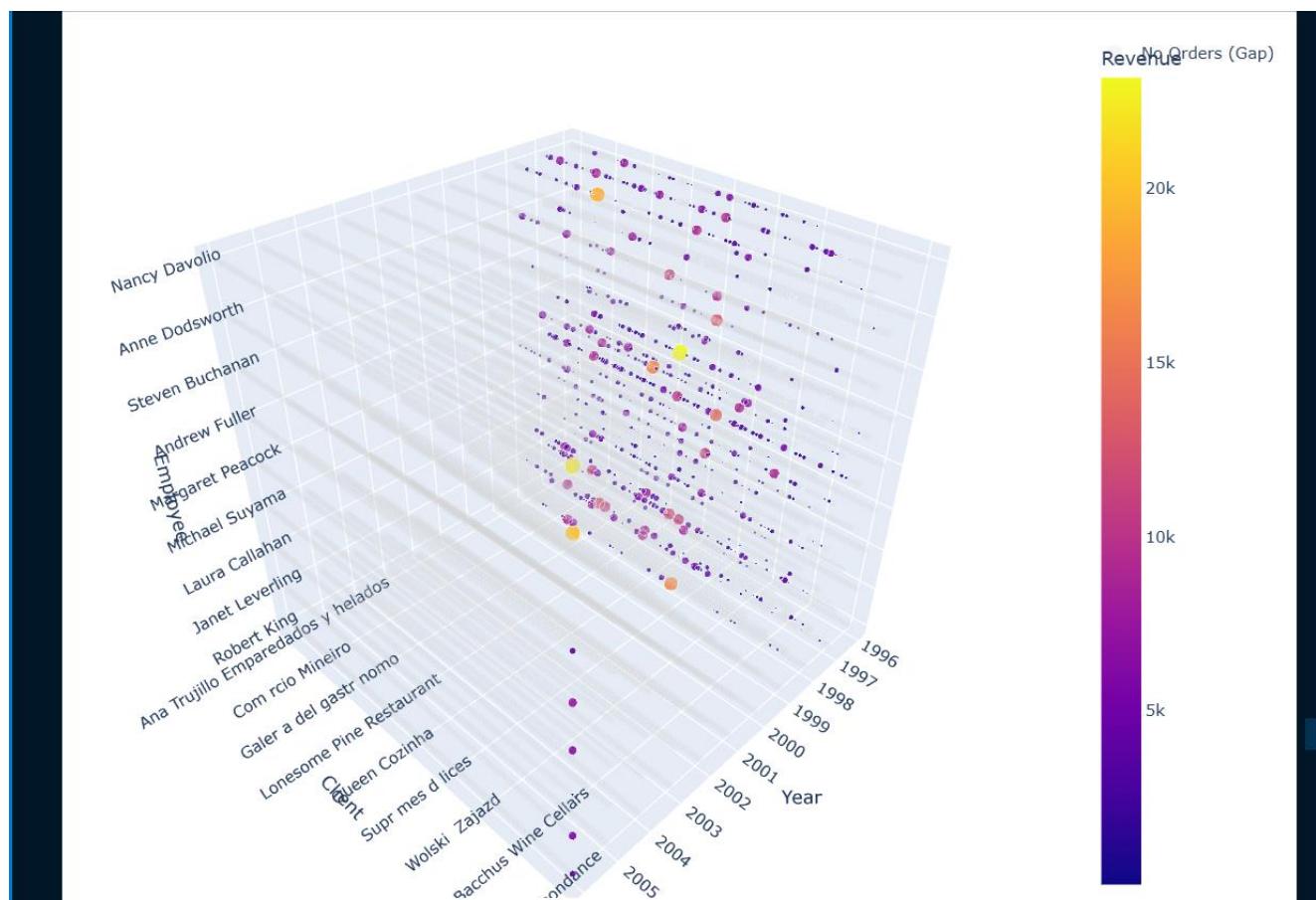


- Analyse par employé



- Visualisation OLAP 3D





Les visualisations sont réalisées à l'aide de Plotly, permettant une interaction avancée (zoom, filtre, rotation).

10. Analyse OLAP 3D

L'analyse OLAP permet une exploration multidimensionnelle des données selon :

- Axe X : Date (année / mois)
- Axe Y : Clients
- Axe Z : Employés
- Couleur : Chiffre d'affaires

Cette analyse facilite l'identification des tendances, des périodes fortes et des acteurs les plus performants.

11. Validation et qualité des données

Des contrôles ont été réalisés afin de garantir :

- L'absence de doublons
- La cohérence des clés
- La correspondance entre données brutes et données du Data Warehouse

Ces vérifications assurent la fiabilité des analyses produites.

12. Conclusion

Ce projet a permis de mettre en œuvre l'ensemble des concepts fondamentaux du Business

Intelligence :

- ETL
- Data Warehouse
- Modélisation multidimensionnelle
- Analyse OLAP
- Tableaux de bord décisionnels

La solution développée fournit une vision claire, synthétique et exploitable des données de vente et constitue une base solide pour l'aide à la décision.