

Module : Business Intelligence

Filière & Semestre : CI-ISBD/S8

Mai 2025

Compte Rendu :

TP Entrepôt de Données - Production Industrielle au Maroc (2005-2013)

Réalisé complètement avec SSIS

Encadré par :

- Prof. Hamid Himech

Réalisé par :

- GARHOUM Hafid

Table des matières

| | |
|-----------------------------------------------------------|----------|
| 1. Introduction | 3 |
| 2. Modélisation de l'entrepôt de données | 3 |
| 3. Phase d'intégration ETL avec SSIS | 4 |
| 4. Chargement de l'entrepôt final via SSIS | 7 |
| 5. Visualisation et Prédiction avec Power BI | 8 |
| 6. Conclusion..... | 9 |

1. Introduction

Ce projet a pour objectif la construction d'un entrepôt de données (Data Warehouse) sur la production industrielle au Maroc entre 2005 et 2013. L'analyse se base sur deux sources principales : les statistiques industrielles et les données pluviométriques.

Les indicateurs étudiés sont :

- Le taux de croissance en %
- Le taux de valeur ajoutée
- Le taux d'investissement
- Le taux d'exportation

Les branches industrielles concernées :

- Agro-alimentation
- Textile et cuir
- Chimie et parachimie
- Mécanique et métallurgique
- Électricité et électronique

2. Modélisation de l'entrepôt de données

2.1. Table de faits : fact_production

Contient les mesures quantitatives liées aux indicateurs.

```
CREATE TABLE fact_production (  
  id_fact INT IDENTITY(1,1) PRIMARY KEY,  
  id_temps INT,  
  id_branche INT,  
  id_indicateur INT,  
  valeur FLOAT,  
  FOREIGN KEY (id_temps) REFERENCES dim_temps(id_temps),  
  FOREIGN KEY (id_branche) REFERENCES dim_branche(id_branche),  
  FOREIGN KEY (id_indicateur) REFERENCES dim_indicateur(id_indicateur)  
);
```

2.2. Tables de dimensions

a. dim_temps : contient les années et la situation pluviométrique

```
CREATE TABLE dim_temps (  
  id_temps INT PRIMARY KEY,  
  annee INT UNIQUE,  
  pluviometrie_annuelle FLOAT,  
  situation NVARCHAR(50)  
);
```

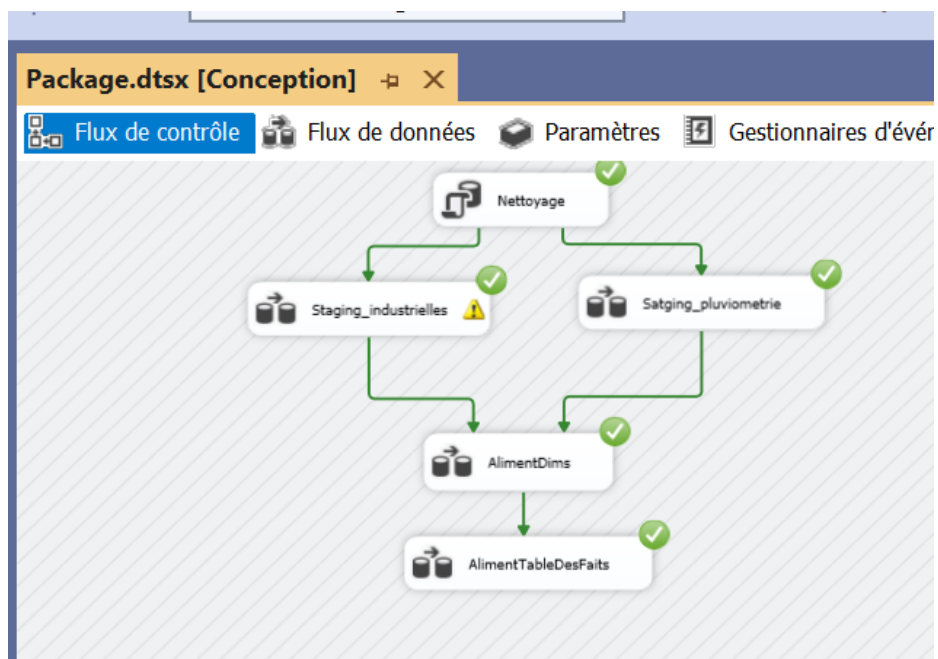
b. dim_branche : contient les branches industrielles

```
CREATE TABLE dim_branche (  
  id_branche INT IDENTITY(1,1) PRIMARY KEY,  
  nom_branche NVARCHAR(100) UNIQUE  
);
```

c. dim_indicateur : contient les types d'indicateurs

```
CREATE TABLE dim_indicateur (  
  id_indicateur INT IDENTITY(1,1) PRIMARY KEY,  
  nom_indicateur NVARCHAR(100) UNIQUE  
);
```

3. Phase d'intégration ETL avec SSIS

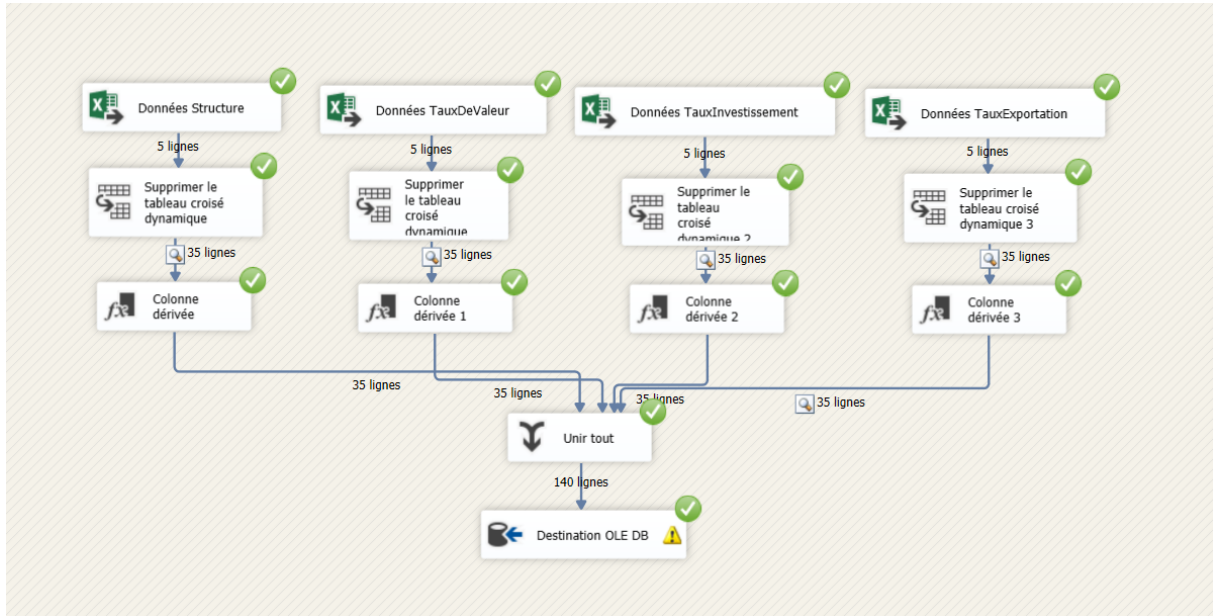


3.1. Source de données

- **Fichier Excel** : statistiques industrielles 2006-2012
- **Fichier Excel** : pr_1991_2015

3.2. Flux de données SSIS : Préparation des données

a. Data Flow Task : Staging_industrielles



1. Excel Source : sélection de plages comme SELECT * FROM [Feuil1\$A4:H8] (par bloc)
2. Unpivot : transformation des années en lignes
3. Derived Column : ajout de colonne "Indicateur" à valeur dans l'ensemble {'Structure', 'TauxDeValeur', 'TauxInvestissement', 'TauxExportation'}
4. Union All : union des 4 indicateurs (Structure, ValeurAjoutée, Investissement, Exportation)
5. OLE DB Destination : insertion dans staging_indust

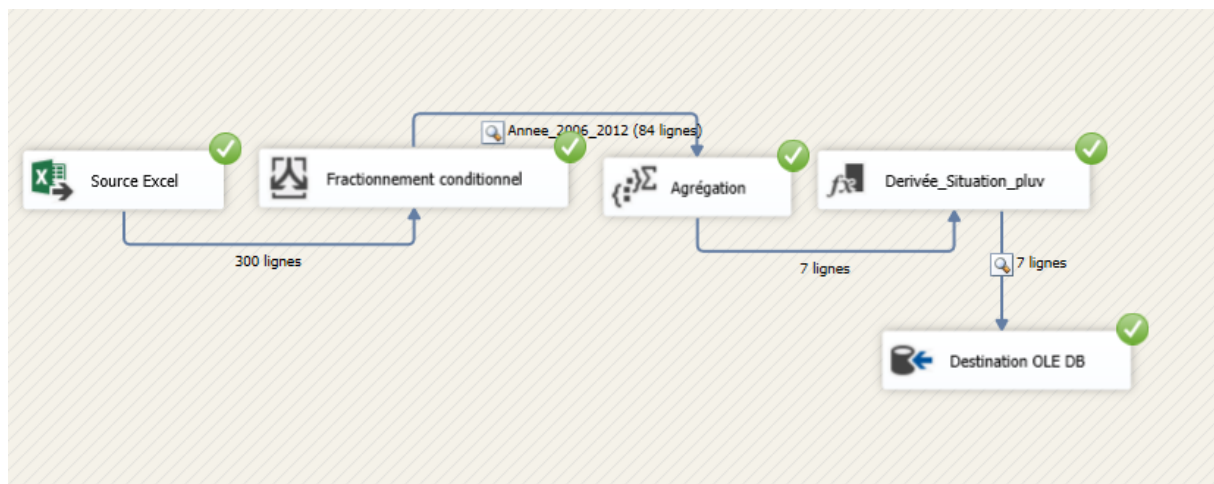
Table staging :

```

CREATE TABLE staging_indust (
  annee INT,
  branche NVARCHAR(100),
  indicateur NVARCHAR(50),
  valeur FLOAT
);
  
```

| | Branche | Annee | Valeur | Indicateur |
|----|-----------------------|-------|------------------|--------------|
| 67 | Électricité et éle... | 2009 | 24,7660626097007 | TauxDeValeur |
| 68 | Électricité et éle... | 2010 | 26,9184268772797 | TauxDeValeur |
| 69 | Électricité et éle... | 2011 | 23,981799350382 | TauxDeValeur |
| 70 | Électricité et éle... | 2012 | 24,2973376841344 | TauxDeValeur |
| 71 | Agro-alimentation | 2006 | 28,1643752890493 | Structure |
| 72 | Agro-alimentation | 2007 | 28,3699666526222 | Structure |
| 73 | Agro-alimentation | 2008 | 26,9439830443455 | Structure |
| 74 | Agro-alimentation | 2009 | 31,3278719917532 | Structure |
| 75 | Agro-alimentation | 2010 | 28,556783237201 | Structure |
| 76 | Agro-alimentation | 2011 | 26,2229843357307 | Structure |
| 77 | Agro-alimentation | 2012 | 27,1388783161015 | Structure |

b. Data Flow Task : Staging_pluviometrie



- Excel File
- Derived Column pour nettoyer ou renommer les colonnes si besoin
- OLE DB Destination vers staging_pluviometrie

Table staging :

```
CREATE TABLE staging_pluviometrie (  
  annee INT,  
  pluviometrie_annuelle FLOAT,  
  situation NVARCHAR(50)  
);
```

| | Annee | PluviometrieAnnuelle | Situation |
|---|-------|----------------------|-----------|
| 1 | 2006 | 347.2393 | Normale |
| 2 | 2007 | 226.7025 | Normale |
| 3 | 2008 | 347.7388 | Normale |
| 4 | 2009 | 360.4693 | Pluvieuse |
| 5 | 2010 | 432.4490 | Pluvieuse |
| 6 | 2011 | 362.9758 | Pluvieuse |
| 7 | 2012 | 332.6006 | Normale |

4. Chargement de l'entrepôt final via SSIS

Etapes dans SSIS :

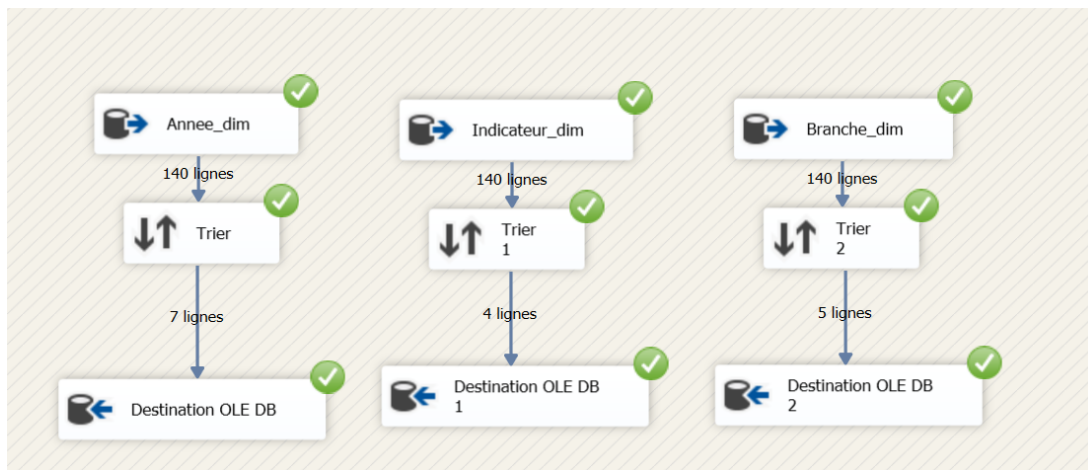
1. Créer un Execute SQL Task pour vider les tables si besoin :

```
DELETE FROM fact_production;  
DELETE FROM dim_temps;  
DELETE FROM dim_branche;  
DELETE FROM dim_indicateur;
```

(Si contraintes FK : supprimer les FK, puis TRUNCATE, puis recréer FK)

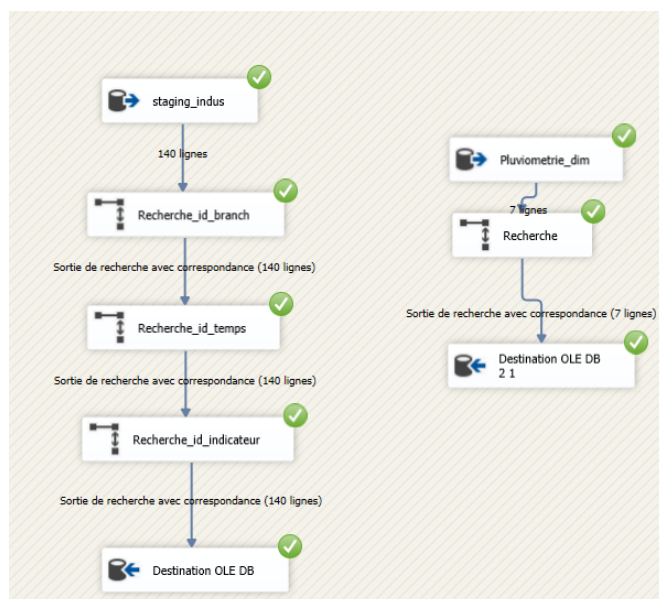
2. **Data Flow Task : AlimentDims**

Utiliser des Lookup (Recherche) pour alimenter les dimensions dim_branche, dim_indicateur, dim_temps



3. **Data Flow Task : AlimentTableDesFaits**

Utiliser des composants Lookup pour faire correspondre les identifiants (ID) aux dimensions existantes dans la table de faits, et intégrer les données de pluviométrie dans la dimension **dim_pluviometrie**, en établissant une relation avec la dimension temporelle via **id_temps** provenant de **dim_temps**.



4. Destination finale : fact_production

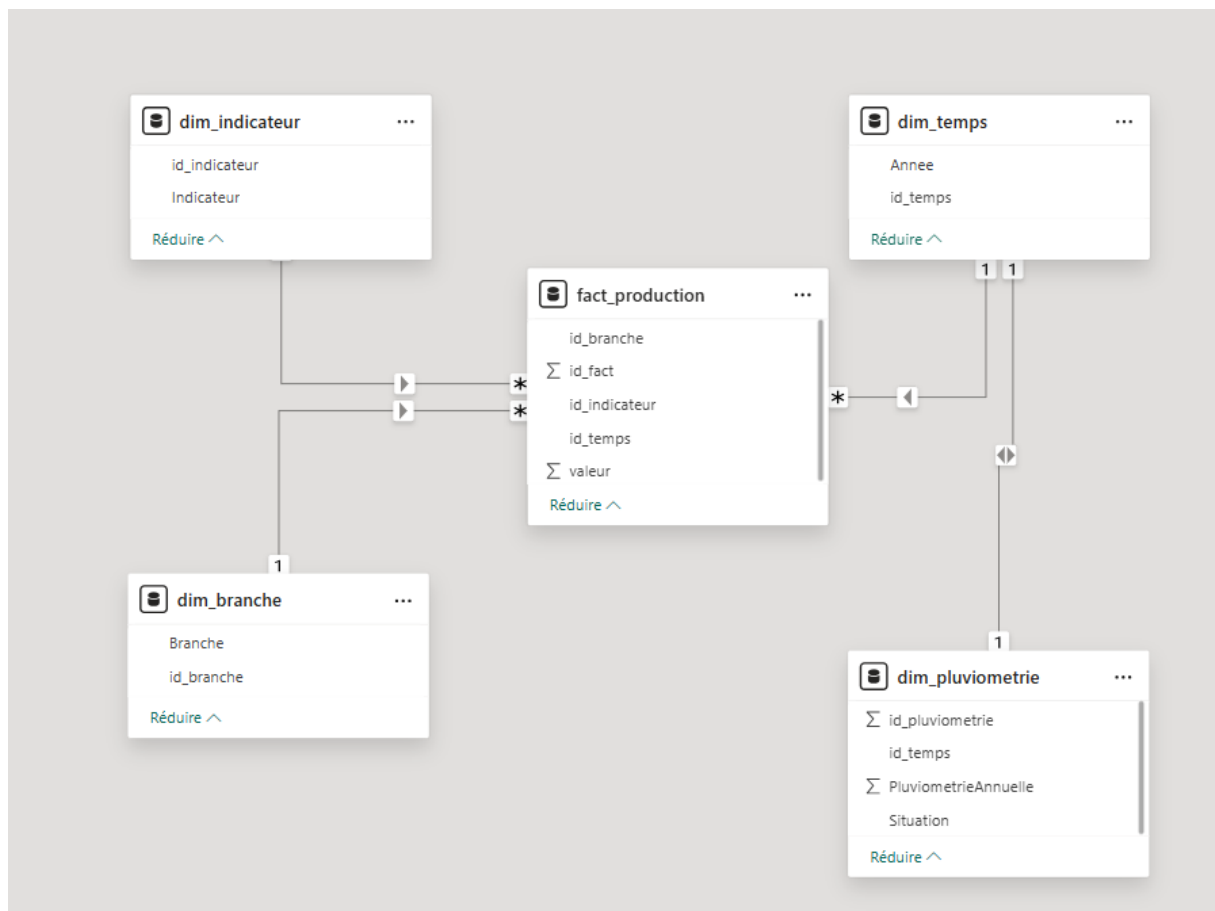
| | id_fact | id_temps | id_branche | id_indicateur | valeur |
|----|---------|----------|------------|---------------|------------------|
| 1 | 1 | 36 | 21 | 19 | 15,7415775677227 |
| 2 | 2 | 37 | 21 | 19 | 15,4075411917842 |
| 3 | 3 | 38 | 21 | 19 | 15,0048966894142 |
| 4 | 4 | 39 | 21 | 19 | 14,0975229361472 |
| 5 | 5 | 40 | 21 | 19 | 12,6417268268278 |
| 6 | 6 | 41 | 21 | 19 | 12,0840850243434 |
| 7 | 7 | 42 | 21 | 19 | 12,6462043366197 |
| 8 | 8 | 36 | 25 | 19 | 65,7989948959283 |
| 9 | 9 | 37 | 25 | 19 | 64,4204815803166 |
| 10 | 10 | 38 | 25 | 19 | 62,9746024627937 |
| 11 | 11 | 39 | 25 | 19 | 60,7950462216327 |

Remarque

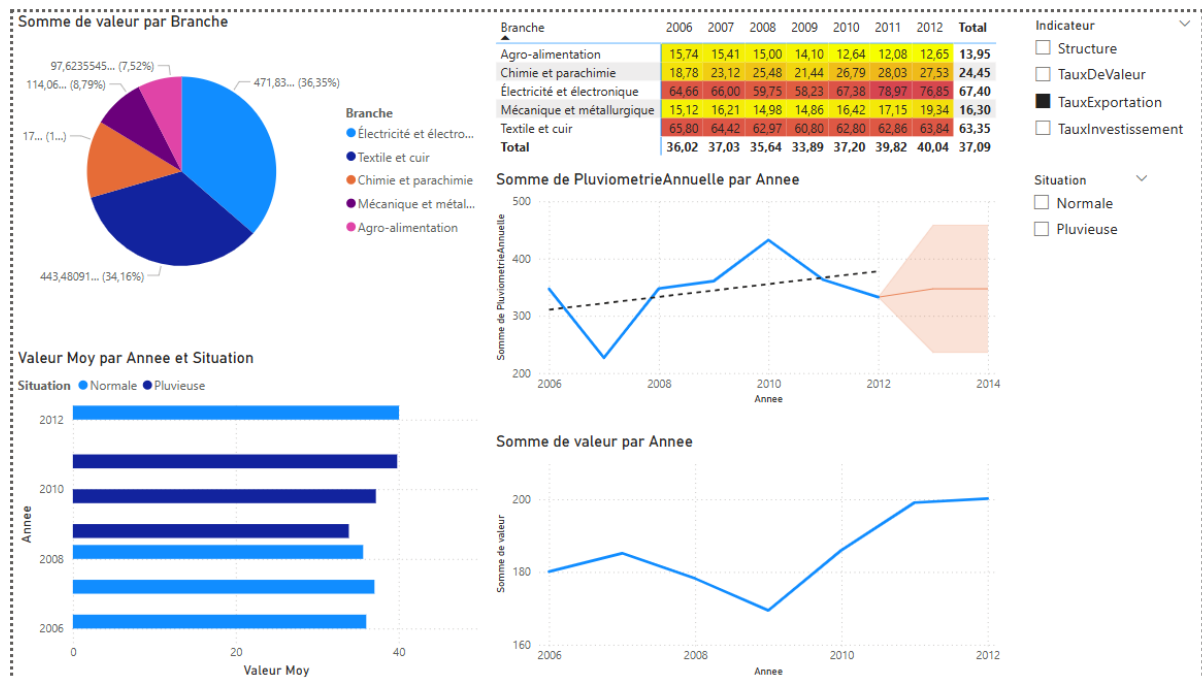
Les données pluviométriques, quand c'est nécessaire, peuvent être reliées à la table de faits via une **jointure avec id_temps** dans dim_temps.

5. Visualisation et Prédiction avec Power BI

5.1. Schéma des données



- Courbes d'évolution par année
- Comparaison entre branches
- Impact de la pluviométrie sur l'investissement ou la production
- Prédiction via Power BI appliquées aux données de pluviométrie



6. Conclusion

Ce projet a permis de maîtriser les différentes étapes de création d'un entrepôt de données, depuis l'intégration jusqu'à la visualisation. SSIS a joué un rôle crucial dans la transformation et l'automatisation des flux de données. Power BI a permis d'interpréter les résultats de manière visuelle et intuitive.