Analyse de la performance des stations de distribution de pétrole en Tunisie

Réalisé par : Yoldez Hfaiedh, Khadija Ennaifer et Amira Moadeb Année universitaire 2024-2025

Table des matières

1	Introduction					
	Étapes du projet					
	2.1 Analyse des sources de données					
	2.2 Préparation et transformation des données					
	2.3 Implémentation ETL sous Talend					
	2.4 Identification des KPIs					
	2.5 Visualisation avec Power BI					
3	Conclusion					

1 Introduction

Dans un environnement économique de plus en plus concurrentiel, les entreprises opérant dans le secteur pétrolier doivent s'appuyer sur des outils décisionnels efficaces pour piloter leur activité et prendre des décisions stratégiques éclairées. Face à la diversité des branches commerciales — Carburants, Gaz, Lubrifiants et Aviation — il devient essentiel de disposer d'un système capable de centraliser, structurer et valoriser les données issues de ces différentes activités.

Ce projet s'inscrit dans cette logique. Il a pour objectif de concevoir une application d'aide à la décision dédiée au suivi et à l'analyse des performances commerciales. À travers une chaîne décisionnelle complète — allant de l'analyse des sources de données à la visualisation des indicateurs via Power BI — ce projet vise à mettre en place un datamart optimisé, alimenté par des flux ETL conçus sous Talend, permettant de restituer des tableaux de bord dynamiques et interactifs.

Ce rapport retrace l'ensemble des étapes de réalisation du projet, depuis l'étude des données sources jusqu'à la création des visualisations finales, en passant par la modélisation, le choix des indicateurs clés (KPIs) et le développement des processus ETL.

2 Étapes du projet

2.1 Analyse des sources de données

La première étape du projet a consisté à explorer les données mises à disposition, fournies sous forme de fichiers plats (CSV/Excel). Ces fichiers contiennent des informations de vente détaillées par transaction.

Les attributs clés de ces fichiers incluent :

- Code produit et libellé produit
- Quantité vendue
- Montant total
- Client et segment client
- Région et point de vente
- Canal de distribution

Voici quelques aperçus des fichiers sources.

Voici un extrait du fichier de codification des clients :

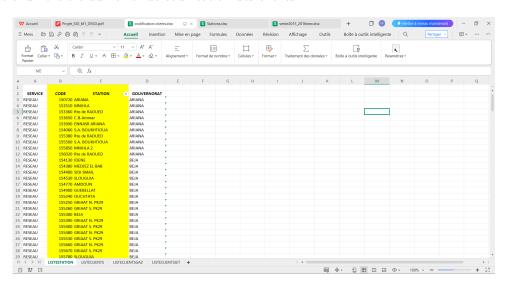


FIGURE 1 – Fichier Excel — Codification des clients

Voici le fichier contenant les informations sur les stations de distribution :

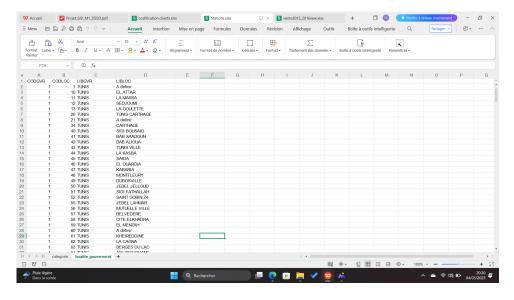


FIGURE 2 – Fichier Excel — Informations sur les stations

Voici le fichier regroupant les données de vente par transaction :

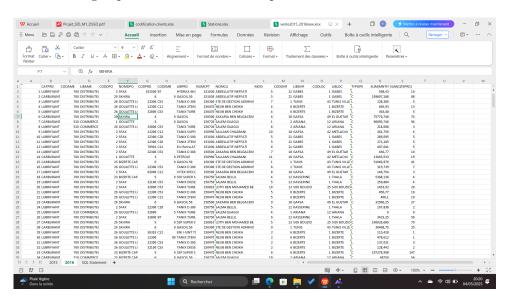


FIGURE 3 – Fichier Excel — Détail des ventes

2.2 Préparation et transformation des données

Afin de garantir la qualité des données, une phase de nettoyage a été réalisée à l'aide d'un script Python. Cette étape correspond à la zone de *staging* dans un processus décisionnel.

Les principales opérations appliquées sont :

- Suppression des lignes et colonnes entièrement vides
- Élimination des doublons
- Standardisation des noms de colonnes (minuscules, sans espaces)
- Suppression manuelle des colonnes non pertinentes

Voici un aperçu du script Python utilisé pour ces transformations :

FIGURE 4 – Extrait du script Python — Partie 1

Figure 5 – Extrait du script Python — Partie 2

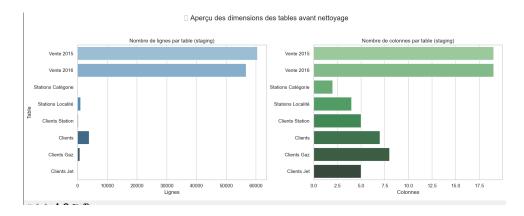


FIGURE 6 – Extrait du script Python — Partie 3

Voici les fichiers générés après nettoyage :

s clients_clean	0	22/04/2025 18:43	Feuille de calcul X	220 Ko	
S clients_station_clean	•	19/04/2025 19:47	Feuille de calcul X	18 Ko	
S clients_gaz_clean	•	19/04/2025 19:37	Feuille de calcul X	53 Ko	
s clients_jet_clean	•	19/04/2025 19:37	Feuille de calcul X	13 Ko	
station_categorie_clean	•	19/04/2025 19:37	Feuille de calcul X	7 Ko	
station_localite_clean	•	19/04/2025 19:37	Feuille de calcul X	28 Ko	
s vente_2016_clean	•	19/04/2025 19:37	Feuille de calcul X	5513 Ko	
vente_2015_clean	•	19/04/2025 19:36	Feuille de calcul X	5834 Ko	

FIGURE 7 – Fichiers CSV générés après nettoyage

2.3 Implémentation ETL sous Talend

La phase suivante a consisté à concevoir les flux ETL (Extract, Transform, Load) avec l'outil Talend Open Studio. Ces flux permettent de charger les données nettoyées dans un schéma en étoile comportant une table de faits et plusieurs dimensions.

Voici un aperçu de l'implémentation des différentes dimensions :

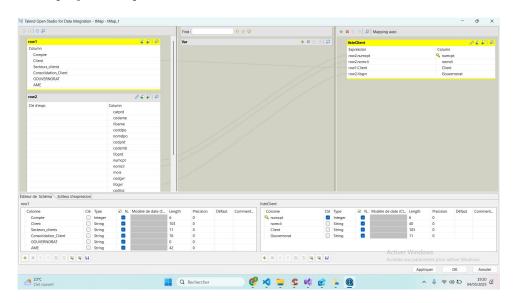


FIGURE 8 – Composant t
Map — Transformation des données client

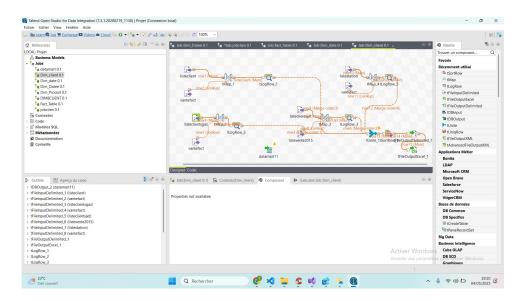


Figure 9 – Dimension Client — Structure cible

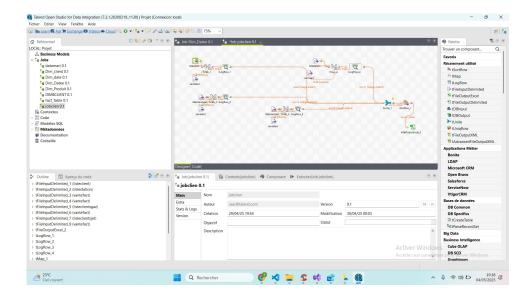


FIGURE 10 – Job Talend — Chargement de la dimension Date

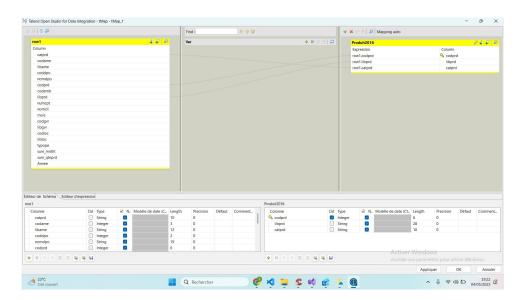


FIGURE 11 – Composant t
Map — Transformation des données produit

Voici la structure de la table de faits vente :

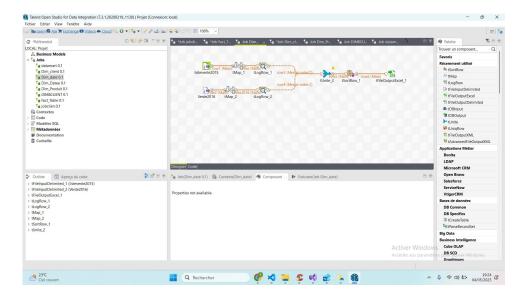


FIGURE 12 – Table de faits — Vente

Voici le flux Talend alimentant cette table de faits :

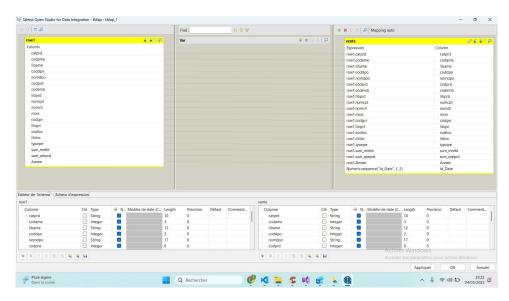


Figure 13 – Composant t
Map — Alimentation de la table de faits vente

2.4 Identification des KPIs

Pour orienter la visualisation et l'analyse, plusieurs indicateurs clés de performance (KPIs) ont été identifiés en collaboration avec les parties prenantes métiers. Ces indicateurs sont les suivants :

- Chiffre d'affaires par région, station et produit
- Volume de vente par canal et type de client
- Top produits en termes de ventes
- Répartition géographique et temporelle des ventes
- Évolution mensuelle du chiffre d'affaires
- Nombre total de transactions
- Nombre de clients par année
- Nombre de ventes
- Nombre de clients par mois
- Nombre de clients par gouvernorat
- Répartition géographique des clients (carte)
- Proportion de clients par type de produit
- Montant HT total
- Quantité vendue totale

2.5 Visualisation avec Power BI

Une fois les données chargées et modélisées, nous avons conçu un tableau de bord interactif dans Power BI. Il est structuré en plusieurs pages thématiques :

- Page d'Accueil : Vue synthétique de la composition du dashboard
- Vue d'ensemble : Analyse globale des ventes par période, région, canal
- **Analyse des mesures** : Exploration par station, produit et client des mesures(montant hors taxe et da quantité totale vendue)

Voici les différentes captures d'écran des pages Power BI créées :

Voici la page d'accueil du tableau de bord Power BI:

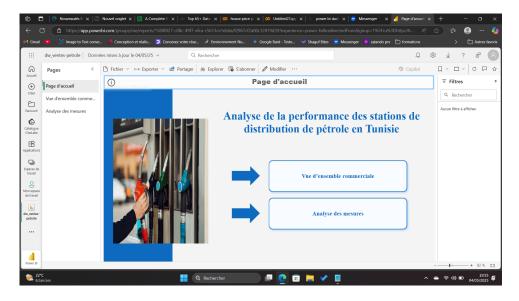


FIGURE 14 – Page d'accueil — Tableau de bord Power BI

Voici la vue d'ensemble :

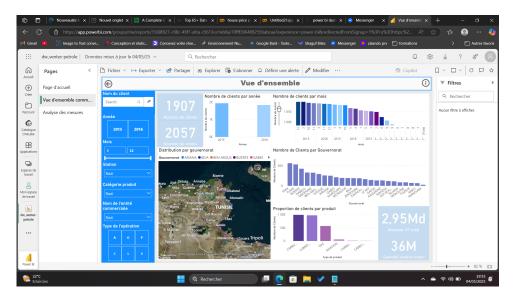


Figure 15 – Vue d'ensemble — Analyse globale

Voici une analyse détaillée :

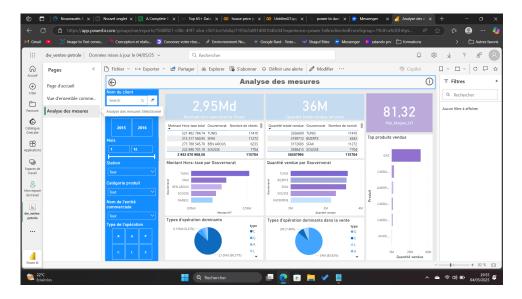


FIGURE 16 – Analyse des mesures

3 Conclusion

Ce projet a permis la mise en œuvre d'une solution complète de Business Intelligence dédiée au secteur pétrolier. Grâce à une architecture décisionnelle bien pensée, l'intégration automatisée via Talend et des visualisations pertinentes sous Power BI, l'entreprise peut désormais suivre et analyser efficacement la performance de ses stations de distribution.

Cette démarche permet non seulement de centraliser les données mais aussi d'améliorer la prise de décision à tous les niveaux de l'organisation. Les indicateurs dynamiques offrent une meilleure compréhension de l'activité commerciale et facilitent le suivi des objectifs.