



- **Mise en place** d'un pipeline ETL automatisé avec SSIS et **Réalisation** d'un tableau de bord du chiffre d'affaire via Power BI –

Année Universitaire : 2024-2025

Ingénierie de l'Information Numérique

Remerciements

Je tiens d'emblée à remercier M. Salaheddine BAHJI, le directeur de l'ESI, pour tous ses efforts afin d'améliorer notre vie estudiantine, ainsi que le personnel responsable du stage, et tous les professeurs qui essayent de nous transmettre le maximum d'informations d'une manière pertinente.

Je veux aussi remercier mon encadrant M. Achraf Karmaoui pour répondre à toutes mes questions afin de livrer un projet de qualité et de répondre à leurs besoins, ainsi que le responsable technique de l'entreprise qui n'a jamais hésiter de nous soutenir.

J'exprime également ma gratitude envers ma famille qui m'a soutenue depuis le début de mon parcours scolaire.

Resumé

Mon expérience du stage PFA effectué à Nouvelle Aveiro Maroc était sous forme d'une observation de la ligne de production de leurs produits conservés (Maquereaux, Sardines ...), d'une compréhension de la structure de l'entreprise, et bien sûr un projet BI où j'ai répondu à leurs besoins et contraintes tout en réalisant un tableau de bord du chiffre d'affaire avec un pipeline ETL automatisé pour extraire les données des fichiers Excel fournis vers une base de données SQL après avoir passé sur un processus de traitement. Et tout cela en communiquant avec le directeur SI afin de collecter les besoins métier en vue de diriger le résultat final vers leurs objectifs.

Table de matières

Chapitre I : Introduction de l'entreprise.....	6
Introduction.....	6
Organigramme de l'entreprise.....	7
Visite au site de production.....	8
Chapitre II : Conception du projet.....	9
Introduction.....	10
Conception du projet.....	11
Conclusion.....	13
Chapitre III : Réalisation du projet.....	14
Introduction.....	14
Les outils abordés.....	14
Réalisation du projet.....	15
1- Implémentation du Pipeline ETL sur SSIS.....	15
2- Déploiement du package ETL.....	24
3- Création d'un Job pour automatiser l'exécution du pipeline.....	25
4- Connexion de Power BI avec la base de données Aveiro.....	26
5- Création du tableau de bord.....	27
Résultats.....	29
Livrables du projet.....	29
Choses a améliorer.....	30
Conclusion.....	30
Conclusion générale.....	30
Webographie.....	31

Table des figures / annexe

Figure 1 : Organigramme de l'entreprise.....	8
Figure 2 : Schéma de la ligne de production.....	9
Figure 3 : Schéma des étapes du projet.....	10
Figure 4: Logo de visual studio.....	14
Figure 5: Logo de SQL Server.....	14
Figure 6 : Logo de SSMS.....	14
Figure 7 : Logo de Power BI.....	14
Figure 8 : Le pipeline ETL.....	15
Figure 9 : Ajout des fichiers Excel de la destination.....	16
Figure 10 : Suppression des ventes nulles.....	17
Figure 11 : Suppression des ventes invalides.....	18
Figure 12 : Remplacement des marques vide par Lukus.....	19
Figure 13 : Elimination des doublons.....	20
Figure 14 : Remplacement des marques vide par Lukus.....	21
Figure 15 : Table destination avec toutes les données traitées.....	22
Figure 17 : la forme du package ETL.....	23
Figure 18 : La tâche Execute SQL.....	23
Figure 19 : Déploiement du package.....	24
Figure 21 : Création d'un calendrier pour le job.....	26
Figure 22 : Création de connexion.....	26
Figure 24 : La mesure CA. par année.....	27
Figure 25: La table mesure pour arranger les mesures calculées.....	27
Figure 26: Tableau de bord initial.....	28
Figure 27: Mesure pour la ligne de proportion.....	29
Figure 28: Tableau de bord final avec menu.....	29

Tableau des acronymes

Acronyme	Signification
ESI	Ecole des Sciences de l'Information
BI	Business Intelligence
ETL	Extract, Transform, Load
SSIS	SQL Server Integration Services
SQL	Structured Query Language
SI	Système d'Information
ERP	Enterprise Resource Planning
KPI	Key Performance Indicators
Qté	Quantité
S/ce	Service
SSMS	SQL Server Management Studio
C.A.	Chiffre d'affaire

Chapitre I : Introduction de l'entreprise

Introduction

Groupe Bicha, leader national dans l'industrie de la conserve.

Le GROUPE BICHA, leader dans la conserve de poissons et végétale, mise sur la qualité, la diversification et des équipements modernes pour produire selon les normes internationales. Avec ses trois unités (**AVEIRO MAROC, DOHA, SESTA**), il jouit d'une forte notoriété et oriente sa stratégie vers des activités à forte valeur ajoutée et durables.

La société AVEIRO Maroc.

La société AVEIRO Maroc, fondée en 1946, est spécialisée dans la production et la commercialisation de produits de la mer en conserve, soigneusement sélectionnés et pêchés sur les côtes marocaines.

Tous les produits de l'entreprise sont fabriqués dans une installation respectant les normes les plus exigeantes des clients internationaux et des chaînes de distribution alimentaire de renommée mondiale, avec une capacité installée d'environ 40 000 tonnes de matières premières.

Nouvelle AVEIRO Maroc propose une gamme complète de conserves de poissons : sardines, maquereaux et thon, commercialisés sous quatre marques internationales : **DELMONACO – LIBERATOR – LIMAO – SAMAKI**, ainsi que d'autres marques, importateurs et distributeurs.

L'expérience, le savoir-faire et l'investissement constant dans un équipement de pointe ont permis à Nouvelle AVEIRO Maroc d'étendre son activité à l'échelle internationale, avec la grande portion de son chiffre d'affaires réalisé à l'export, vers l'Europe, les États-Unis, l'Amérique latine, l'Afrique et le Moyen-Orient.

Organigramme de l'entreprise

Cet organigramme explique la structure organisationnelle de l'entreprise
Nouvelle Aveiro Maroc :



Figure 1 : Organigramme de l'entreprise

Visite au site de production

Pour comprendre plus la mission de l'entreprise, on a effectué une visite à l'usine de production tout en passant par tous les points existants un par un dès la réception de la matière première, jusqu'à l'emballage final et la distribution.

Voici une représentation de la ligne de production au sein de Nouvelle Aveiro Maroc :

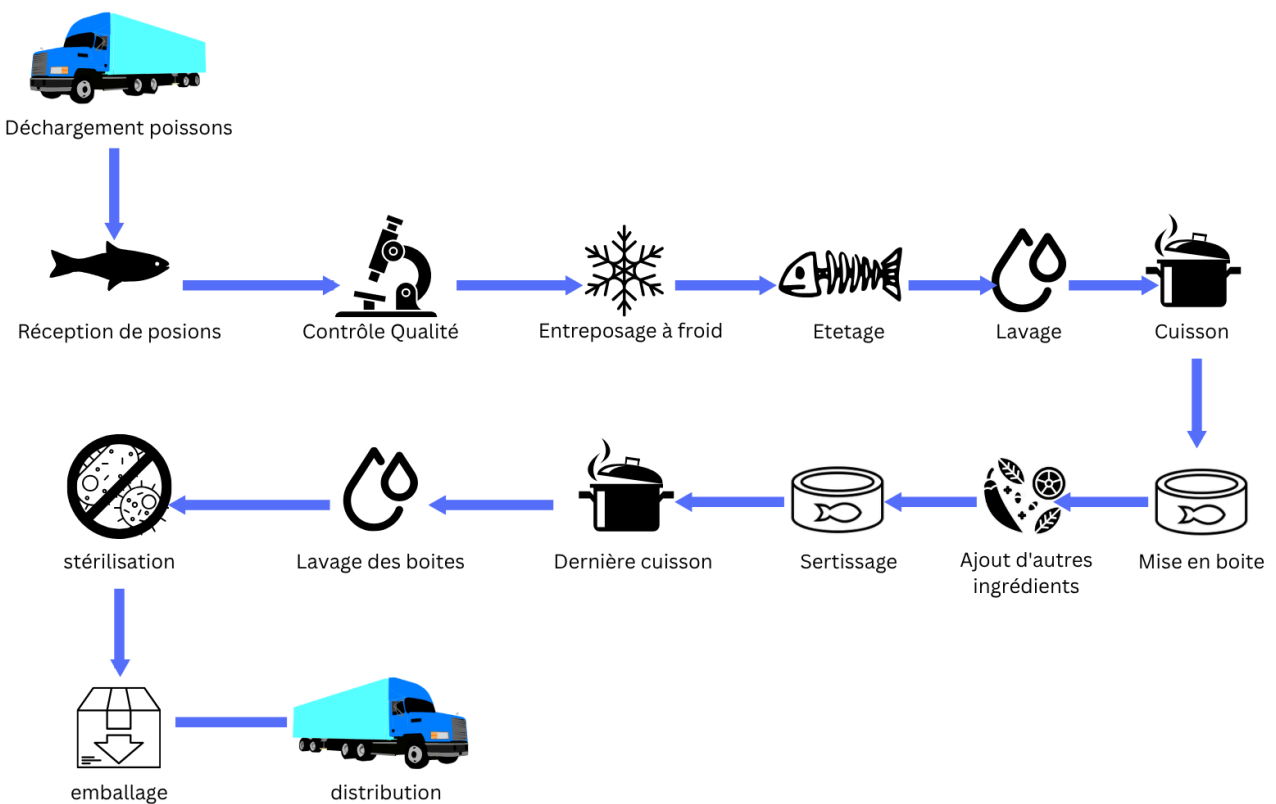


Figure 2 : Schéma de la ligne de production

Chaque étape passe par des contrôles pour vérifier le respect des règles de production (dosage, hygiène, stockage ...)

Chapitre II : Conception du projet

Introduction

La société Nouvelle Aveiro Maroc aborde l'outil français YourCegid comme solution ERP pour contrôler l'ensemble de leurs activités, cet outil est en cours de mise à jour afin pour intégrer des tableaux de bord automatisés en temps réel, d'un autre côté cet organisme aborde Power BI pour l'instant afin de réaliser leurs tableaux de bord de vente ou d'autres objectifs.

Dans ce projet, j'ai réalisé un tableau de bord visualisant l'ensemble des KPIs qui influencent le chiffre d'affaires de l'entreprise. Selon des besoins exprimés par l'entreprise, j'ai compris qu'actuellement ils ont besoin d'une étude du chiffre d'affaires entre 2021 et 2025 pour observer son progrès ainsi que pour voir les facteurs qui l'affectent le plus (clients, marques, ...) avec un tableau de bord clair et compréhensible par tout lecteur autorisé.

Voici ci-dessous un schéma expliquant l'ensemble des étapes du projet:

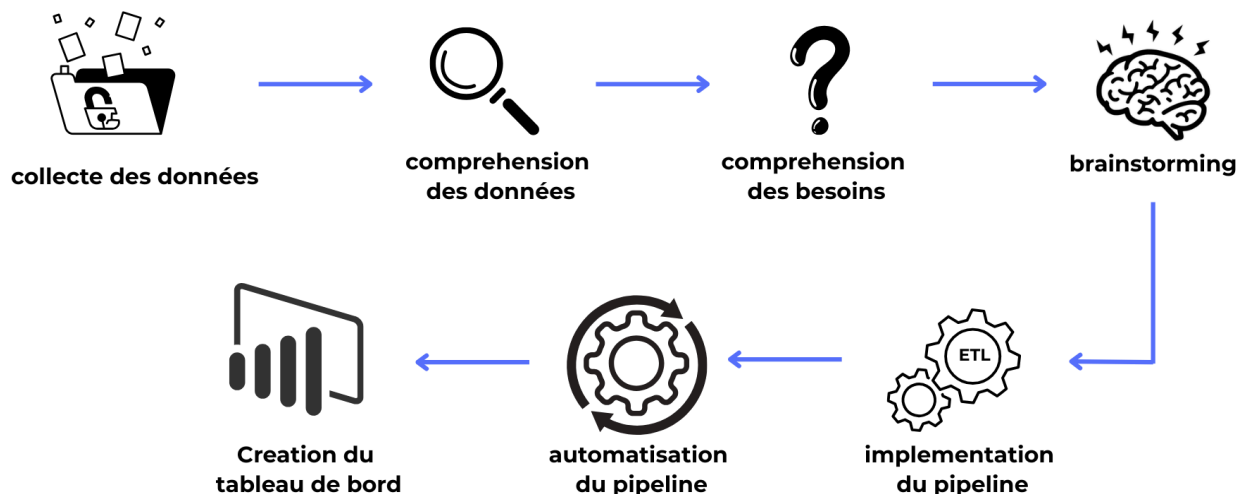


Figure 3 : Schéma des etapes du projet

Conception du projet

1- Collecte des données nécessaires

J'ai reçu comme données 4 fichiers Excel chacun correspond à des ventes durant une année. (2021-2025)

2- Compréhension des données fournies

Afin de comprendre ainsi les données et leur signification métier, j'ai collecté ces informations en posant des questions sur le directeur SI:

- **Désignation** => Nom complet des produits (ex: BB Alum FM HO Bipackx27 Hacendado/ST)
- **Total HT** => Le chiffre d'affaire fractionné par chaque vente
- **Famille Niv 2** => Nom de marque
- **Etablissement** => Type de distribution : Locale ou export
- **Famille Niv 3** => Type de poisson : Maquereaux, thon ou sardines
- **Tiers** => Clients avec lesquels ils travaillent

3- Compréhension profonde des besoins

J'ai eu aussi un discours avec le directeur SI pour collecter les exigences et besoins d'une manière plus détaillée en vue de diriger le résultat final vers ce qu'ils attendent de mon intervention et voici ce que j'ai eu comme résultat de cette conversation :

- Besoin d'un tableau de bord récapitulatif du chiffre d'affaire entre 2021 et 2025
- Indiquer les désignations existantes
- Indiquer les types d'établissement
- Indiquer les tiers
- Indiquer les familles des produits

4- Conclusion des transformations à faire

Tout au long de ce processus de découverte des données, j'ai aussi essayer d'analyser et conclure quelles transformations et nettoyage à faire et voici les actions que j'ai conclu lors de cette étude préalable :

- Enlever les valeurs NULL et les doublons
- Enlever les colonnes qui n'illustrent pas les besoins exprimés (ex. Qtés, Montant Devise ...)
- Enlever les ventes qui contiennent "sampling, sapmling ou sapling" pour signifier échantillon) car elles ne sont pas prises en considération dans le chiffre d'affaire
- Enlever les ventes avec num. de référence invalid (après avoir compris d'eux que déjà elles sont que des ventes insignifiantes)
- Dans Famille Niv. 2, remplacer "vide" par la marque correspondante "Lukus"

5- Décision d'aborder SSIS pour implémenter un pipeline ETL

Après avoir compris la structure des données et la signification de chaque colonne et aussi après avoir eu une idée sur la manipulation à faire, j'ai décidé d'opter pour l'implémentation d'un pipeline ETL sur SSIS au lieu de nettoyer les données directement sur Power BI car SSIS:

- Offre un flux intuitif et facile à comprendre
- Est plus automatisé et rapide
- Permet la validation de chaque étape et une analyse d'erreur plus profonde
- A plus d'options de transformation que Power BI
- Permet la manipulation de tous les fichiers soit de source ou destination dans un seul canvas facilement sans avoir besoin de traiter chacun séparément
- Facilite la répétition des mêmes transformations pour chaque fichier (les copier en changeant juste les sources ou destinations)
- Possibilité d'automatisation des pipelines (surtout quand les données ne sont pas complètes [2025 s'est arrêté au mois mai])

6- Choix des visuels à aborder dans le tableau de bord

Comme déjà connu, chaque type de données à un visuel qui lui va le plus, ainsi que l'objectif aussi joué un rôle primordial dans le choix des graphiques (distribution, comparaison, proportion ...) et ce n'est pas un choix arbitraire ou selon ses préférences personnelles. C'est pour cette raison que j'ai pris du temps pour choisir les visuels pour un résultat plus professionnel :

- **Chiffre d'affaire au fil du temps** : graphe linéaire
- **Etablissement** : graphe circulaire
- **Désignation** : liste déroulante (dans un menu)
- **Tiers** : liste déroulante (dans un menu)
- **Types de poissons** : graphe donut
- **Chiffre d'affaire par année** : carte
- **Top 10 marque** : graphe en bâton horizontal
- **Chiffre d'affaire de chaque année** : graphe en bâton
- Ajout d'une ligne chronologique pour plus de contrôle

Conclusion





L'étape de conception permet d'avoir une idée générale sur ce que sont les actions à faire ainsi que les décisions à prendre. Certes elle peut sembler n'être qu'une étape ajoutée sans objectif, mais elle facilitera la réalisation par la suite, donc il faut prendre du temps pour réfléchir à tous les détails du projet pour éviter des va et vient durant l'exécution et pour avoir un flux plus fluide.

Chapitre III : Réalisation du projet

Introduction

Durant la réalisation du projet, j'ai essayé de suivre tout ce que j'ai conclu de la phase de conception, du pipeline ETL à réaliser jusqu'au choix des visuels et les modifications à effectuer dans le tableau de bord pour le rendre lisible.

Les outils abordés

Pour l'implémentation du pipeline ETL, j'ai installé SSIS sur Visual Studio	Pour créer la base de données destination du pipeline, et pour automatiser l'implémentation de ce dernier, j'ai installé SQL Server	j'ai abordé SSMS pour créer la base de données et tous les autres éléments nécessaires sur l'instance de SQL Server.	Afin de réaliser le tableau de bord, j'ai utilisé Microsoft Power BI pour son interface Low Code/No Code facile.
			 Power BI
Figure 4: Logo de visual studio	Figure 5: Logo de SQL Server	Figure 6 : Logo de SSMS	Figure 7 : Logo de Power BI

Réalisation du projet

1- Implémentation du Pipeline ETL sur SSIS

Voici une Figure qui illustre la forme du Pipeline ETL réalisé :

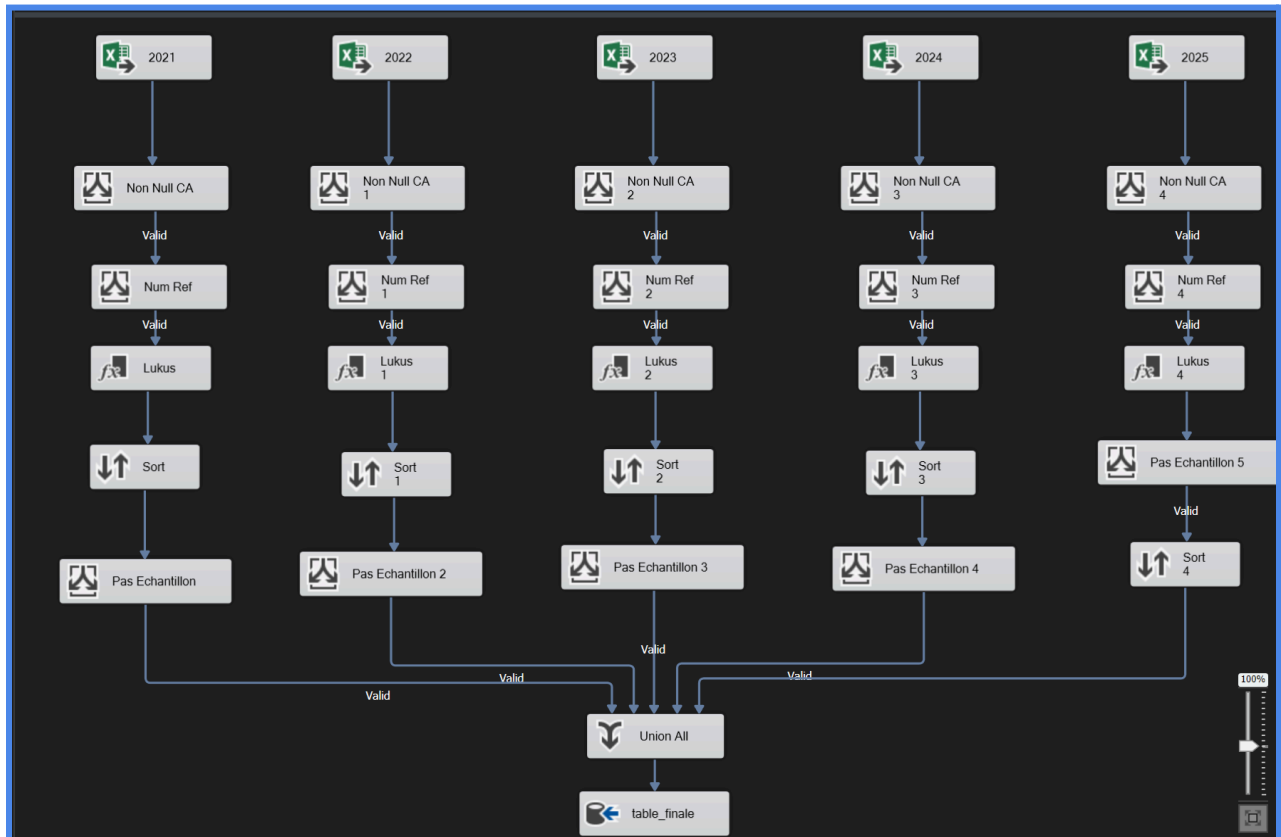


Figure 8 : Le pipeline ETL

- La source (Fichier Excel de l'année):
Pour l'intégrer, j'ai créé une connexion au chemin du fichier comme suit :

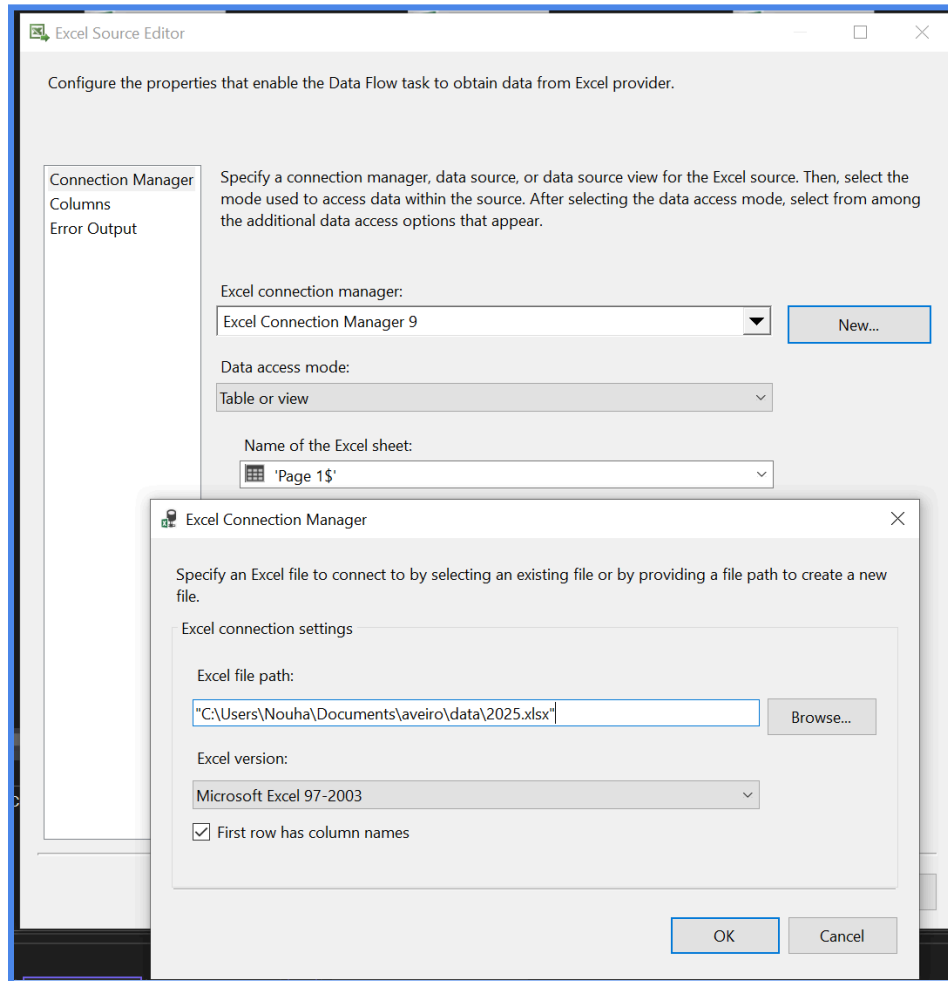


Figure 9 : Ajout des fichiers Excel de la destination

- Non NULL CA:

Dans cette étape, j'ai enlevé les lignes qui ont des ventes avec des valeurs nulles = 0

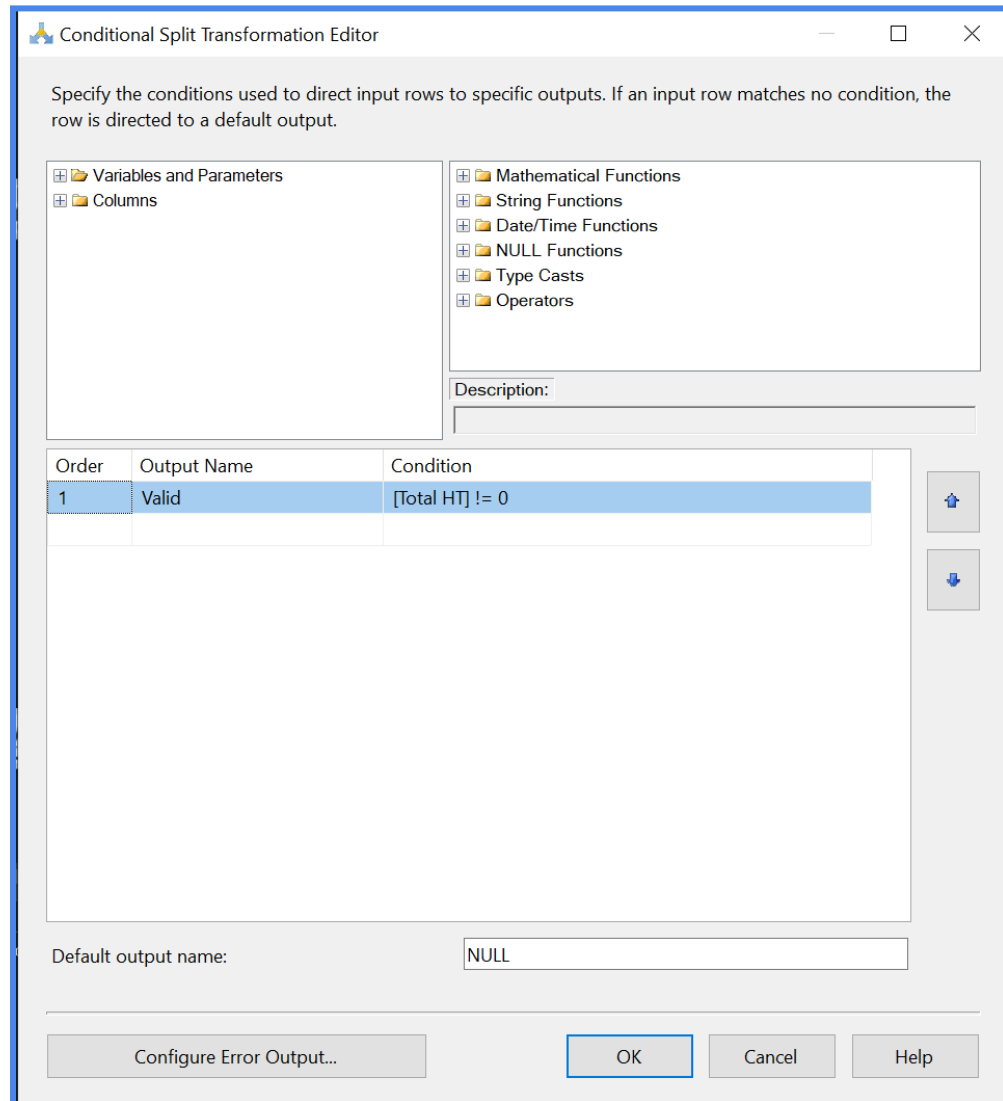


Figure 10 : Suppression des ventes nulles

Formule : [Total HT] != 0

- Num Ref:

Ici j'ai enlevé les lignes avec les num. de référence invalides (ventes insignifiantes)

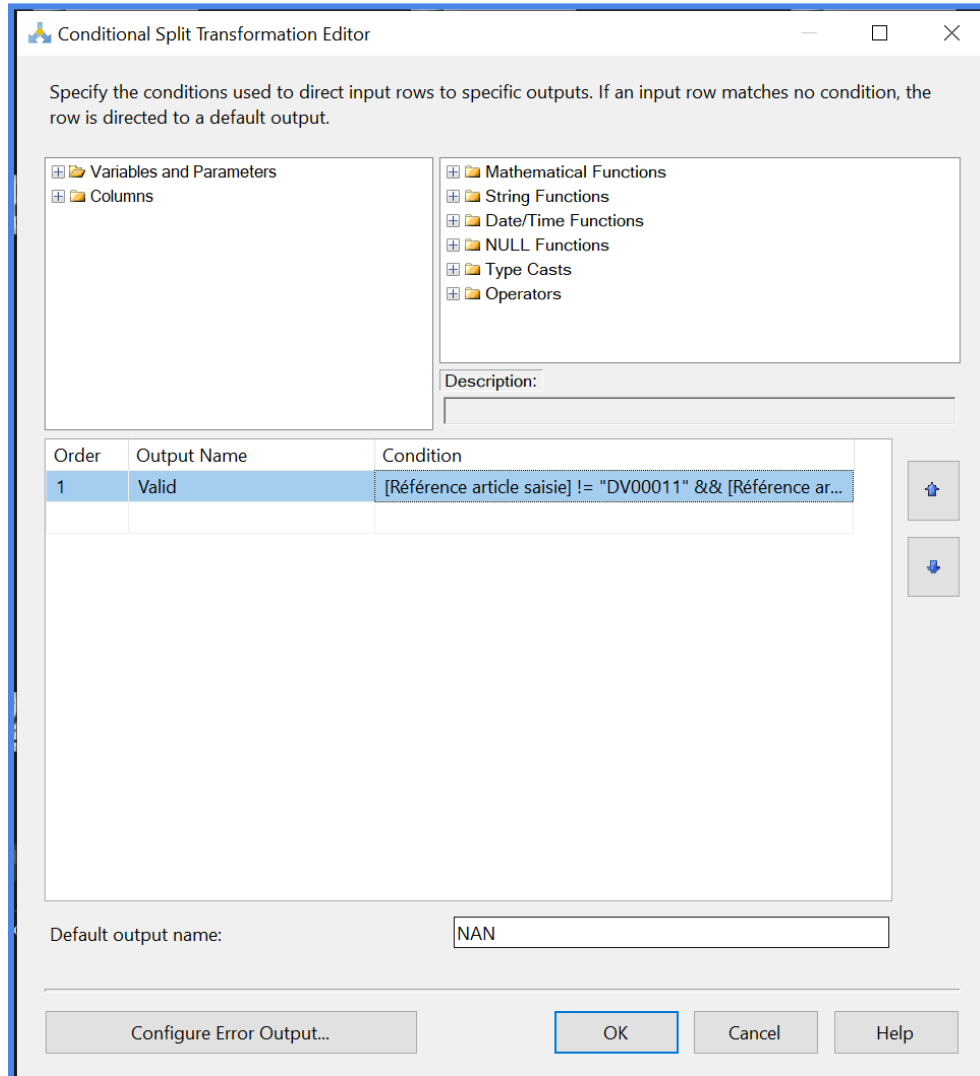


Figure 11 : Suppression des ventes invalides

Formule : [Référence article saisie] != "DV00011" && [Référence article saisie] != "DV00012" && [Référence article saisie] != "DV00019"

- **Lukus:**

Durant cette étape, j'ai remplacé chaque colonne avec le nom de marque vide par Lukus

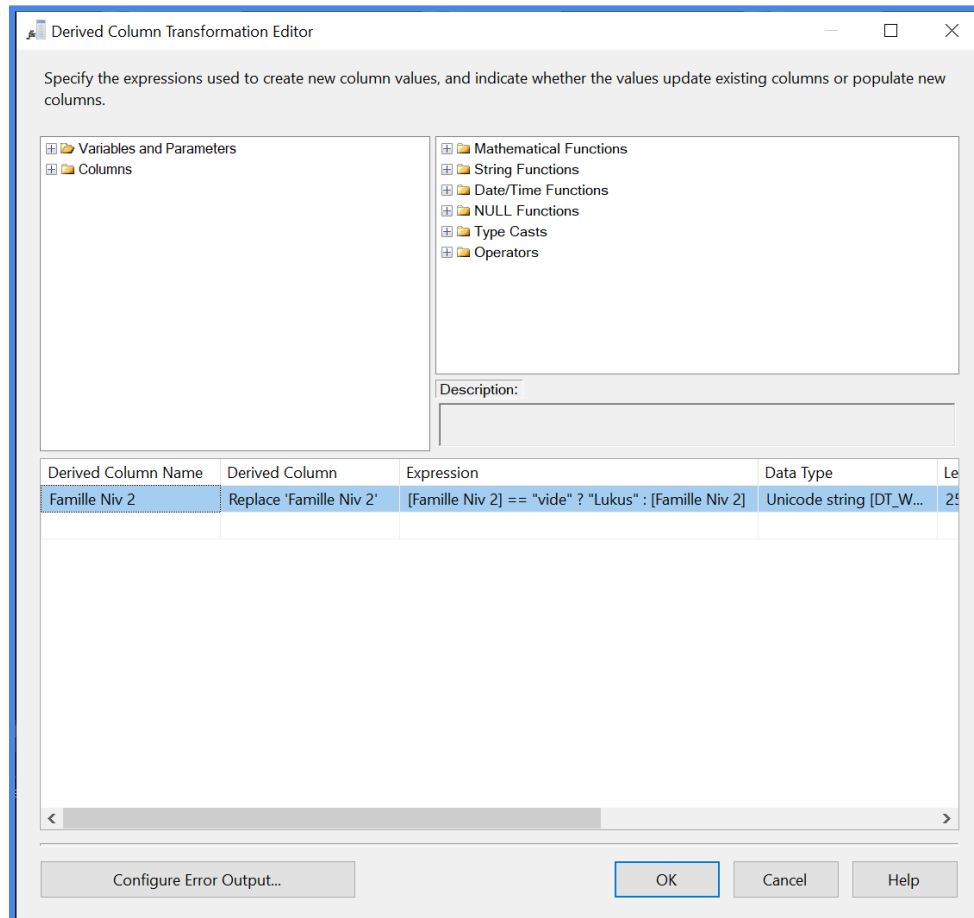


Figure 12 : Remplacement des marques vide par Lukus

Formule : [Famille Niv 2] == "vide" ? "Lukus" : [Famille Niv 2]

- Sort:

Dans cette transformation, j'ai gardé qu'une occurrence de chaque vente sans doublons

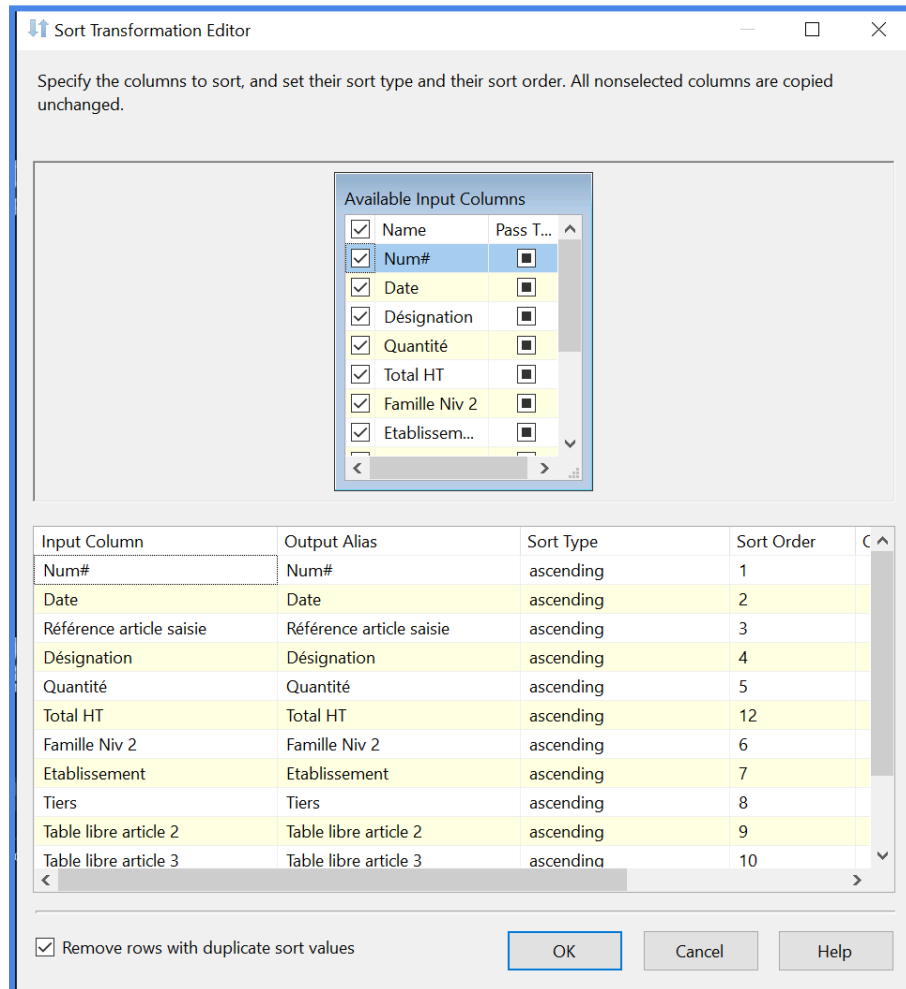


Figure 13 : Elimination des doublons

- Pas Echantillon:

Durant cette étape, j'ai enlevé les echantillons

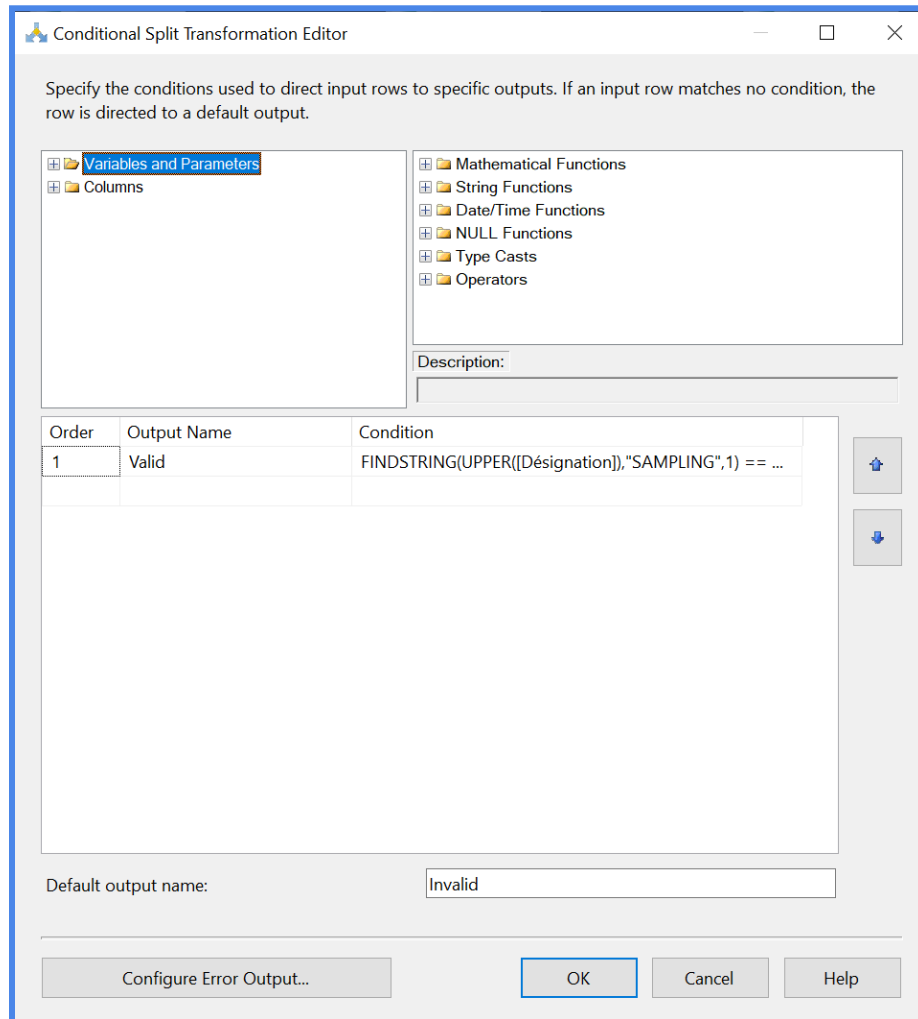


Figure 14 : Suppression des echantillons

Toutes ces étapes sont répétées pour chaque fichier Excel, puis j'ai effectué une union des données nettoyées et transformées vers une table SQL.

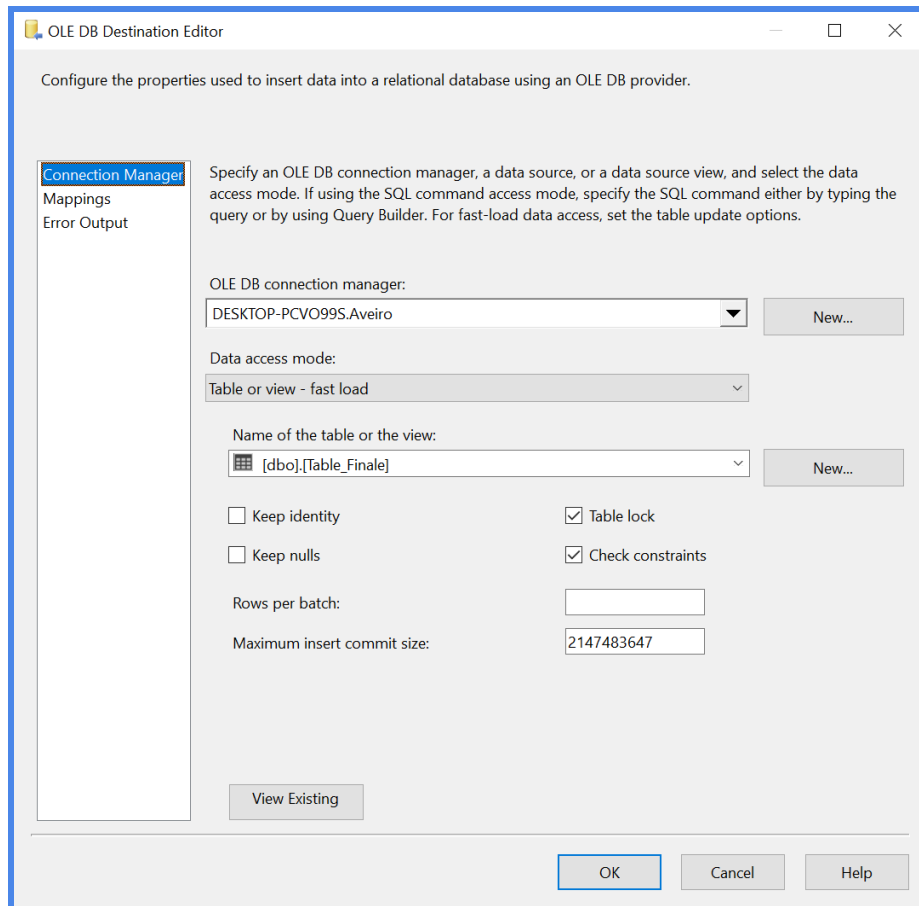


Figure 15 : Table destination avec toutes les données traitées

Cette table est par la suite créée sur SSMS pour recevoir les données ainsi que la base de donnée dont elle appartient

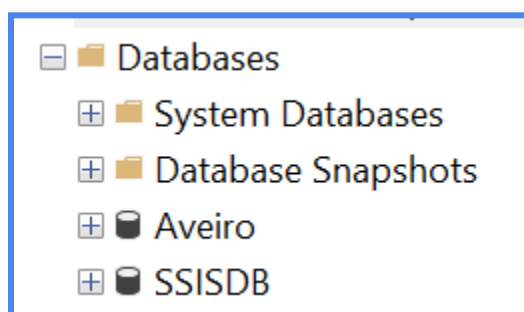


Figure 16 : Base de données et Table finale sur SSIS

Mais afin d'éviter que les données se concatènent à chaque fois le package ETL se lance, j'ai ajouté une tâche SQL avant Data Flow sur SSIS pour retirer et recréer la table pour avoir toujours une table mise à jour.

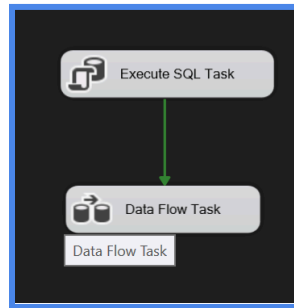


Figure 17 : la forme du package ETL

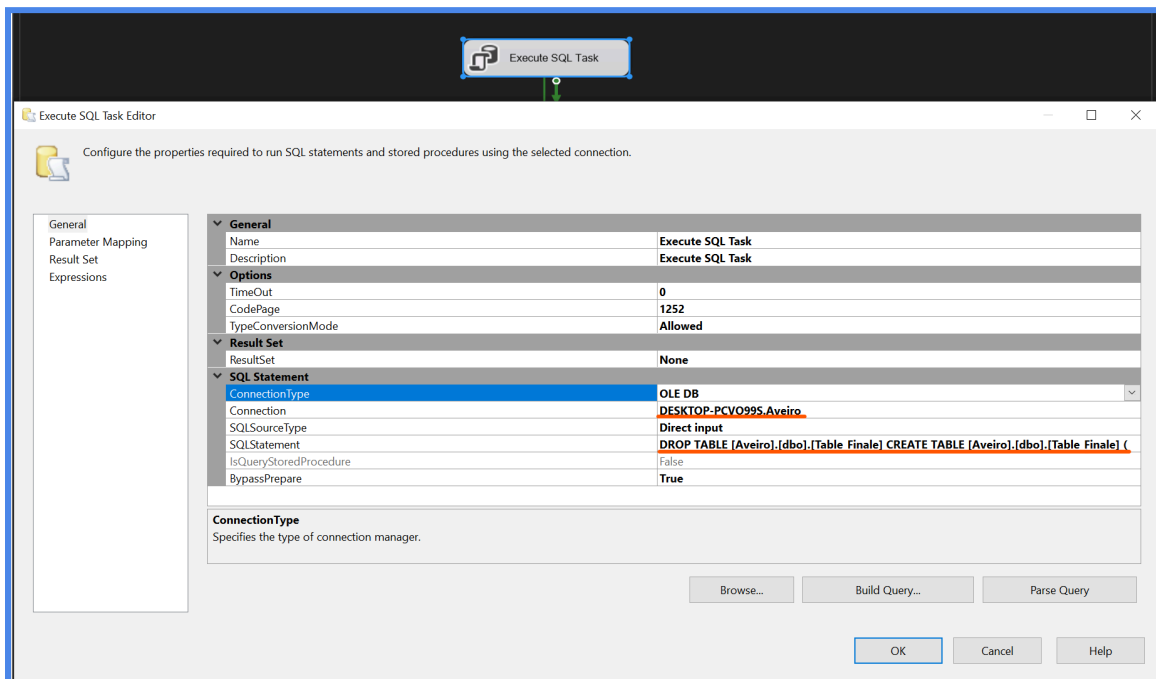


Figure 18 : La tâche Execute SQL

2- Déploiement du package ETL

Après avoir réalisé le pipeline, je l'ai déployé dans SSIS Catalog pour créer un job qui automatise son exécution avec une fréquence ajustable.

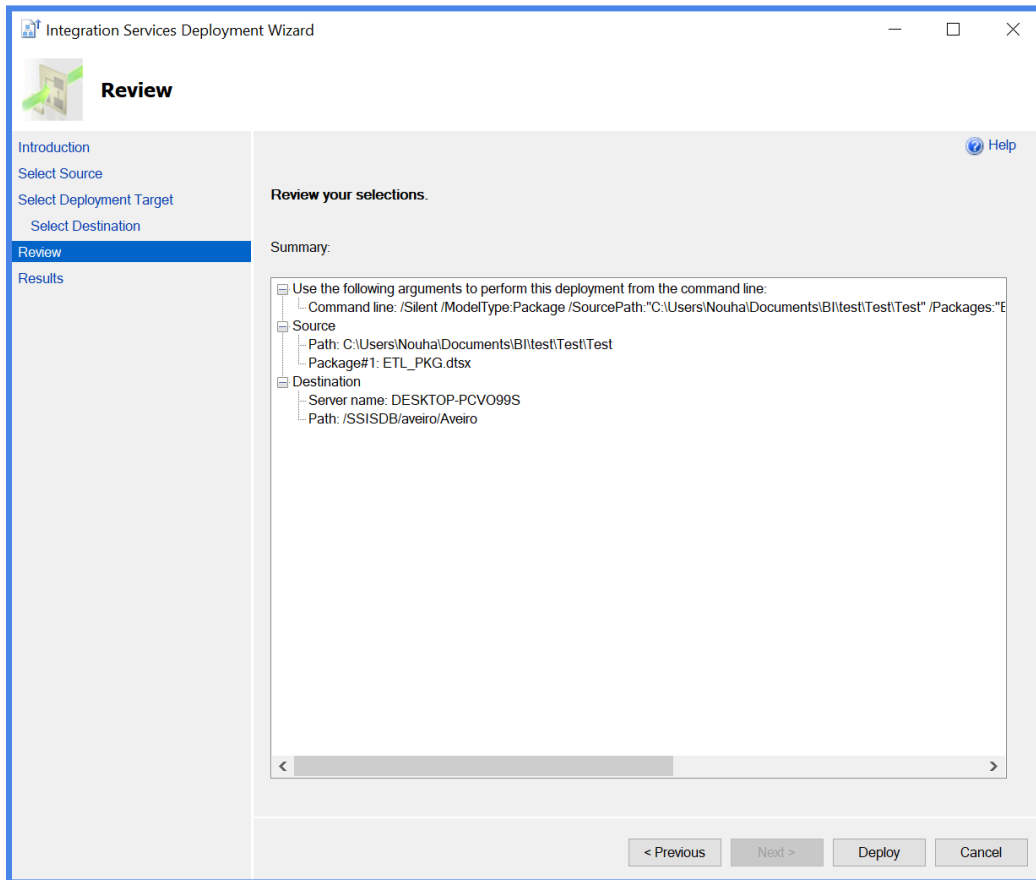


Figure 19 : Déploiement du package

Si on exécute le package ETL sur SSMS, la table (Table_Finale) se remplit par les données traitées

3- Création d'un Job pour automatiser l'exécution du pipeline

Après avoir déployé le pipeline sur SSIS Catalog, j'ai créé un job pour automatiser son exécution avec une fréquence ajustable selon l'activité de l'entreprise :

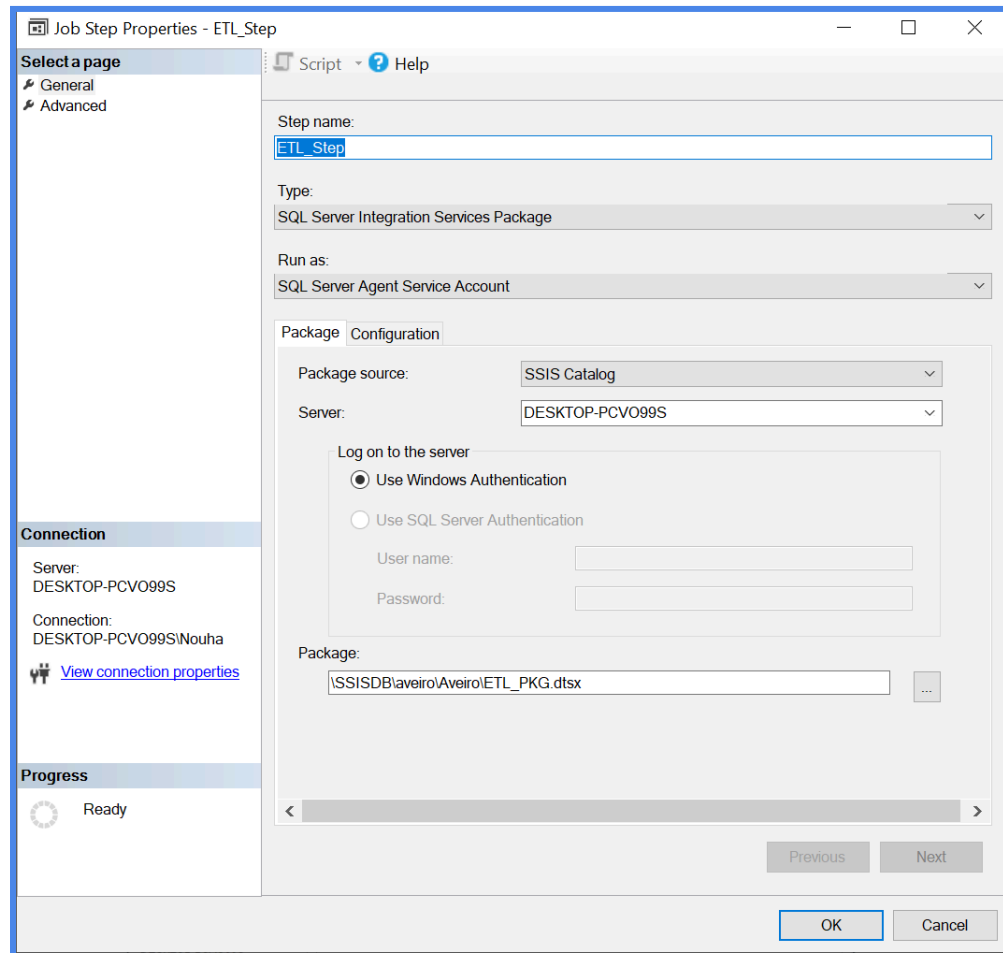


Figure 20 : Création du job sur SSMS

Figure 21 : Création d'un calendrier pour le job

4- Connexion de Power BI avec la base de données Aveiro

Pour créer le tableau de bord contenant les KPI liés au chiffre d'affaire de l'entreprise, je l'ai lié avec la base de données pour tirer les données actualisées de la table (Table_Finale)

Figure 22 : Création de connexion

5- Création du tableau de bord

Après avoir placé les visuels choisis dans la phase conception, en appliquant les filtres nécessaires (ex : Top 10 clients avec le filtre Top N, ...), tout en créant les mesures calculées nécessaires :

```
1 % CA_ParAnnee =  
2 DIVIDE(  
3     SUM(Table_Finale[Total HT]),  
4     CALCULATE(  
5         SUM(Table_Finale[Total HT]),  
6         ALL('Table_Finale'[Date].[Year]))  
7 )  
8
```

Figure 24 : La mesure CA. par année

Cette mesure a pour but la création d'une carte ou en visualise le pourcentage que chaque année préoccupe des données fournies, il sert ainsi à comparer généralement entre la performance des années.

J'ai créé aussi une table mesure pour organiser le projet BI sur Power BI :

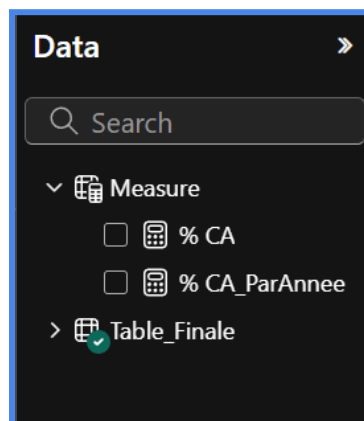


Figure 25: La table mesure pour arranger les mesures calculées

On obtient par la suite un tableau fonctionnel mais qui manque plusieurs modifications pour le rendre facile à exploiter :

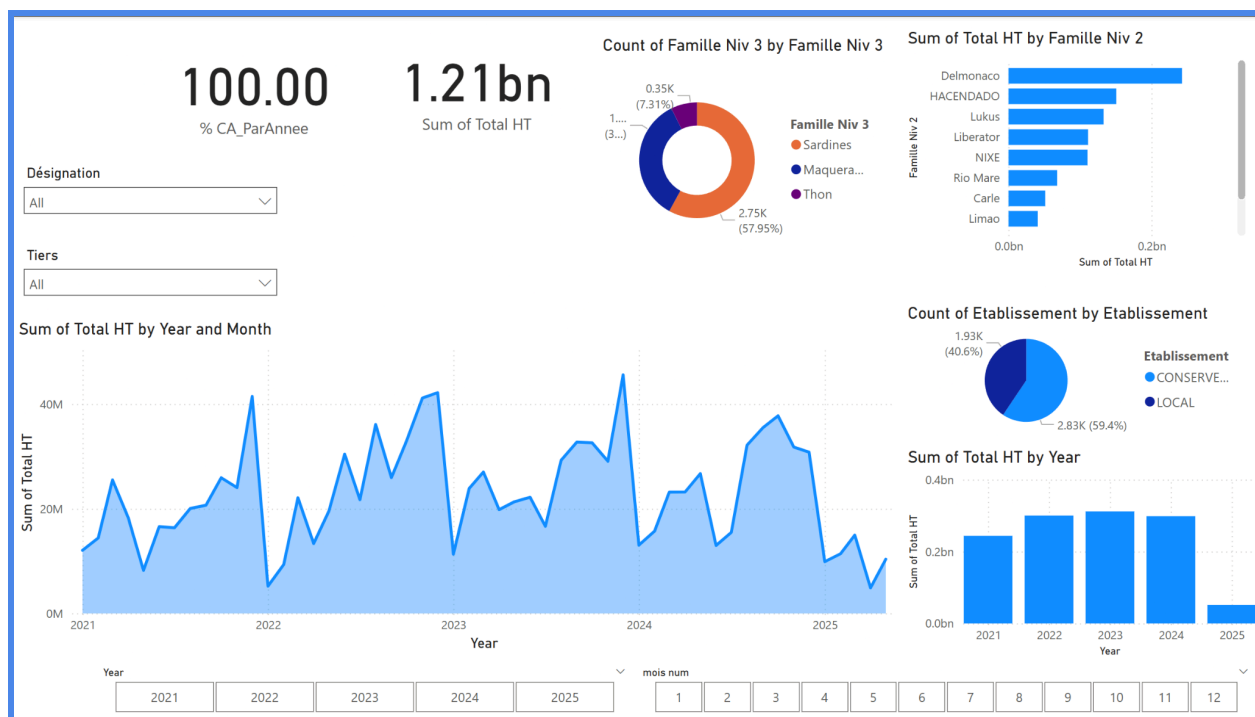


Figure 26: Tableau de bord initial

Voici la liste des modifications que j'ai effectué pour avoir une expérience utilisateur plus améliorée :

- Avoir une section titre et logo pour comprendre le contexte
- Changer l'arrière plan pour un contrast aisé
- Changer les titres des graphiques
- Avoir un gradient de couleur dans le graphe Top 10 marques pour donner ses valeurs plus de dimension
- Simplifier les visuels en enlevant des labels d'axes et les remplacer par des labels de données
- Diviser le graphe C.A. pour chaque année en deux secteurs (local et conserverie) pour comprendre plus la source
- Créer un menu pour les listes déroulantes selon le besoin

- Ajouter une ligne sous les cartes pour visualiser plus leur proportion, cette visualisation est basée sur un graphe empilé horizontal avec les deux mesures %CA_ParAnnee précédente et %CA pour représenter la partie vide :

$$1 \% CA = 1 - [\% CA_ParAnnee]$$

Figure 27: Mesure pour la ligne de proportion

- Ajouter les unités (% et MAD pour les ventes)
- Des icons pour plus de signification

Voici le tableau de bord final après la liste des modifications effectuées :

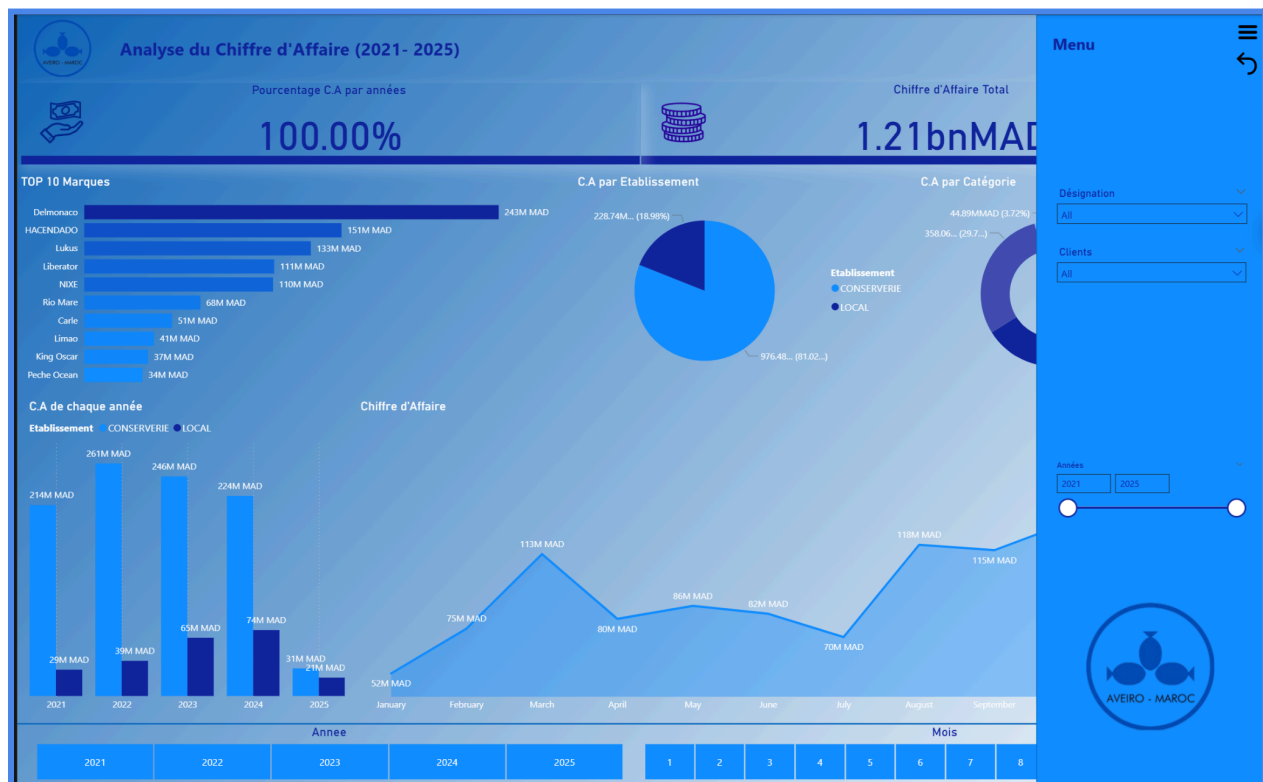


Figure 28: Tableau de bord final avec menu

Pour avoir un tableau de bord actualisé, il faut rafraichir dans Power BI pour reprendre les données récentes de la connexion créée

Résultats

Dans la fin de chaque projet BI, le développeur doit avoir l'aptitude d'exploiter les données finales dans le tableau de bord et de savoir comment les interpréter et non pas juste s'arrêter dans la visualisation, afin de transmettre le message aux experts métier et de montrer qu'il a bien comprendre leurs besoins.

Dans ce tableau de bord, on observe :

- 2022 a le plus de ventes internationales 261 M MAD
- 2024 a le plus de ventes locales 74 M MAD
- Delmonaco est la 1^{ere} marques vendue 243 M MAD
- Aveiro a plus de profit à l'international qu'au local (81% par rapport a 19%)
- Chiffre d'affaire jusqu'à la période saisie de 2025 (mai) est 1.21 Milliard DH
- Les sardines préoccupe la plus grandes proportion du chiffre d'affaire (66.1%), puis les maquereaux (~30%) puis les thon (~4%)

Livrables du projet

En raison de sécurité, j'ai pas eu l'accès au serveur de l'entreprise pour configurer tout ce processus directement là-bas, c'est pour cette raison que j'ai tout testé sur mon ordinateur, et après avoir assuré son succès, j'ai livré au directeur SI :

- **Le project du package ETL**
- **Le fichier .pbix du tableau de bord sur Power BI**
- **Un guide de configuration et d'installation des outils nécessaires**

Choses à améliorer

Durant cette expérience, j'ai constaté qu'il y a une marge d'amélioration qu'on peut effectuer pour aboutir à un résultat plus affiné et un workflow plus fluide :

- Avoir un système de saisie des données avec des listes déroulantes au lieu de la saisie manuelle pour éviter les fautes de saisie ou d'orthographe qui rendent le processus de nettoyage plus épuisant où on risque de commettre plus d'erreurs
- Avoir une licence Power BI pour un rafraichissement automatique des tableaux de bord

Conclusion

Certes cette phase est la plus nécessaires dans le projet car c'est dans laquelle on met toute conception en réalité tangible pour livrer un projet de qualité, mais elle ne doit pas se limiter à l'exécution, elle doit toujours se baser sur une conversation continue avec les parties prenantes.

Conclusion générale

Ce projet m'a permis d'avoir une expérience réelle avec un besoin réel aussi où j'ai pu appliquer ce que j'ai appris avant, mais aussi d'apprendre plus, de savoir comment analyser les besoins métiers, de faire une conception détaillée et de réaliser un projet BI d'une manière optimisée. Certes on apprend la plupart des aspects théoriques dans notre cursus académique, mais avoir une expérience professionnelle au sein d'un organisme nous permet d'apprendre comment interagir avec les autres cadres, et pour extraire autant d'informations que possible pour réussir son projet professionnel.

Webographie

Documentation

Microsoft Community : <https://learn.microsoft.com/en-us/>

Power BI Community : <https://community.fabric.microsoft.com/>

Stack Overflow : <https://stackoverflow.com/>

Aveiro Maroc : <http://aveiro-maroc.ma/contacts.html>

Installation

Visual Studio : <https://visualstudio.microsoft.com/downloads/>

SQLServer: <https://www.microsoft.com/en-us/sql-server/sql-server-downloads>

Power BI: <https://www.microsoft.com/en-us/download/details.aspx?id=58494>

SSMS: <https://learn.microsoft.com/en-us/ssms/release-notes-20>