



كلية العلوم القانونية والاقتصادية
والاجتماعية - طنجة
FSJES TANGER

Royaume du Maroc

UNIVERSITÉ ABDELMALEK ESSAÂDI

Faculté des Sciences Juridiques, Economiques et Sociales

-Tanger



Master : Intelligence Artificielle pour l'Economie Numérique et la Gestion

Analyse de l'impact des campagnes marketing sur les ventes en ligne



Encadré par :

Pr. BENNANI ANAS

Réaliser par :

Eyoub Nidal

2025 - 2026

Résumé

Ce projet s'inscrit dans le cadre du Master Spécialisé en **Intelligence Artificielle pour l'Économie Numérique et la Gestion**. Il porte sur *l'analyse de l'impact des campagnes marketing sur les ventes en ligne*, un sujet à forte valeur ajoutée pour les entreprises qui cherchent à optimiser leur stratégie digitale.

L'objectif principal de ce travail est de concevoir une chaîne complète de traitement et d'analyse des données marketing, depuis leur extraction brute jusqu'à la production de tableaux de bord interactifs. Pour cela, nous avons mobilisé plusieurs outils complémentaires:

- **SSIS (SQL Server Integration Services)** pour l'extraction, la transformation et le chargement (ETL) des données issues d'une base brute vers un modèle en étoile,
- **SSMS (SQL Server Management Studio)** pour la gestion des tables de faits et de dimensions, ainsi que pour la vérification par requêtes SQL,
- **Power BI** pour la modélisation et la visualisation des données, permettant la création d'indicateurs clés de performance (KPIs) tels que le *Conversion Rate*, le *Coût d'Acquisition* et l'*Engagement Score*.

Les résultats obtenus démontrent l'importance d'une bonne gouvernance des données et d'une visualisation efficace pour comprendre l'impact réel des campagnes marketing. Le modèle mis en place offre une vue claire sur la performance des différents canaux, permettant aux décideurs de mieux allouer leurs ressources et d'orienter leurs actions stratégiques.

Au-delà de l'aspect technique, ce projet illustre la valeur ajoutée de l'**intelligence artificielle et de l'analytique avancée** dans le pilotage de l'économie numérique. Il constitue une première étape vers l'intégration d'outils d'IA plus évolués (prédiction, segmentation, recommandation) pour renforcer la prise de décision et accroître la compétitivité des entreprises dans un marché digital en constante évolution.

Table des matières

Résumé	2
Introduction	4
Contexte et Justification du Sujet.....	4
Problématique et Objectifs de l'Étude	4
Méthodologie et Outils Choisis	4
Description des données	6
Source des données	6
Conception du Modèle de Données.....	14
Description des Tables de Dimensions.....	14
Description de la Table de Faits	16
Relation entre les Tables de Dimensions et la Table de Faits	18
Création du Projet SSIS et Test de Connexion	19
Vérification du Chargement des Données	27
Requête SQL pour chaque Table de Dimension.....	27
Requête SQL pour la Table de Faits.....	30
Requêtes de Validation Générales pour Chaque Table.....	31
Intégration et visualisation dans Power BI	32
Transformations de Données	34
Power Query — nettoyage :	34
Supprimer les doublons sur les clés (si présents).	36
Renommer les colonnes pour cohérence (ex. LocationName).	36
Power Query — transformations métier :	37
Mesures et KPIs calculés.....	39
Analyse et visualisations Power BI.....	40
Visualisations Finales du Dashboard Power BI.....	52
Conclusion Générale.....	54

Introduction

Contexte et Justification du Sujet

Les campagnes marketing en ligne jouent un rôle crucial dans l'optimisation des performances commerciales des entreprises, notamment dans le secteur de l'e-commerce. Dans un environnement numérique en constante évolution, les entreprises cherchent constamment à maximiser le retour sur investissement (ROI) de leurs actions marketing. Cependant, la véritable efficacité de ces campagnes n'est pas toujours évidente sans une analyse approfondie des données générées.

L'impact des campagnes marketing sur les ventes est une question centrale pour les départements marketing et financier des entreprises. Les décideurs ont besoin de preuves concrètes sur l'efficacité de chaque canal de communication (emailing, réseaux sociaux, publicités payantes, etc.). Ce projet se concentre précisément sur l'analyse de cet impact en utilisant des outils modernes d'analyse de données.

Problématique et Objectifs de l'Étude

Ce projet vise à analyser l'impact des différentes campagnes marketing sur les ventes en ligne. L'objectif principal est de répondre aux questions suivantes :

1. Quelle est la performance globale de chaque compagnie ?
2. Quelles sont les compagnies les plus rentables ?
3. Quelle est la compagnie avec le coût d'acquisition le plus faible ?
4. Quel segment de clients génère le meilleur ROI ?
5. Quelle audience réagit le mieux aux campagnes marketing ?
6. Quelle localisation apporte le plus de conversions ?
7. Quelle est la tendance des conversions dans le temps (mois/année) ?
8. Est-ce que le ROI est en amélioration ou en baisse au fil des mois ?
9. Quels sont les meilleurs mois et jours de la semaine pour les performances marketing ?
10. Quel est le coût moyen par conversion et que révèle-t-il sur l'efficacité des campagnes ?
11. Quel canal marketing génère le meilleur retour sur investissement (ROI) ?

Pour ce faire, nous utiliserons des données extraites du Marketing Campaign Performance Dataset sur Kaggle, que nous traiterons et analyserons à l'aide de SSIS pour l'ETL (Extract, Transform, Load) et de Power BI pour la visualisation des résultats.

Méthodologie et Outils Choisis

Dans le cadre de cette étude, nous avons opté pour une approche basée sur l'analyse des données avec les outils suivants :

- SSIS (SQL Server Integration Services) : Utilisé pour l'extraction, la transformation et le nettoyage des données. Nous avons conçu un flux ETL permettant d'extraire les données brutes de Kaggle, de les transformer pour les rendre exploitables et de les charger dans une base de données SQL Server.
- Power BI : Utilisé pour la visualisation des données. Power BI nous permettra de créer des tableaux de bord interactifs afin de mieux comprendre les tendances des campagnes marketing et de générer des rapports détaillés sur le ROI.

Description des données

Source des données

Les données utilisées dans ce projet proviennent du dataset intitulé **Marketing Campaign Performance Dataset**, disponible sur la plateforme **Kaggle**.

<https://www.kaggle.com/datasets/manishabhatt22/marketing-campaign-performance-dataset>

The screenshot shows the Kaggle interface for the 'Marketing Campaign Performance Dataset'. The left sidebar includes links for Home, Competitions, Datasets (selected), Models, Benchmarks, Code, Discussions, Learn, and More. Under 'Your Work', there are links for VIEWED (Marketing Campaign ...), EDITED (MAGIC Gamma Teles...), and E-Commerce Data. Basic Data Exploration is also listed under EDITED. The main content area displays the dataset's summary, showing it contains 27.22 MB of data across 16 columns. A preview table shows data for various cities, languages, and campaign details. The right sidebar shows the 'Data Explorer' with 1 file and 16 columns.

Location	Language	Clicks	Impressions	Engagement	Customer Segm.	Date
Chicago	Spanish	506	1922	6	Health & Wellness	2021-01-01
New York	German	116	7523	7	Fashionistas	2021-01-02
Los Angeles	French	584	7698	1	Outdoor Adventurers	2021-01-03
Miami	Mandarin	217	1820	7	Health & Wellness	2021-01-04
Los Angeles	Mandarin	379	4281	3	Health & Wellness	2021-01-05
New York	German	108	1643	1	Foodies	2021-01-06
Los Angeles	Spanish	817	8749	18	Tech Enthusiasts	2021-01-07
Los Angeles	Mandarin	624	7854	7	Outdoor Adventurers	2021-01-08
Chicago	German	861	1754	6	Tech Enthusiasts	2021-01-09
Los Angeles	English	642	3856	3	Tech Enthusiasts	2021-01-10
New York	Spanish	321	6628	10	Tech Enthusiasts	2021-01-11
Los Angeles	Mandarin	659	8948	1	Foodies	2021-01-12

Ce jeu de données contient **200 000 lignes uniques** couvrant une période de **deux ans**, et regroupe des campagnes issues de plusieurs entreprises fictives dans différents secteurs.

Volume des données

- **Nombre de lignes :** 200 000
- **Nombre de colonnes :** 14 principales variables
- **Période couverte :** 2 ans

Variables principales

Le tableau ci-dessous présente les colonnes principales du dataset et leur signification :

Colonne	Description
Company	Nom de l'entreprise fictive responsable de la campagne
Campaign_Type	Type de campagne (email, réseaux sociaux, influenceur, display, search, etc.)
Target_Audience	Segment de clientèle visé (ex. : femmes 25-34, hommes 18-24, tous âges)
Duration	Durée de la campagne (en jours)
Channels_Used	Canaux utilisés (email, YouTube, réseaux sociaux, Google Ads, etc.)
Conversion_Rate	Pourcentage de prospects ayant effectué l'action souhaitée (ventes, inscriptions...)
Acquisition_Cost	Coût moyen pour acquérir un client
ROI	Retour sur investissement (mesure de rentabilité)
Location	Ville ou région où la campagne a été menée (New York, Los Angeles, etc.)
Language	Langue utilisée dans la communication (anglais, espagnol, français, etc.)
Clicks	Nombre de clics générés par la campagne
Impressions	Nombre total d'affichages de la campagne
Engagement_Score	Score de 1 à 10 représentant le niveau d'engagement du public
Customer_Segment	Catégorie de clients ciblée (mode, tech, santé, foodies, etc.)
Date	Date de la campagne (permet l'analyse chronologique des tendances)

Intérêt analytique

Ces données sont particulièrement adaptées pour :

- analyser la performance des différentes campagnes marketing,
- calculer des **KPIs essentiels** (Conversion Rate, ROI, Acquisition Cost, Engagement Score),
- comparer l'efficacité des canaux (réseaux sociaux, email, etc.),
- identifier les segments de clientèle les plus réactifs,
- et comprendre l'impact global des stratégies marketing sur les ventes en ligne.

Chargement initial dans Excel

Afin de vérifier la structure et la qualité des données, le fichier CSV a d'abord été ouvert dans **Microsoft Excel**.

- Cette étape permet de visualiser les colonnes, détecter d'éventuelles incohérences (valeurs manquantes, doublons, etc.), et confirmer que les données correspondent bien à la description fournie par Kaggle.
- Dans Excel, on peut déjà effectuer un premier aperçu du volume et du format des données.

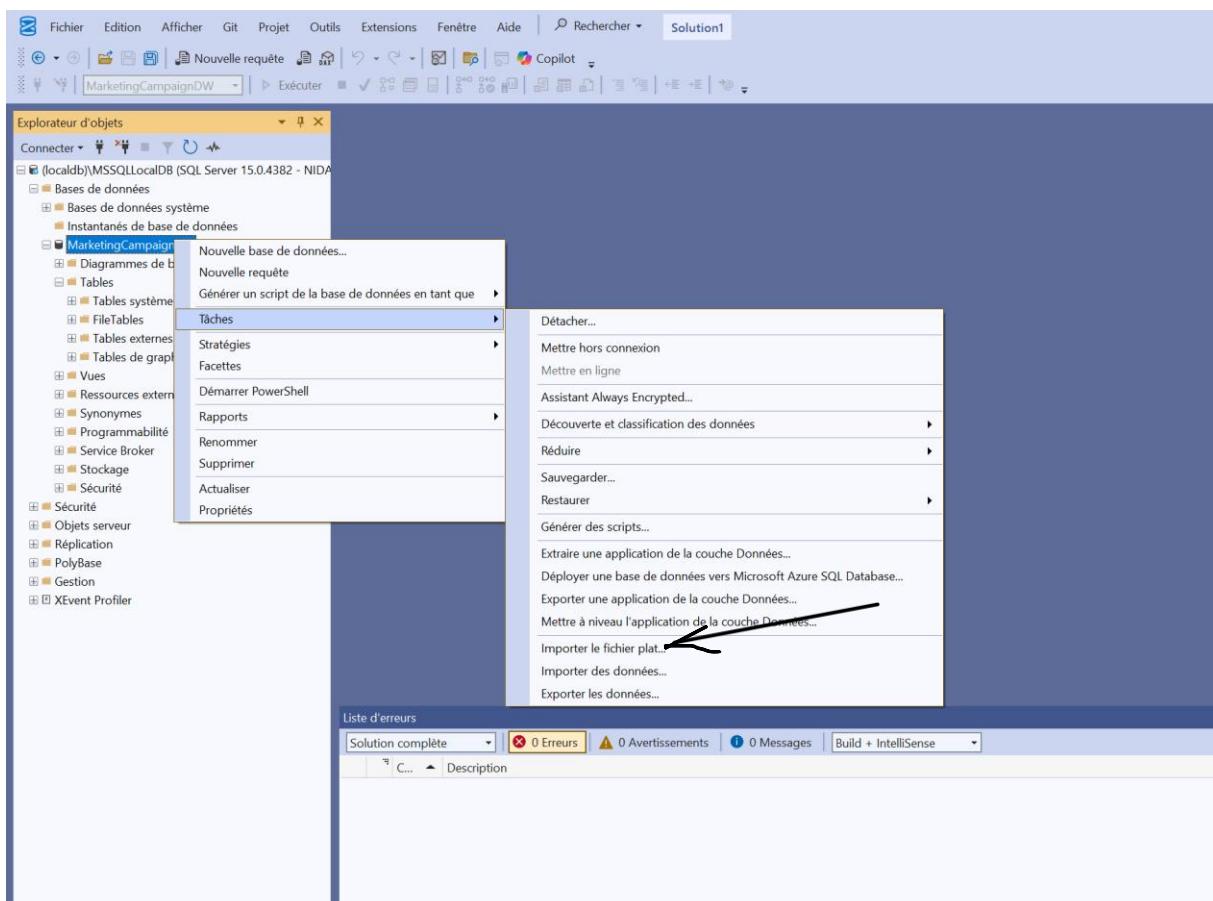
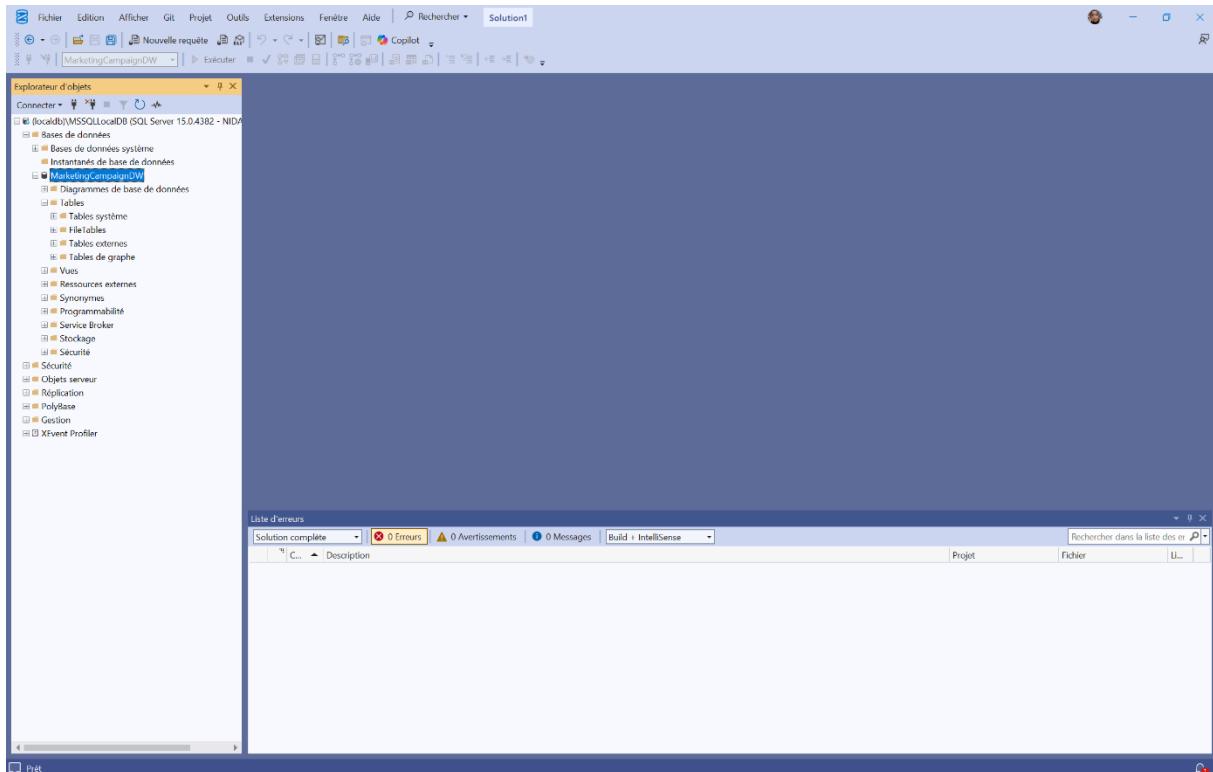
Campaign_ID	Company	Campaign_Type	Target_Audience	Duration	Channel_Used	Conversion_Rate	Acquisition_Cost	ROI	Location	Language	Clicks	Impressions	Engagement	Customer_Sat	Date
1	Innovate Industries	Email	Men 18-24	30 days	Google Ads	0.04	\$16,174.00	6.29	Chicago	Spanish	506	1922	6	Health & We	1/1/2021
2	NexGen Systems	Email	Women 35-44	60 days	Google Ads	0.12	\$11,566.00	5.61	New York	German	116	7523	7	Fashionistas	1/2/2021
3	Alpha Innovations	Influencer	Men 25-34	30 days	YouTube	0.07	\$10,200.00	7.18	Los Angeles	French	584	7698	1	Outdoor Adv	1/3/2021
4	DataTech Solutions	Display	All Ages	60 days	YouTube	0.11	\$12,724.00	5.55	Miami	Mandarin	217	1820	7	Health & We	1/4/2021
5	NexGen Systems	Email	Men 25-34	15 days	YouTube	0.05	\$16,452.00	6.5	Los Angeles	Mandarin	379	4201	3	Health & We	1/5/2021
6	DataTech Solutions	Display	All Ages	15 days	Instagram	0.07	\$9,716.00	4.36	New York	German	100	1643	1	Foodies	1/6/2021
7	NexGen Systems	Email	Women 35-44	60 days	Website	0.13	\$11,067.00	2.86	Los Angeles	Spanish	817	8749	10	Tech Enthusi	1/7/2021
8	DataTech Solutions	Search	Men 18-24	45 days	Google Ads	0.06	\$13,280.00	5.55	Los Angeles	Mandarin	624	7854	7	Outdoor Adv	1/8/2021
9	Alpha Innovations	Social Media	Women 35-44	15 days	Facebook	0.09	\$18,066.00	6.73	Chicago	German	861	1754	6	Tech Enthusi	1/9/2021
10	TechCorp	Email	Women 35-44	15 days	Instagram	0.09	\$13,766.00	3.78	Los Angeles	English	642	3856	3	Tech Enthusi	1/10/2021
11	NexGen Systems	Display	Men 25-34	45 days	Email	0.12	\$8,590.00	3.49	New York	Spanish	321	6628	10	Tech Enthusi	1/11/2021
12	Innovate Industries	Influencer	Men 25-34	60 days	Google Ads	0.05	\$17,502.00	3.59	Los Angeles	Mandarin	659	8948	1	Foodies	1/12/2021
13	TechCorp	Social Media	Men 25-34	60 days	Facebook	0.09	\$17,189.00	4.91	Chicago	Mandarin	677	8817	10	Tech Enthusi	1/13/2021
14	TechCorp	Email	Men 25-34	45 days	Instagram	0.14	\$9,975.00	7.06	New York	German	994	2201	4	Health & We	1/14/2021
15	TechCorp	Display	All Ages	45 days	Website	0.06	\$11,346.00	5.28	Chicago	Spanish	482	8470	1	Outdoor Adv	1/15/2021
16	Innovate Industries	Social Media	Women 35-44	60 days	YouTube	0.11	\$9,407.00	2.91	New York	German	299	1512	5	Health & We	1/16/2021
17	Innovate Industries	Display	Women 35-44	45 days	Website	0.08	\$5,478.00	4.53	Houston	English	931	2488	3	Fashionistas	1/17/2021
18	Alpha Innovations	Influencer	Men 18-24	15 days	Instagram	0.14	\$9,485.00	4.48	Miami	French	218	9264	9	Health & We	1/18/2021
19	Alpha Innovations	Social Media	Men 25-34	60 days	Google Ads	0.07	\$19,224.00	6.08	New York	French	182	5798	1	Foodies	1/19/2021
20	DataTech Solutions	Influencer	Men 25-34	15 days	Google Ads	0.09	\$10,258.00	3.83	Miami	French	193	3677	1	Tech Enthusi	1/20/2021
21	DataTech Solutions	Search	Women 25-34	15 days	Email	0.04	\$16,580.00	7.99	New York	Spanish	975	1561	3	Outdoor Adv	1/21/2021
22	TechCorp	Influencer	All Ages	30 days	Facebook	0.02	\$12,824.00	6.21	New York	French	319	8586	2	Foodies	1/22/2021
23	NexGen Systems	Social Media	Women 35-44	15 days	YouTube	0.13	\$8,699.00	2.12	Miami	French	646	3841	5	Health & We	1/23/2021
24	Innovate Industries	Email	Women 25-34	15 days	Facebook	0.04	\$17,608.00	7.31	Miami	Spanish	764	2157	8	Health & We	1/24/2021
25	NexGen Systems	Social Media	Men 18-24	60 days	YouTube	0.07	\$8,773.00	6.03	Miami	Mandarin	527	1733	1	Fashionistas	1/25/2021

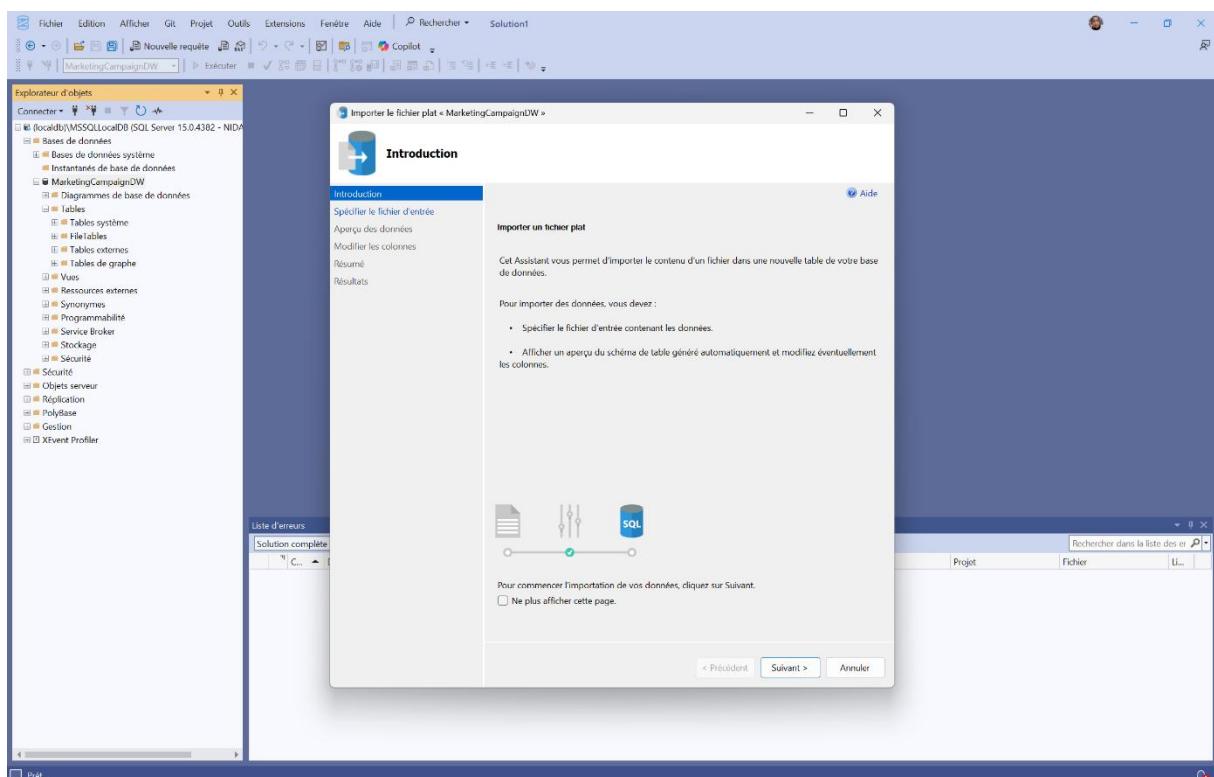
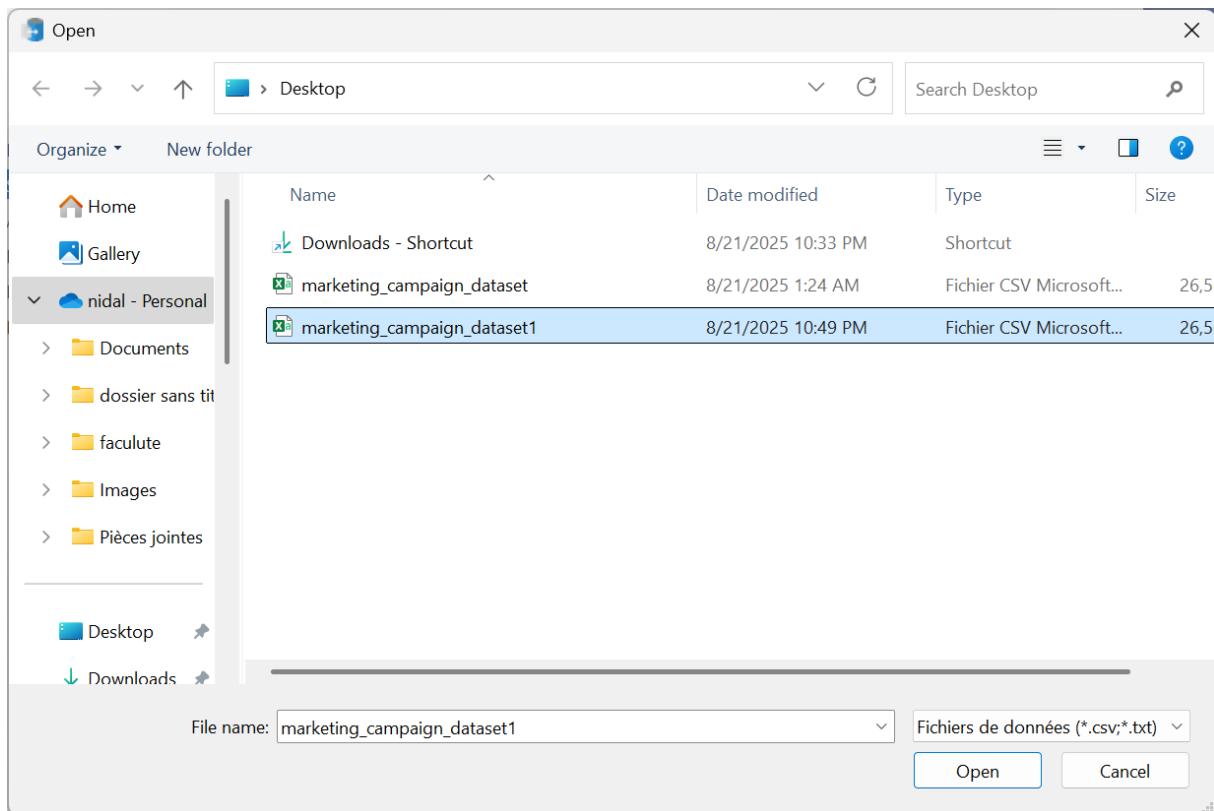
Importation dans SQL Server (SSMS)

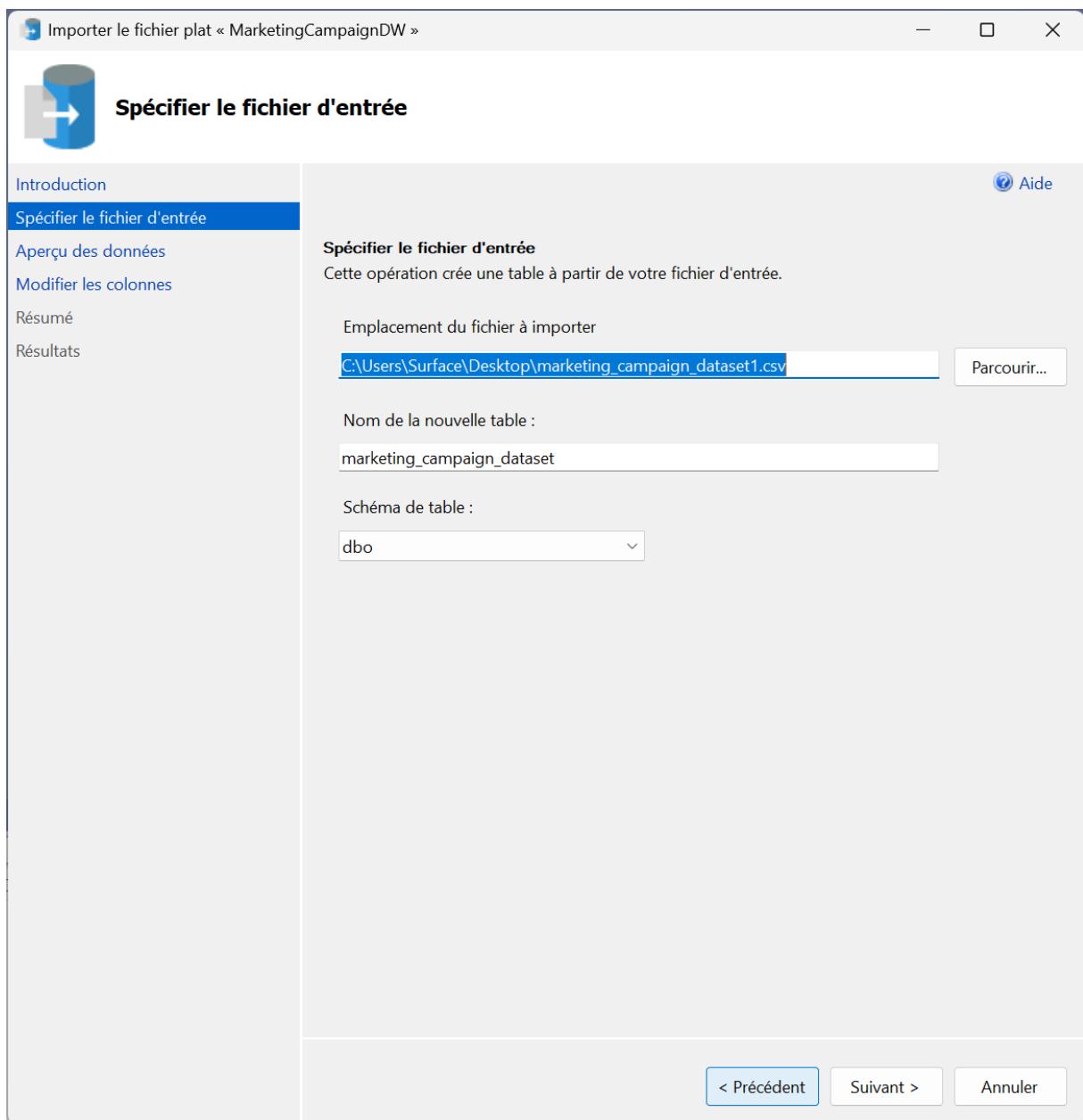
Après validation dans Excel, les données ont été importées dans **SQL Server Management Studio (SSMS)** pour la création d'un **entrepôt de données (Data Warehouse)** basé sur un schéma en étoile.

Les étapes principales de l'importation sont :

1. Création d'une **base de données dédiée [MyDataBase]**
2. Utilisation de l'option **Import Data** dans SSMS pour charger le fichier CSV.
3. Sélection de la source (fichier CSV) et de la destination (SQL Server).
4. Définition des noms de tables (par exemple [dbo].[marketing_campaign_dataset1]).
5. Vérification du type de données attribué à chaque colonne.
6. Chargement final et contrôle de la bonne insertion des données.







Importer le fichier plat « MyDataBase »

Modifier les colonnes

[Introduction](#) [Spécifier le fichier d'entrée](#) [Aperçu des données](#) [Aide](#)

Modifier les colonnes

Cette opération a généré le schéma de table suivant. Vérifiez si le schéma est correct et, dans le cas contraire, apportez-y les modifications nécessaires.

Nom de la colonne	Type de données	Clé primaire	Autoriser les valeurs Null
Campaign_ID	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Company	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Campaign_Type	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Target_Audience	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Duration	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Channel_Used	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Conversion_Rate	decimal(18, 10)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Acquisition_Cost	decimal(18, 10)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
ROI	decimal(18, 10)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Location	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Language	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Clicks	int	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Impressions	int	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Engagement_Score	int	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Customer_Segment	nvarchar(50)	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Date	date	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Granularité des lignes des rapports d'erreurs (impact sur les performances avec des plages plus petites) [Aucune plage](#)

[< Précédent](#) [Suivant >](#) [Annuler](#)

Vérification des données importées

Une fois les données chargées dans SSMS, des requêtes simples ont été exécutées afin de vérifier :

- le nombre total de lignes importées,
- la présence des colonnes principales,
- l'intégrité des valeurs (ex. : aucune colonne critique n'est totalement vide).

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. At the top, the menu bar includes Fichier, Edition, Afficher, Requête, Git, Projet, Outils, Extensions, Fenêtre, Aide, and Rechercher. Below the menu is a toolbar with various icons. The main window has tabs for 'Explorateur d'objets' and 'SQLQuery2.sql...Surface (67)'. The query results grid displays a table with 22 rows of data from the 'marketing_campaign_dataset1' table. The columns include Campaign_ID, Company, Campaign_Type, Target_Audience, Duration, Channel_Used, Conversion_Rate, Acquisition_Cost, ROI, Location, Language, Clicks, Impressions, Engagement_Score, Customer_Segment, and Date. The status bar at the bottom shows 'Execution réussie' (Execution successful), '(localdb)\MSSQLLocalDB (SQL Server 15.0.4382 - NID)' as the connection, 'MyDataBase' as the database, and '00:00:00 1,000 lignes' (00:00:00 1,000 lines).

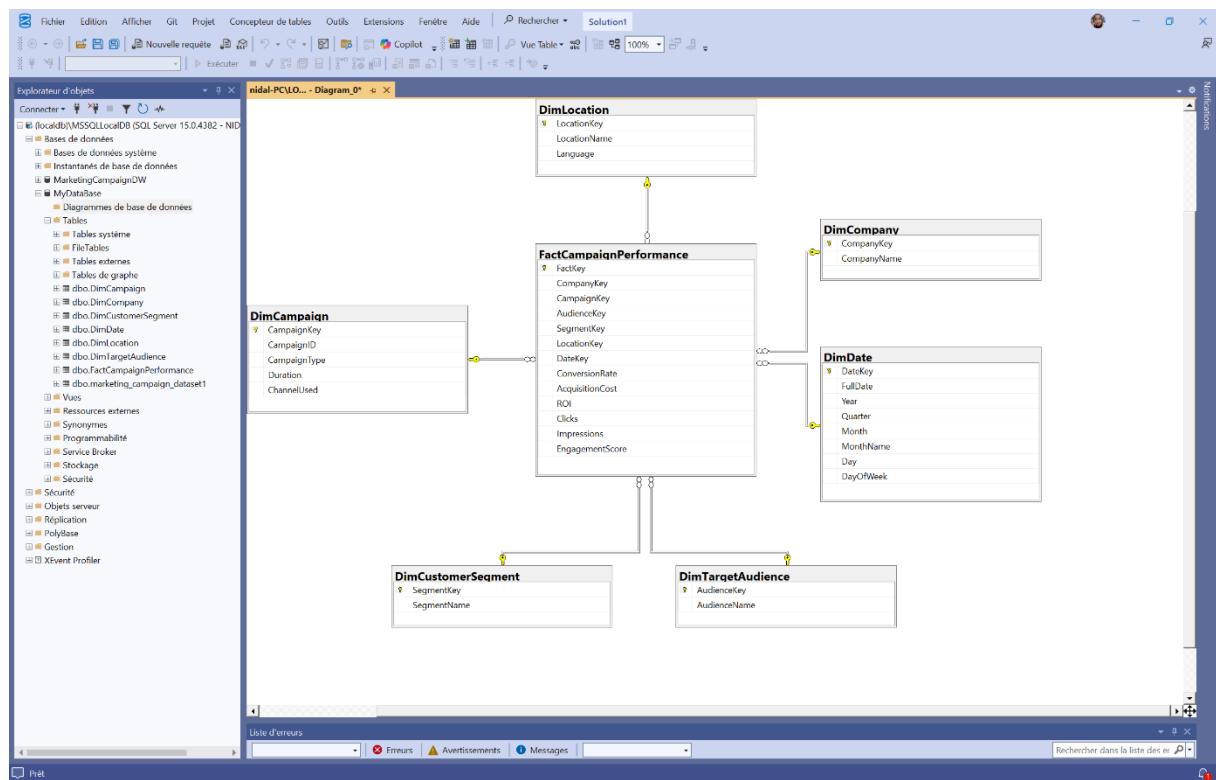
Campaign_ID	Company	Campaign_Type	Target_Audience	Duration	Channel_Used	Conversion_Rate	Acquisition_Cost	ROI	Location	Language	Clicks	Impressions	Engagement_Score	Customer_Segment	Date
1	Innovate Industries	Email	Men 18-24	30 days	Google Ads	0.04000000	\$16,174.00	6.290	Chicago	Spanish	508	1922	6	Health & Wellness	2021-01-01
2	NetGen Systems	Email	Women 35-44	60 days	Google Ads	0.12000000	\$11,566.00	5.610	New York	German	116	7523	7	Fashionista	2021-01-02
3	Alpha Innovations	Influencer	Men 25-34	30 days	YouTube	0.07000000	\$10,200.00	7.180	Los Angeles	French	564	7698	1	Outdoor Adventur.	2021-01-03
4	DataTech Solutions	Display	All Ages	60 days	YouTube	0.11000000	\$12,724.00	5.550	Miami	Mandarin	217	1820	7	Health & Wellness	2021-01-04
5	NetGen Systems	Email	Men 25-34	180 days	Facebook	0.08000000	\$1,000.00	0.500	Los Angeles	Spanish	22	1200	2	Foodies	2021-01-05
6	DataTech Solutions	Display	All Ages	15 days	Instagram	0.07000000	\$9,716.00	4.360	New York	German	100	1643	1	Foodies	2021-01-06
7	NetGen Systems	Email	Women 35-44	60 days	Website	0.13000000	\$11,067.00	2.960	Los Angeles	Spanish	617	6749	10	Tech Enthusiasts	2021-01-07
8	DataTech Solutions	Search	Men 18-24	45 days	Google Ads	0.08000000	\$13,280.00	5.550	Los Angeles	Mandarin	624	7654	7	Outdoor Adventur.	2021-01-08
9	Alpha Innovations	Social Media	Women 35-44	15 days	Facebook	0.09000000	\$18,066.00	6.730	Chicago	German	861	1754	6	Tech Enthusiasts	2021-01-09
10	TechCorp	Email	Women 35-44	15 days	Instagram	0.09000000	\$13,766.00	3.780	Los Angeles	English	642	3856	3	Tech Enthusiasts	2021-01-10
11	NetGen Systems	Display	Men 25-34	45 days	Email	0.12000000	\$8,590.00	3.490	New York	Spanish	321	6628	10	Tech Enthusiasts	2021-01-11
12	Innovate Industries	Influencer	Men 25-34	60 days	Facebook	0.08000000	\$10,000.00	3.000	Los Angeles	Mandarin	292	5931	1	Foodies	2021-01-12
13	TechCorp	Social Media	Men 25-34	60 days	Facebook	0.09000000	\$17,188.00	4.910	Chicago	Spanish	677	8817	10	Tech Enthusiasts	2021-01-13
14	TechCorp	Email	Men 25-34	45 days	Instagram	0.14000000	\$9,975.00	7.060	New York	German	994	2301	4	Health & Wellness	2021-01-14
15	TechCorp	Display	All Ages	45 days	Website	0.04000000	\$11,346.00	5.280	Chicago	Spanish	452	8470	1	Outdoor Adventur.	2021-01-15
16	Innovate Industries	Social Media	Women 35-44	60 days	YouTube	0.11000000	\$9,407.00	2.910	New York	German	299	1512	5	Health & Wellness	2021-01-16
17	Innovate Industries	Display	Women 35-44	45 days	Website	0.08000000	\$5,476.00	4.530	Houston	English	931	2488	3	Fashionista	2021-01-17
18	Alpha Innovations	Influencer	Men 18-24	15 days	Instagram	0.14000000	\$9,495.00	4.480	Miami	French	218	9264	9	Health & Wellness	2021-01-18
19	Alpha Innovations	Social Media	Men 25-34	60 days	Google Ads	0.09000000	\$19,224.00	6.000	New York	French	162	5798	1	Foodies	2021-01-19
20	20 DataTech Solutions	Search	Women 25-34	15 days	Google Ads	0.08000000	\$10,200.00	3.500	Miami	French	160	3677	1	Tech Enthusiasts	2021-01-20
21	DataTech Solutions	Search	Women 25-34	15 days	Email	0.04000000	\$8,500.00	7.000	New York	Spanish	975	1661	3	Outdoor Adventur.	2021-01-21
22	TechCorp	Influencer	All Ages	30 days	Facebook	0.02000000	\$12,824.00	6.210	New York	French	319	0596	2	Foodies	2021-01-22

Conception du Modèle de Données

Choix du Schéma en Étoile

Pour ce projet, nous avons choisi un schéma en étoile pour organiser les données. Le schéma en étoile est adapté pour les systèmes de décision, car il permet de structurer les données de manière claire et efficace en séparant les tables de faits et les tables de dimensions.

Le modèle est conçu de manière à faciliter l'analyse des performances des campagnes marketing à travers différentes dimensions, telles que la campagne, le segment de clients, la date, et l'audience ciblée.



Description des Tables de Dimensions

Les tables de dimensions permettent de stocker les attributs descriptifs relatifs aux données des campagnes marketing. Voici les principales tables de dimensions utilisées dans le modèle :

```

USE MarketingCampaignDW;
GO
-- 1. DIMENSIONS
-- Dimension Date
CREATE TABLE DimDate (
    DateKey INT PRIMARY KEY,
    FullDate DATE,
    Day INT,
    Month INT,
    MonthName NVARCHAR(20),
    Quarter INT,
    Year INT
);
-- Dimension Company
CREATE TABLE DimCompany (
    CompanyKey INT IDENTITY(1,1) PRIMARY KEY,
    CompanyName NVARCHAR(100)
);
-- Dimension Campaign
CREATE TABLE DimCampaign (
    CampaignKey INT IDENTITY(1,1) PRIMARY KEY,
    CampaignType NVARCHAR(100),
    TargetAudience NVARCHAR(100),
    Duration NVARCHAR(50)
);
-- Dimension Channel
CREATE TABLE DimChannel (
    ChannelKey INT IDENTITY(1,1) PRIMARY KEY,
    ChannelName NVARCHAR(50)
);
-- Dimension Location
CREATE TABLE DimLocation (
    LocationKey INT IDENTITY(1,1) PRIMARY KEY,
    Country NVARCHAR(100),
    Language NVARCHAR(50)
);
-- Dimension Customer Segment
CREATE TABLE DimCustomerSegment (
    SegmentKey INT IDENTITY(1,1) PRIMARY KEY,
    SegmentName NVARCHAR(100)
);

```

1. DimCampaign

Cette table contient des informations détaillées sur les campagnes marketing. Elle inclut :

- CampaignKey : Clé primaire pour identifier chaque campagne.
- CampaignID : Identifiant unique de la campagne.
- CampaignType : Type de campagne (par exemple, emailing, publicité en ligne, etc.).
- Duration : Durée de la campagne.
- ChannelUsed : Canal utilisé pour la campagne (réseaux sociaux, email, etc.).

2. DimCompany

Cette table contient des informations sur l'entreprise :

- CompanyKey : Clé primaire pour identifier chaque entreprise.
- CompanyName : Nom de l'entreprise.

3. DimCustomerSegment

Cette table décrit les segments de clients visés par les campagnes :

- SegmentKey : Clé primaire du segment.
- SegmentName : Nom du segment (par exemple, jeunes adultes, seniors).

4. DimDate

La table DimDate contient des informations sur les dates associées aux campagnes :

- DateKey : Clé primaire pour identifier chaque date.

- FullDate : Date complète.
- Year : Année de la campagne.
- MonthName : Mois de la campagne.
- Day : Jour de la campagne.
- DayOfWeek : Jour de la semaine.

5. DimLocation

Cette table représente la localisation de l'audience cible :

- LocationKey : Clé primaire pour identifier chaque localisation.
- LocationName : Nom de la localisation (par exemple, pays, ville).
- Language : Langue utilisée dans la campagne.

6. DimTargetAudience

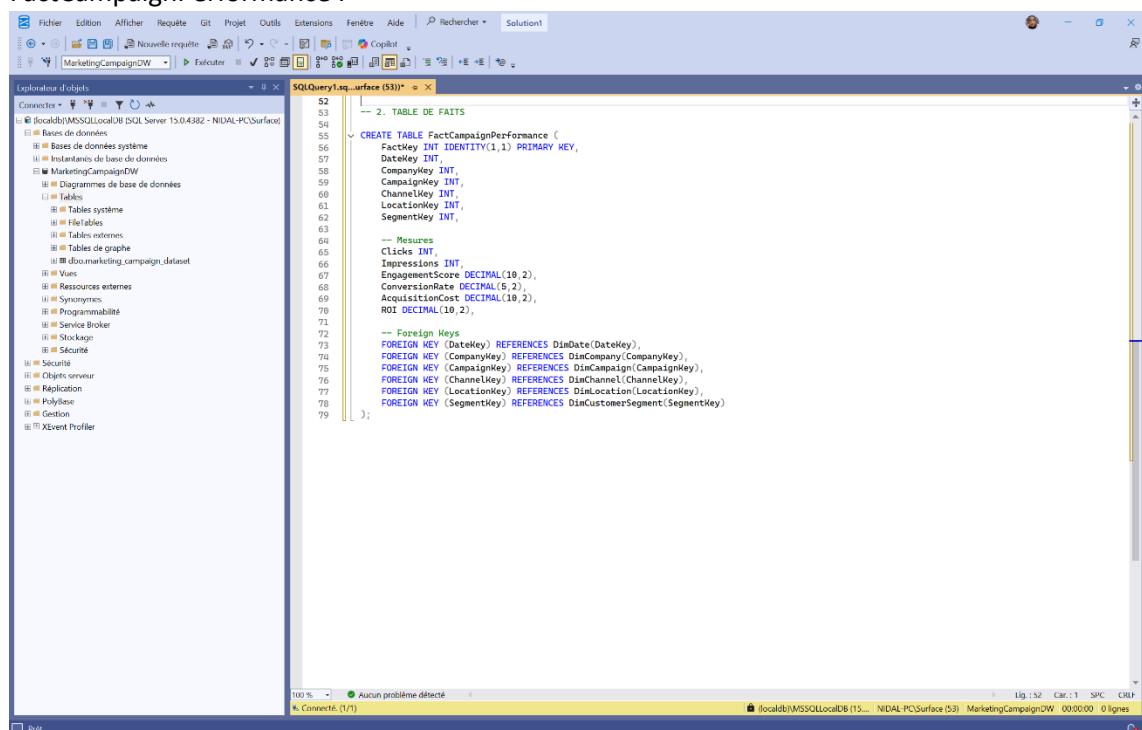
Cette table contient des informations sur le public cible de la campagne :

- AudienceKey : Clé primaire de l'audience.
- AudienceName : Nom de l'audience (par exemple, jeunes adultes, professionnels).

Description de la Table de Faits

La table de faits centralise les données quantitatives associées aux campagnes marketing. Elle permet d'analyser l'impact des campagnes sur les ventes. Voici la description de la table

FactCampaignPerformance :



```

Fichier Edition Afficher Requête Git Projet Outils Extensions Fenêtre Aide Rechercher Solution1
[localdb]\MSSQLLocalDB SQL Server 15.0.4382 - NIDAL-PC\Surface
Explorateur d'objets
Connexions Nouvelle requête Générateur de rapports Copilot
MarketingCampaignDW Exécuter
[localdb]\MSSQLLocalDB (53) MarketingCampaignDW
2. TABLE DE FAITS
CREATE TABLE FactCampaignPerformance (
    FactKey INT IDENTITY(1,1) PRIMARY KEY,
    DateKey INT,
    CompanyKey INT,
    MarketingKey INT,
    CampaignKey INT,
    ChannelKey INT,
    LocationKey INT,
    SegmentKey INT,
    -- Measures
    Clicks INT,
    Impressions INT,
    EngagementScore DECIMAL(10, 2),
    ConversionRate DECIMAL(5, 2),
    AcquisitionCost DECIMAL(10, 2),
    ROI DECIMAL(10, 2),
    -- Foreign Keys
    FOREIGN KEY (DateKey) REFERENCES DimDate(DateKey),
    FOREIGN KEY (CompanyKey) REFERENCES DimCompany(CompanyKey),
    FOREIGN KEY (MarketingKey) REFERENCES DimMarketing(MarketingKey),
    FOREIGN KEY (ChannelKey) REFERENCES DimChannel(ChannelKey),
    FOREIGN KEY (LocationKey) REFERENCES DimLocation(LocationKey),
    FOREIGN KEY (SegmentKey) REFERENCES DimCustomerSegment(SegmentKey)
);

```

Aucun problème détecté

100 % 6. Connecté. (1/1)

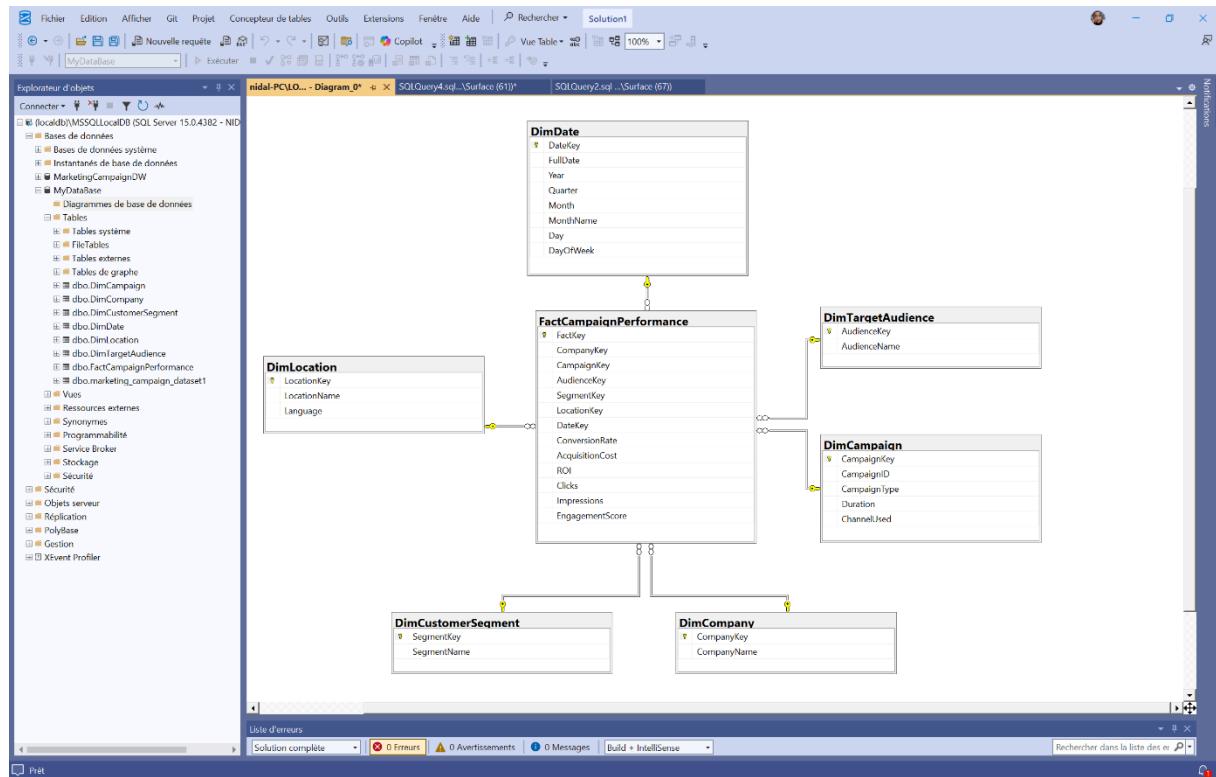
Lig.: 52 Car.: 1 SPC: 0 CRF: 0

1. FactCampaignPerformance

Cette table est au cœur du modèle et contient les mesures clés permettant de mesurer la performance des campagnes :

- FactKey : Clé primaire pour identifier chaque enregistrement de performance de campagne.
- CompanyKey : Clé étrangère liant cette table à la table DimCompany.
- CampaignKey : Clé étrangère liant cette table à la table DimCampaign.
- AudienceKey : Clé étrangère liant cette table à la table DimTargetAudience.
- SegmentKey : Clé étrangère liant cette table à la table DimCustomerSegment.
- LocationKey : Clé étrangère liant cette table à la table DimLocation.
- DateKey : Clé étrangère liant cette table à la table DimDate.
- ConversionRate : Taux de conversion des campagnes.
- AcquisitionCost : Coût d'acquisition par campagne.
- ROI (Return on Investment) : Retour sur investissement des campagnes.
- Clicks : Nombre de clics générés par la campagne.
- Impressions : Nombre d'impressions générées par la campagne.
- EngagementScore : Score d'engagement basé sur les interactions (clics, partages, etc.).

Relation entre les Tables de Dimensions et la Table de Faits



Les tables de dimensions sont reliées à la table de faits par des clés étrangères, ce qui permet de réaliser des analyses détaillées sur les campagnes marketing en fonction de plusieurs critères, tels que le type de campagne, le segment de clients, la localisation géographique, et la date.

Les relations entre les tables de dimensions et la table de faits sont les suivantes :

- DimCampaign est liée à FactCampaignPerformance via CampaignKey.
- DimCompany est liée à FactCampaignPerformance via CompanyKey.
- DimTargetAudience est liée à FactCampaignPerformance via AudienceKey.
- DimCustomerSegment est liée à FactCampaignPerformance via SegmentKey.
- DimLocation est liée à FactCampaignPerformance via LocationKey.
- DimDate est liée à FactCampaignPerformance via DateKey.

Création du Projet SSIS et Test de Connexion

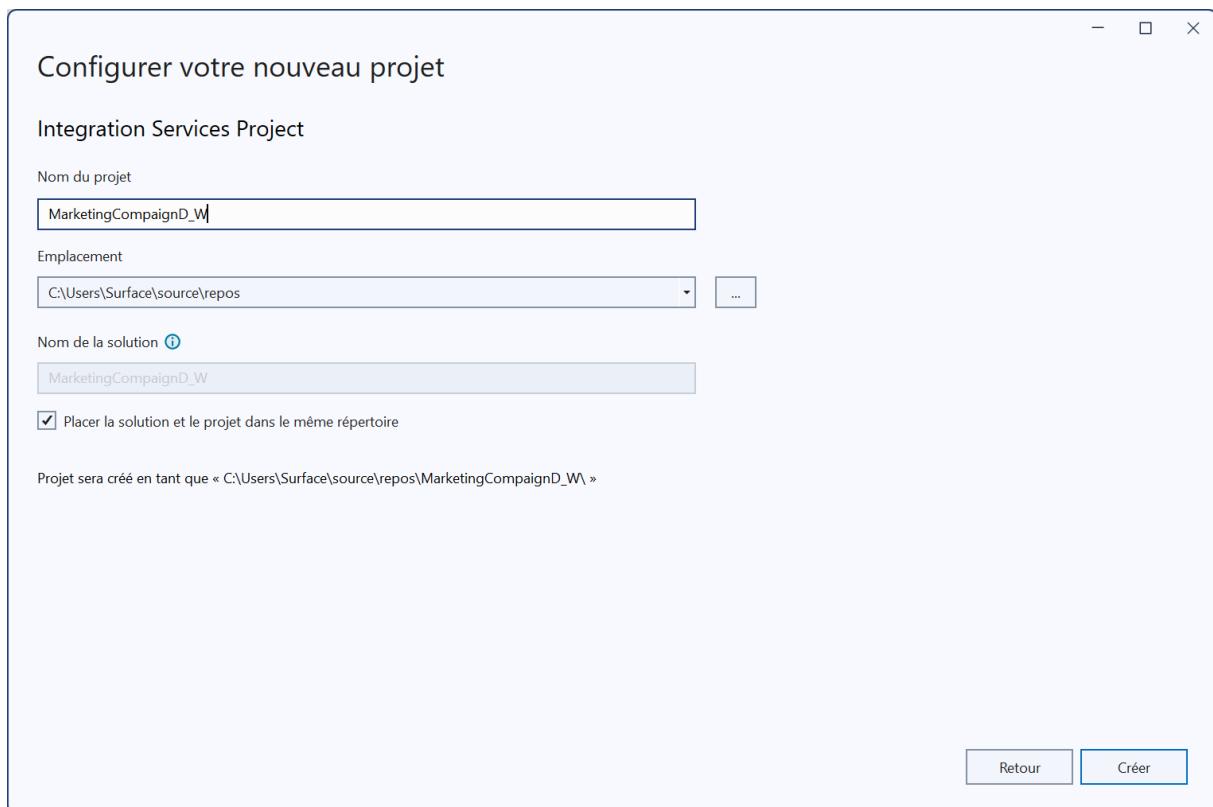


1. Création d'un Nouveau Projet SSIS

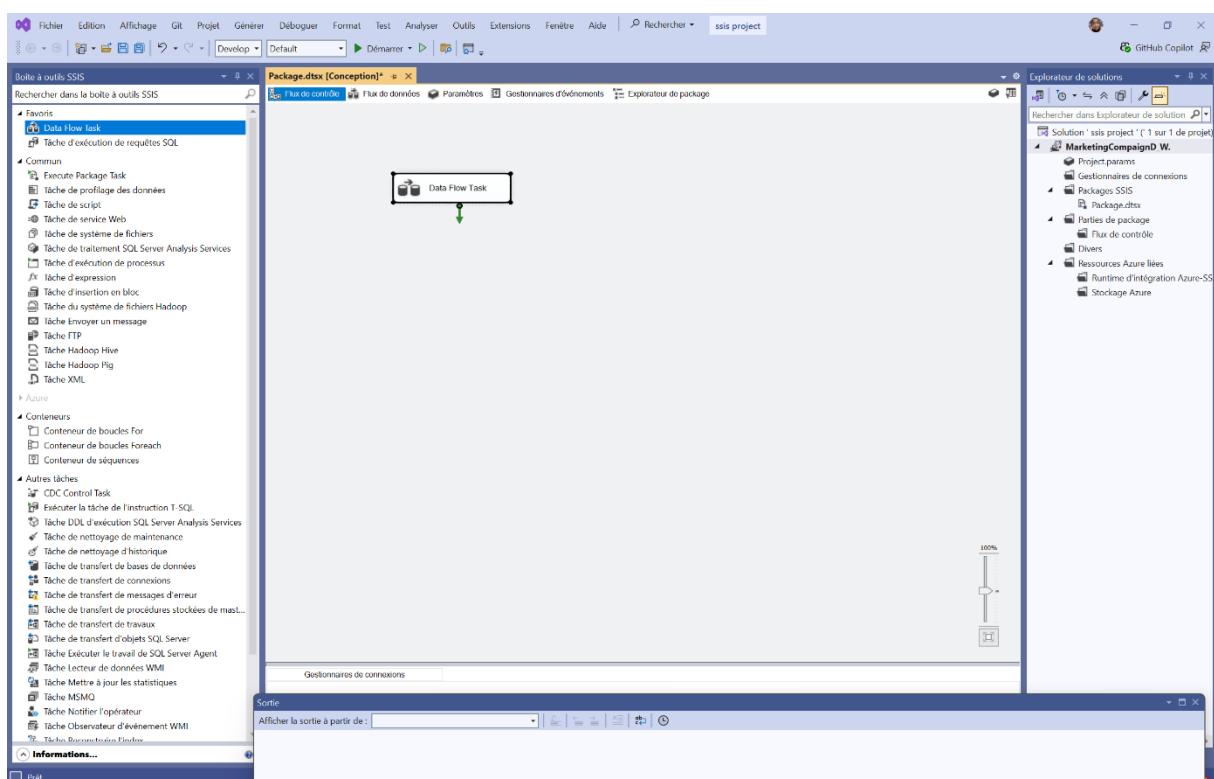
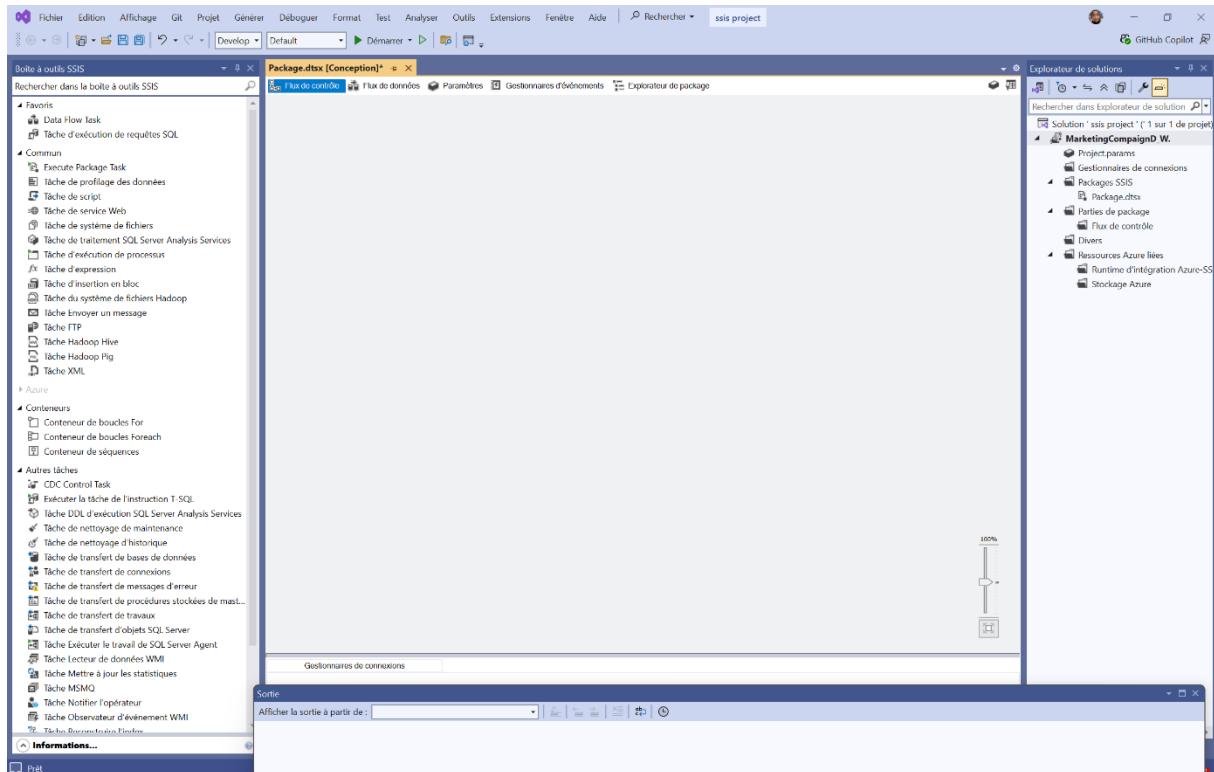
A screenshot of the "Créer un projet" (Create a project) dialog in Microsoft Visual Studio. The search bar at the top contains the text "ssis". The results list shows three items: "Integration Services Project", "Projet de base de données SQL Server", and "Integration Services Project (Azure-Enabled)". The first item is highlighted. A tooltip for the first item provides a brief description: "This project may be used for building high performance data integration and workflow solutions that can be run on SSIS catalog, including extraction, transformation, and loading (ETL) operations for data warehousing." At the bottom of the dialog, there is a link "Vous n'arrivez pas à trouver ce que vous cherchez ? Installer plus d'outils et de fonctionnalités" and a "Suivant" (Next) button.

Comme mentionné précédemment, la première étape a été de créer un nouveau projet SSIS dans Visual Studio. Cela a permis de préparer l'environnement pour le processus d'ETL.

1. Ouvrir Visual Studio et sélectionner l'option "Integration Services Project".
2. Nommer le projet "MarketingCampaignD_W" et définir l'emplacement et la solution pour le projet.



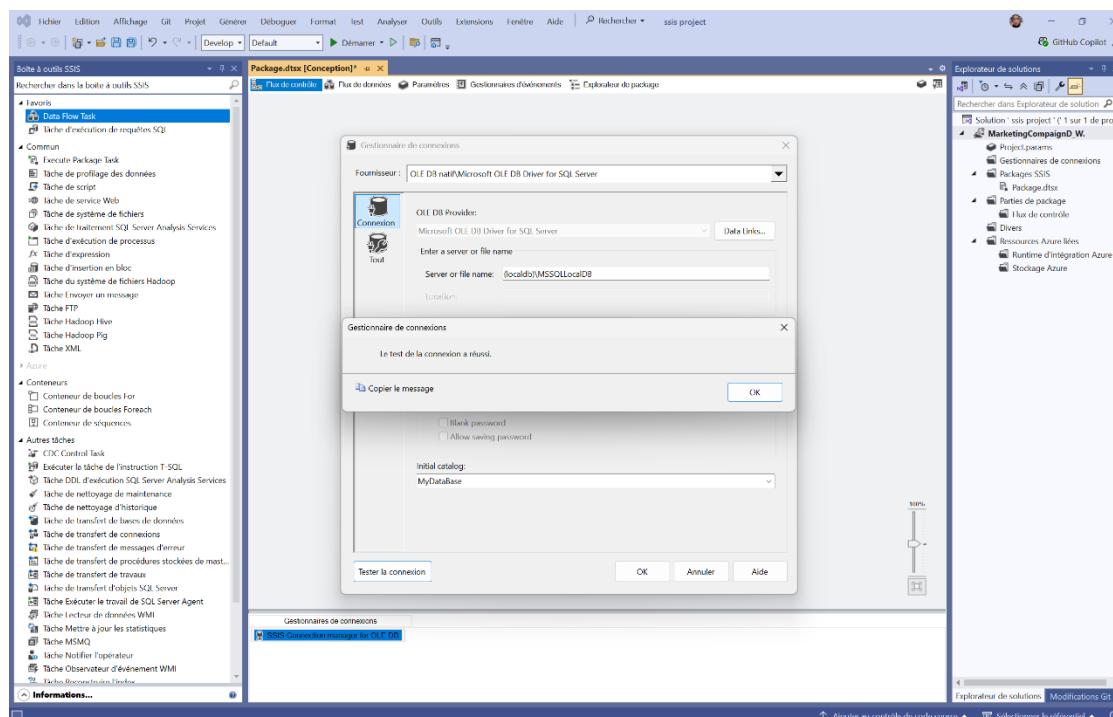
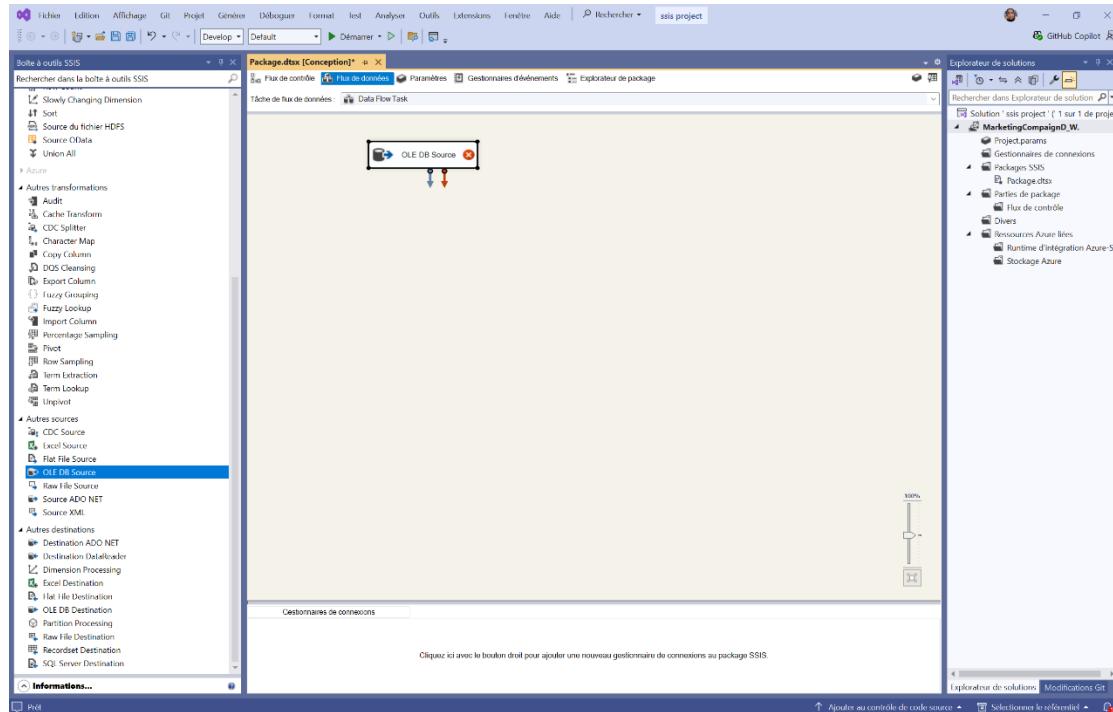
Une fois le projet créé, j'ai commencé à configurer les Data Flow Tasks pour l'extraction et le chargement des données.



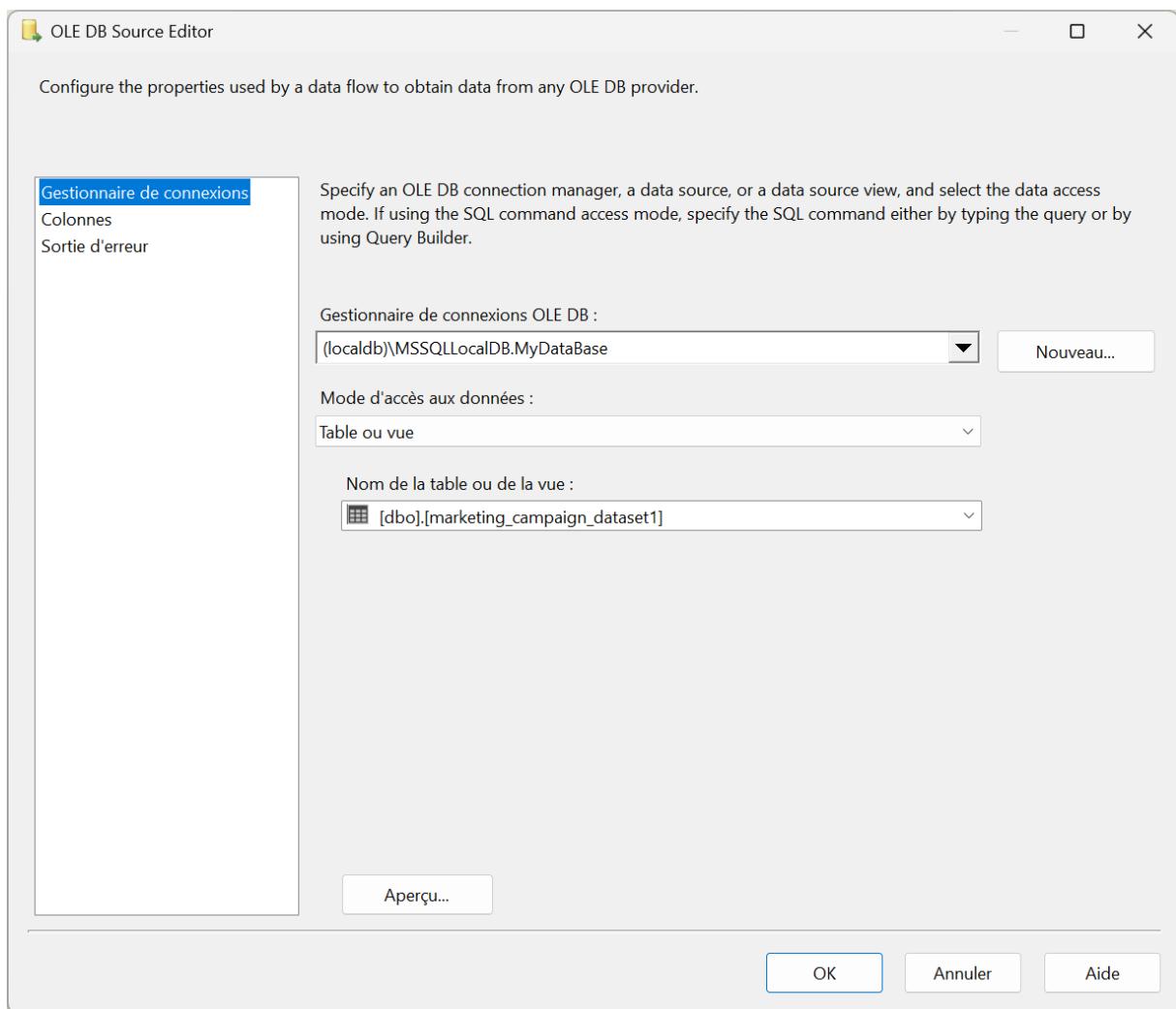
Test de la Connexion OLE DB

Avant de commencer à configurer les flux de données, il était essentiel de tester la connexion entre la source de données OLE DB Source et la base de données SQL Server pour garantir que la connexion fonctionnait correctement.

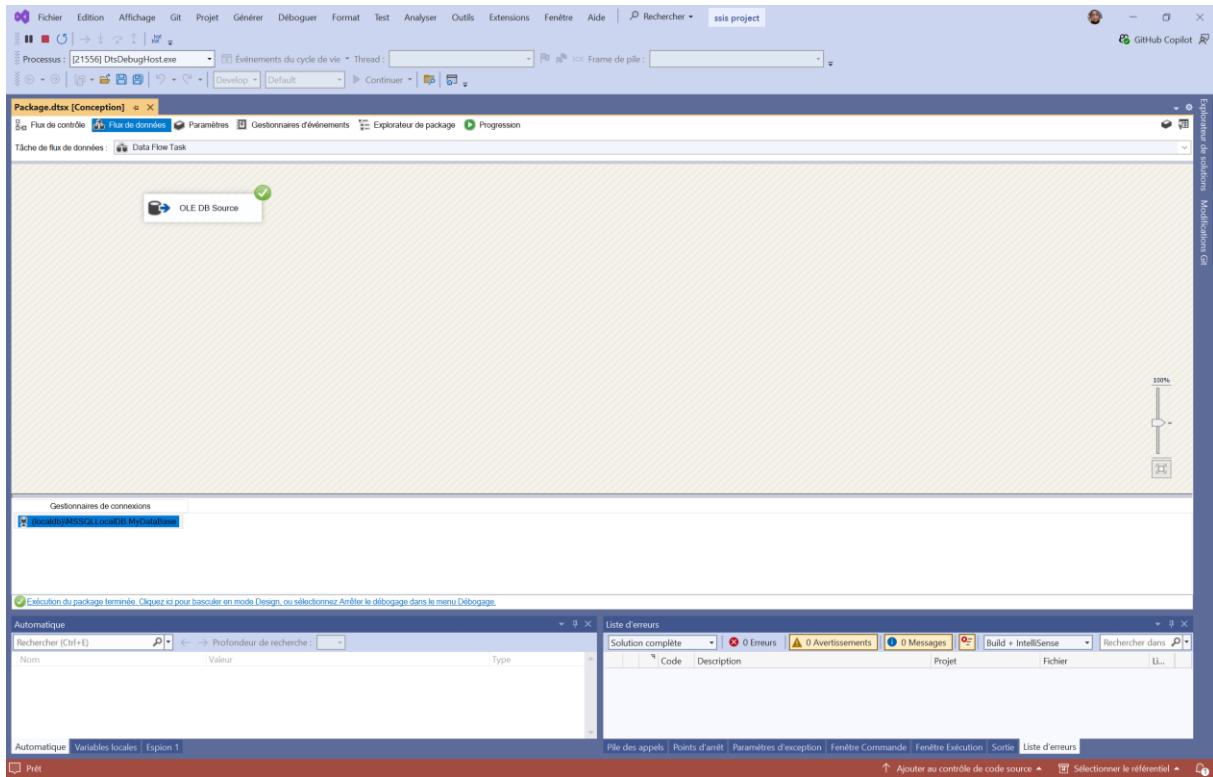
Dans SSIS, j'ai ajouté une connexion OLE DB Source pour me connecter à ma base de données SQL Server



Un test de connexion a été effectué pour s'assurer que la source de données OLE DB était bien connectée à ma base de données SQL Server.

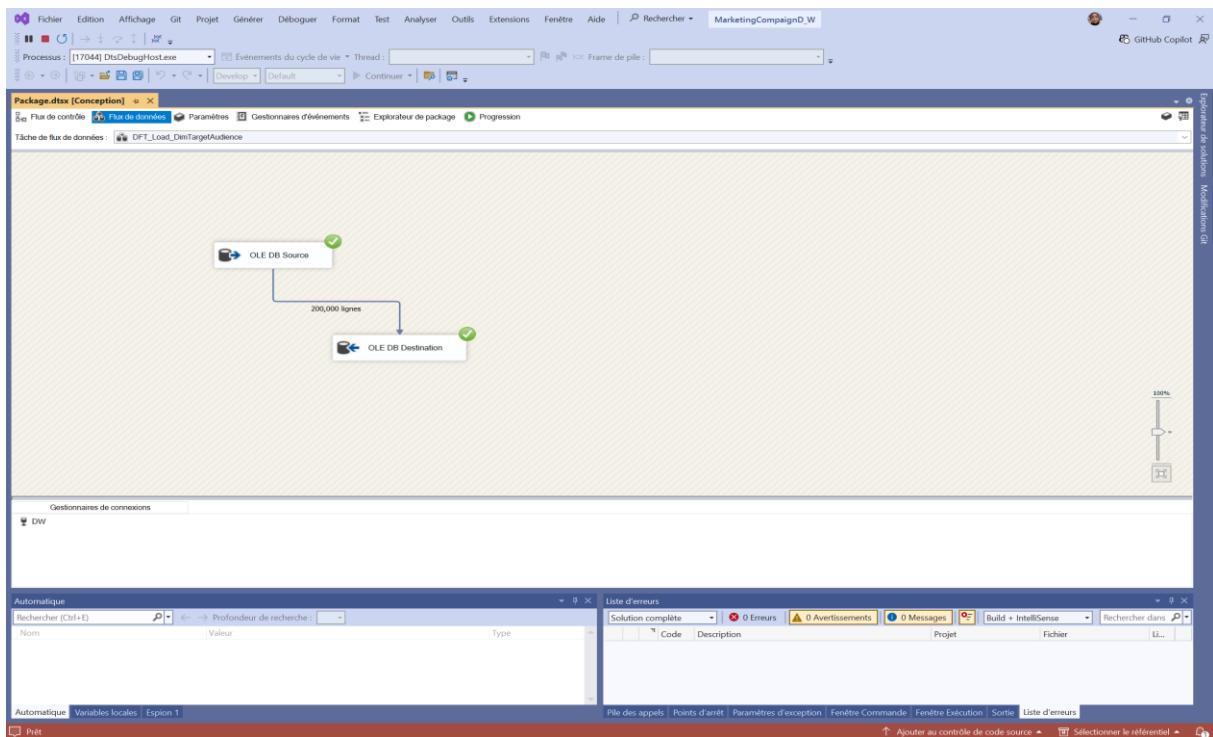


Une fois le test effectué avec succès, j'ai pu vérifier que la source de données était correctement configurée pour l'extraction des données.

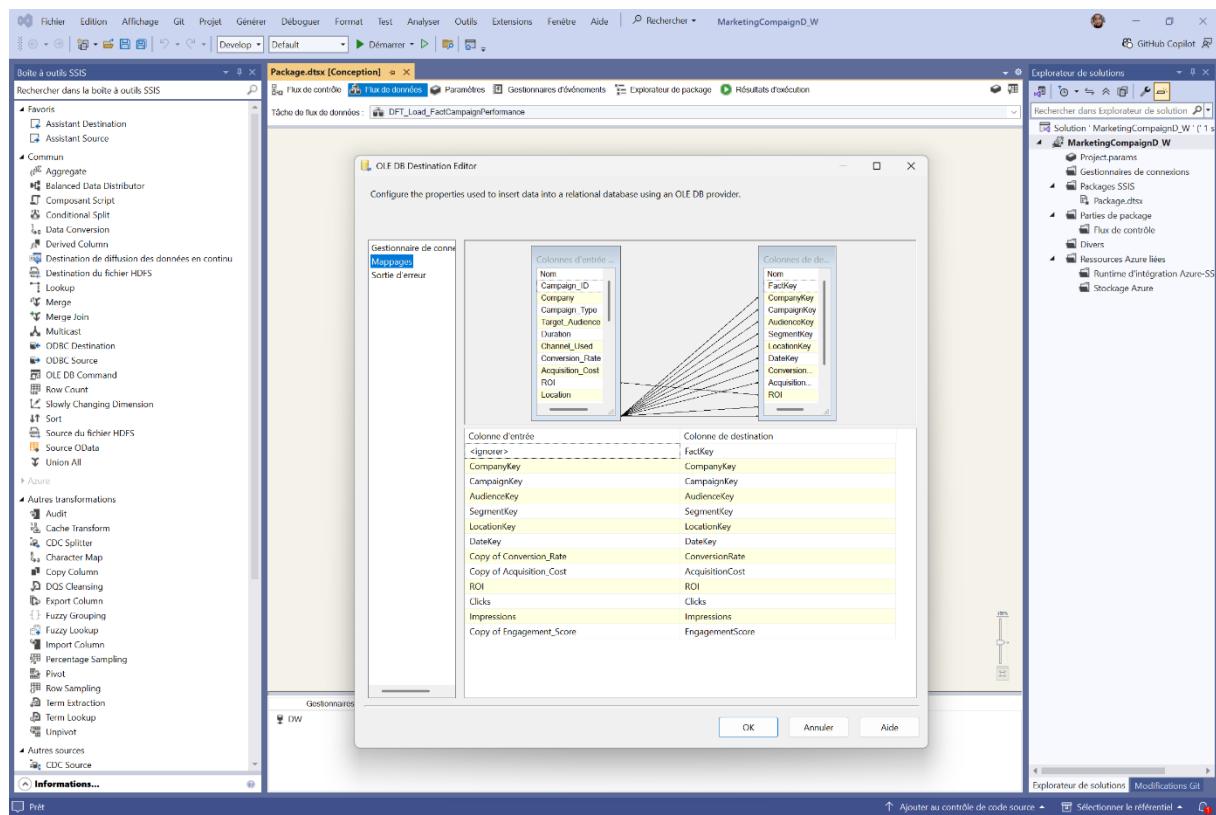
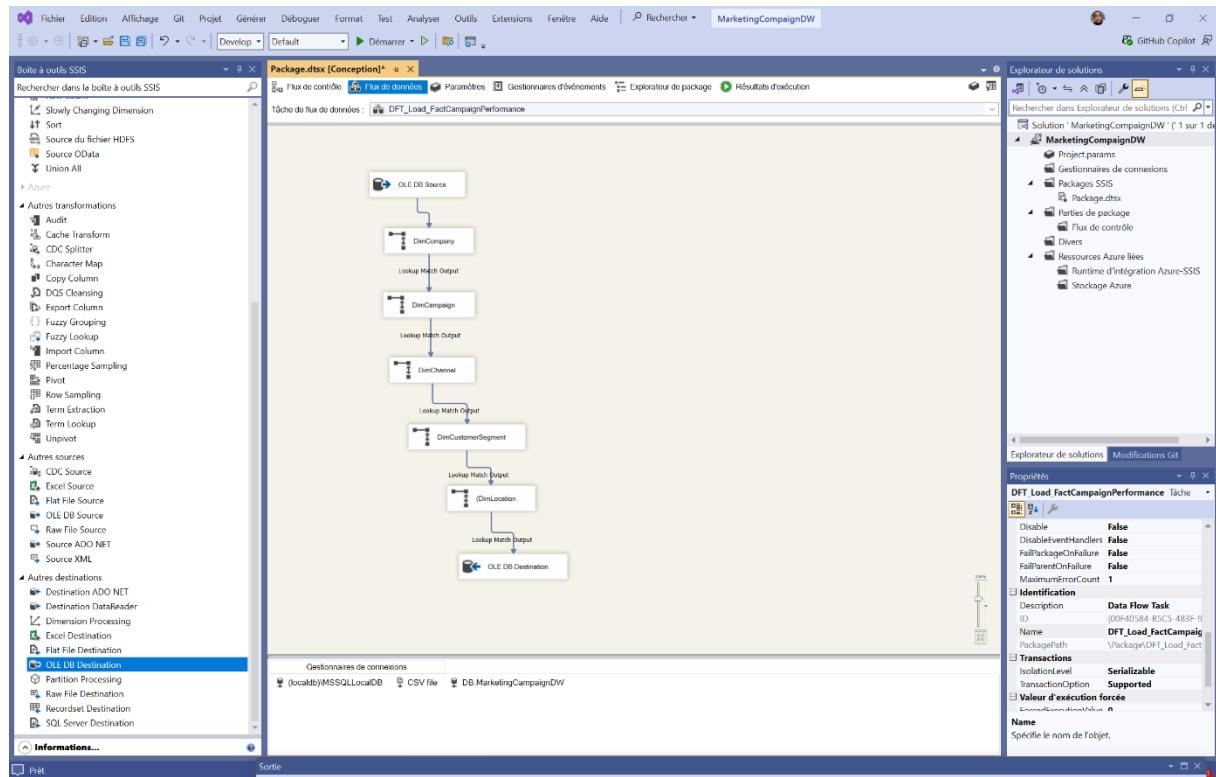


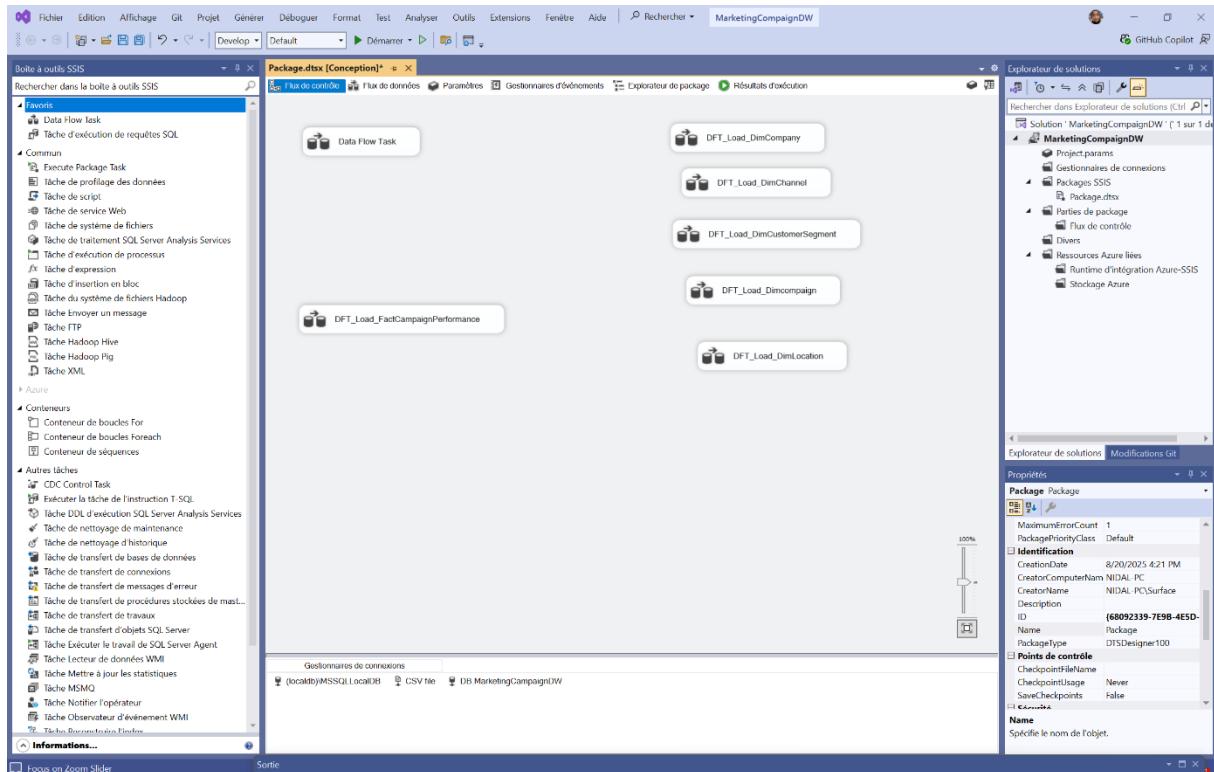
Extraction des Données Brutes

Dans cette phase, l'objectif principal était d'extraire les données de la source OldDB et de les charger directement dans la destination OldDB. Ce processus a été effectué à l'aide de SSIS avec des Data Flow Tasks. Le choix a été fait de ne pas inclure de transformations complexes à ce stade, afin de maintenir la simplicité du flux de données et de transférer les données telles quelles, sans modification.

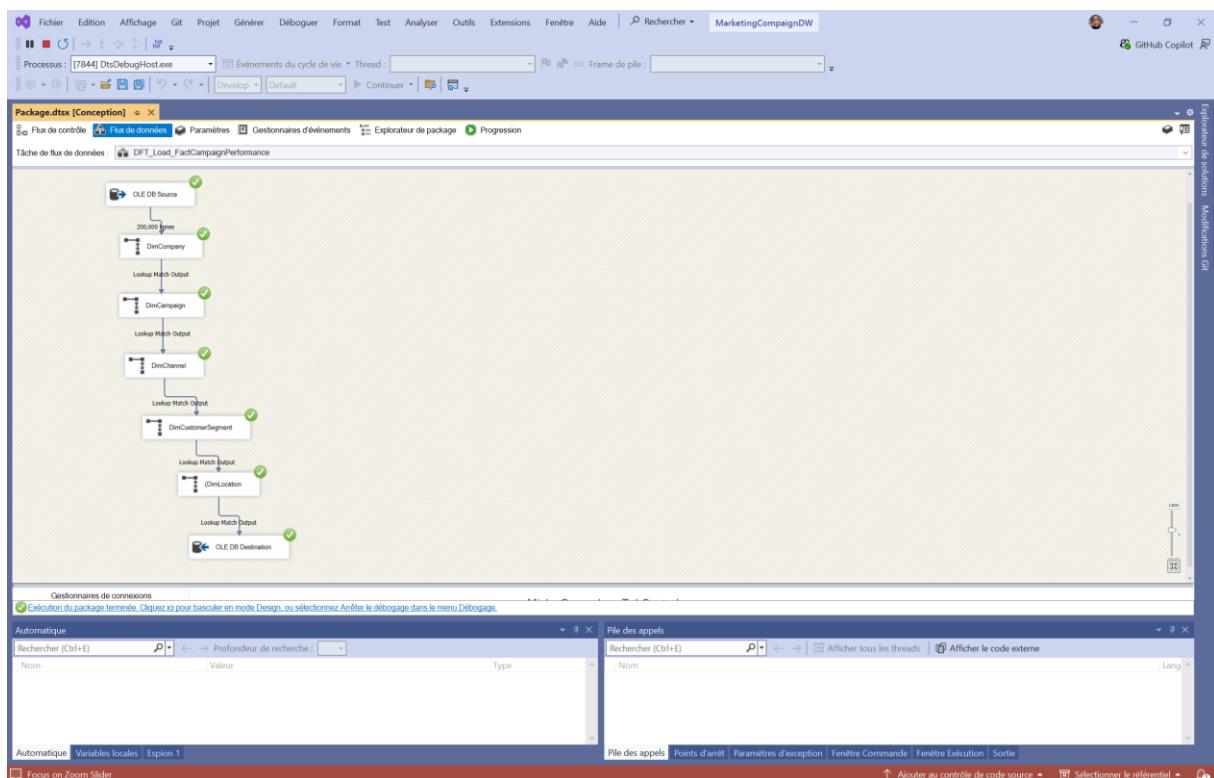


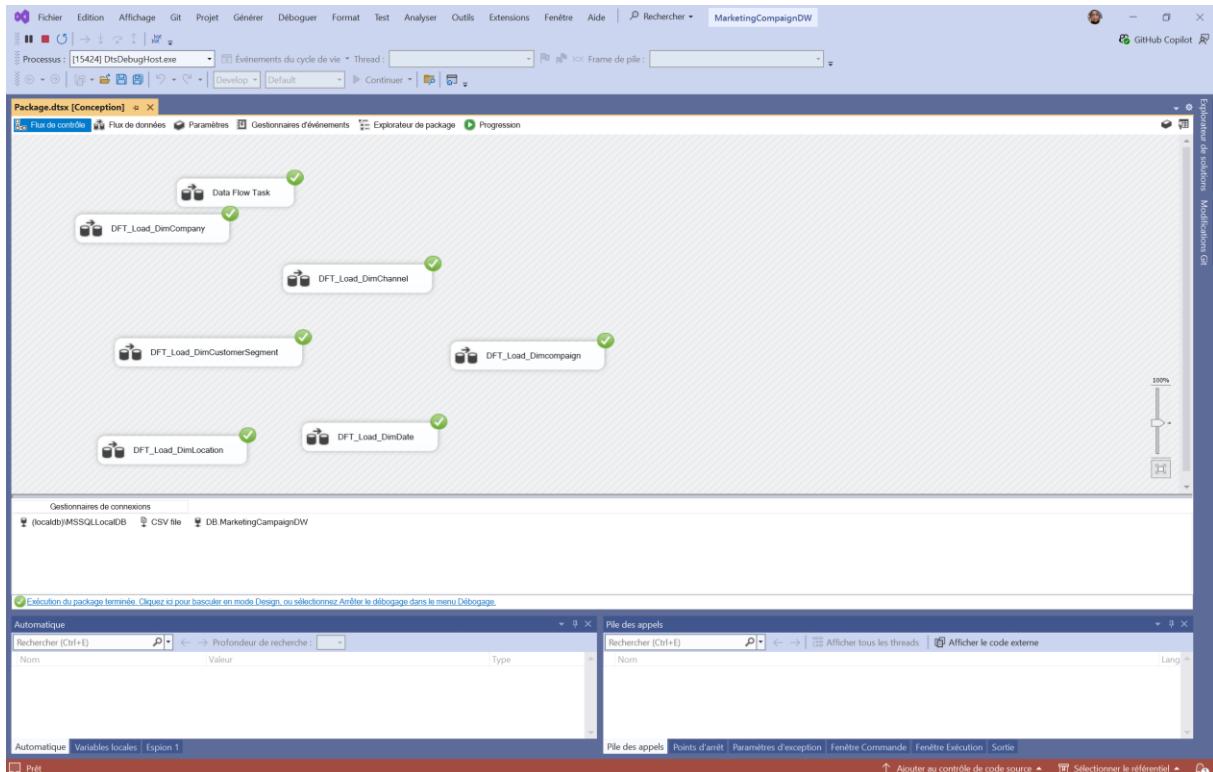
Chargement des Données dans la Base de Données





Exécution du Package SSIS





Le package SSIS a été exécuté avec succès pour effectuer l'extraction et le chargement des données.

Vérification du Chargement des Données

(Sélection des 10 Premières Lignes)

Pour chaque table, nous allons récupérer les 10 premières lignes pour vérifier que les données sont bien chargées. Voici les requêtes SQL pour cela :

Requête SQL pour chaque Table de Dimension

Fichier Edition Afficher Requête Git Projet Outils Extensions Fenêtre Aide | Rechercher | Solution1

Nouvelle requête Exécuter Copier

Explorateur d'objets SQLQuery20.sql - Surface (62) mynewdatabase..surface (69)* SQLQuery19.sql - Surface (52)* nidal-PC\OC_c - Diagram 1* SQLQuery18.sql - Surface (60)

```

73
74 -- Vérification des 10 premières lignes de DimCampaign
75 SELECT TOP 10 * FROM dbo.DimCampaign;
76
77 -- Vérification des 10 premières lignes de DimCompany
78 SELECT TOP 10 * FROM dbo.DimCompany;
79
80 -- Vérification des 10 premières lignes de DimCustomerSegment
81 SELECT TOP 10 * FROM dbo.DimCustomerSegment;
82
83 -- Vérification des 10 premières lignes de DimDate
84 SELECT TOP 10 * FROM dbo.DimDate;
85
86 -- Vérification des 10 premières lignes de DimLocation
87 SELECT TOP 10 * FROM dbo.DimLocation;
88
89 -- Vérification des 10 premières lignes de DimTargetAudience
90 SELECT TOP 10 * FROM dbo.DimTargetAudience;
91

```

Réultats Messages

CampaignKey	CampaignID	CampaignType	Duration	ChannelUsed
1	1	Email	30 days	Google Ads
2	2	Email	60 days	Google Ads
3	3	Influencer	30 days	YouTube
4	4	Display	60 days	YouTube
5	5	Email	15 days	YouTube
6	6	Display	15 days	Instagram
7	7	Email	60 days	Website
8	8	Search	45 days	Google Ads
9	9	Social Media	15 days	Facebook
10	10	Email	15 days	Instagram

Exécution de requête réussie

Liste d'erreurs

Rechercher dans la liste des erreurs

Fichier Edition Afficher Requête Git Projet Outils Extensions Fenêtre Aide | Rechercher | Solution1

Explorateur d'objets

```

73
74    -- Vérification des 10 premières lignes de DimCampaign
75    SELECT TOP 10 * FROM dbo.DimCampaign;
76
77    -- Vérification des 10 premières lignes de DimCompany
78    SELECT TOP 10 * FROM dbo.DimCompany;
79
80    -- Vérification des 10 premières lignes de DimCustomerSegment
81    SELECT TOP 10 * FROM dbo.DimCustomerSegment;
82
83    -- Vérification des 10 premières lignes de DimDate
84    SELECT TOP 10 * FROM dbo.DimDate;
85
86    -- Vérification des 10 premières lignes de DimLocation
87    SELECT TOP 10 * FROM dbo.DimLocation;
88
89    -- Vérification des 10 premières lignes de DimTargetAudience
90    SELECT TOP 10 * FROM dbo.DimTargetAudience;
91

```

Réultats Messages

SegmentKey	SegmentName
1	Health & Wellness
2	Fitness
3	Outdoor Adventures
4	Health & Wellness
5	Health & Wellness
6	Foodies
7	Tech Enthusiasts
8	Outdoor Adventures
9	Tech Enthusiasts
10	Tech Enthusiasts

Dataset

DateKey	FullDate
1	2023-01-01
2	2023-01-02
3	2023-01-03
4	2023-01-04
5	2023-01-05
6	2023-01-06
7	2023-01-07
8	2023-01-08
9	2023-01-09
10	2023-01-10

Exécution de requête réussie.

Liste d'erreurs

Solution complète | 10 Erreurs | 0 Avertissements | 0 Messages | Build + IntelliSense | Rechercher dans la liste des erreurs

Fichier Edition Afficher Requête Git Projet Outils Extensions Fenêtre Aide | Rechercher | Solution1

Explorateur d'objets

```

73
74    -- Vérification des 10 premières lignes de DimCampaign
75    SELECT TOP 10 * FROM dbo.DimCampaign;
76
77    -- Vérification des 10 premières lignes de DimCompany
78    SELECT TOP 10 * FROM dbo.DimCompany;
79
80    -- Vérification des 10 premières lignes de DimCustomerSegment
81    SELECT TOP 10 * FROM dbo.DimCustomerSegment;
82
83    -- Vérification des 10 premières lignes de DimDate
84    SELECT TOP 10 * FROM dbo.DimDate;
85
86    -- Vérification des 10 premières lignes de DimLocation
87    SELECT TOP 10 * FROM dbo.DimLocation;
88
89    -- Vérification des 10 premières lignes de DimTargetAudience
90    SELECT TOP 10 * FROM dbo.DimTargetAudience;
91

```

Réultats Messages

LocationKey	LocationName	Language
1	Chicago	Spanish
2	New York	German
3	Los Angeles	French
4	Miami	Mandarin
5	Los Angeles	Mandarin
6	New York	German
7	Los Angeles	Spanish
8	Los Angeles	Mandarin
9	Chicago	German
10	Los Angeles	English

AudienceKey AudienceName

AudienceKey	AudienceName
1	Men 18-24
2	Women 35-44
3	Men 25-34
4	All Ages
5	Men 25-34
6	All Ages
7	Women 35-44
8	Men 18-24
9	Women 35-44
10	Women 35-44

Exécution de requête réussie.

Liste d'erreurs

Solution complète | 10 Erreurs | 0 Avertissements | 0 Messages | Build + IntelliSense | Rechercher dans la liste des erreurs

Ces requêtes permettent de vérifier que chaque table de dimension contient bien des données et que les enregistrements sont correctement insérés.

Requête SQL pour la Table de Faits

```

92
93
94
95
96    -- Vérification des 10 premières lignes de FactCampaignPerformance
97    SELECT TOP 10 * FROM dbo.FactCampaignPerformance;

```

FactKey	CompanyKey	CampaignKey	AudienceKey	SegmentKey	LocationKey	DateKey	ConversionRate	AcquisitionCost	ROI	Clicks	Impressions	EngagementScore
981	1	1	1	1	1	20210101	0.0400	16174.00	6.2000	398	1022	6
982	2	1	2	2	2	20210102	0.1200	11568.00	5.6100	118	7523	7
983	3	3	3	3	3	20210103	0.0700	10200.00	7.1800	564	7898	1
984	4	4	4	1	4	20210104	0.1100	12724.00	5.5500	217	1820	7
985	2	1	3	1	3	20210105	0.0500	16452.00	6.5000	379	4201	3
986	4	4	4	6	2	20210106	0.0700	9716.00	4.3600	100	1843	1
987	2	1	2	7	3	20210107	0.0300	11800.00	2.8000	20	8749	10
988	4	8	1	3	3	20210108	0.0800	13280.00	5.5500	824	7654	7
989	3	9	2	7	1	20210109	0.0900	18698.00	6.7300	881	1754	6
990	10	1	2	7	3	20210110	0.0900	13798.00	3.7800	842	3858	3

Execution de requête réussie.

Cela permet de vérifier que la table de faits contient bien les 10 premiers enregistrements et que les données sont correctement chargées.

```

114
115    -- Vérification de la relation entre FactCampaignPerformance et DimCampaign
116    SELECT f.FactKey, f.CampaignKey, d.CampaignKey
117        FROM dbo.FactCampaignPerformance f
118            LEFT JOIN dbo.DimCampaign d ON f.CampaignKey = d.CampaignKey
119            WHERE d.CampaignKey IS NULL;
120
121    -- Vérification de la relation entre FactCampaignPerformance et DimCompany
122    SELECT f.FactKey, f.CompanyKey, c.CompanyKey
123        FROM dbo.FactCampaignPerformance f
124            LEFT JOIN dbo.DimCompany c ON f.CompanyKey = c.CompanyKey
125            WHERE c.CompanyKey IS NULL;
126
127    -- Vérification de la relation entre FactCampaignPerformance et DimCustomerSegment
128    SELECT f.FactKey, f.SegmentKey, s.SegmentKey
129        FROM dbo.FactCampaignPerformance f
130            LEFT JOIN dbo.DimCustomerSegment s ON f.SegmentKey = s.SegmentKey
131            WHERE s.SegmentKey IS NULL;
132
133    -- Vérification de la relation entre FactCampaignPerformance et DimLocation
134    SELECT f.FactKey, f.LocationKey, l.LocationKey
135        FROM dbo.FactCampaignPerformance f
136            LEFT JOIN dbo.DimLocation l ON f.LocationKey = l.LocationKey
137            WHERE l.LocationKey IS NULL;
138
139    -- Vérification de la relation entre FactCampaignPerformance et DimTargetAudience
140    SELECT f.FactKey, f.AudienceKey, a.AudienceKey
141        FROM dbo.FactCampaignPerformance f
142            LEFT JOIN dbo.DimTargetAudience a ON f.AudienceKey = a.AudienceKey
143            WHERE a.AudienceKey IS NULL;
144
145    -- Vérification de la relation entre FactCampaignPerformance et DimDate
146    SELECT f.FactKey, f.DateKey, d.DateKey
147        FROM dbo.FactCampaignPerformance f
148            LEFT JOIN dbo.DimDate d ON f.DateKey = d.DateKey
149            WHERE d.DateKey IS NULL;

```

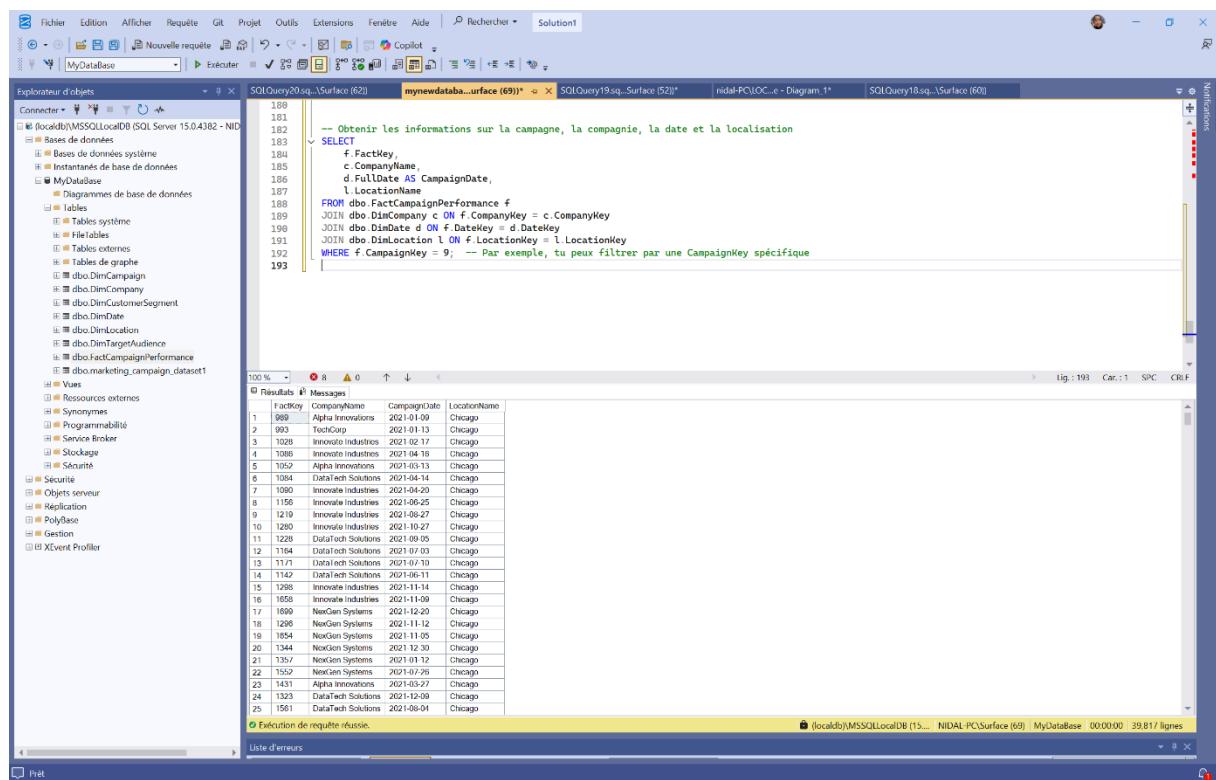
Execution de requête réussie.

Les résultats montrent que la relation entre les clés primaires et les clés étrangères est correcte, et il n'y a pas de valeurs manquantes ou incohérentes dans la table de faits par rapport aux tables de dimensions. Cela garantit l'intégrité des données et confirme que toutes les relations entre les tables sont correctement établies.

Requêtes de Validation Générales pour Chaque Table

Vérification des Données sur la Compagnie, la Date, et la Localisation

Obtenir les noms des compagnies, dates et locations (jointure entre FactCampaignPerformance, DimCompany, DimDate et DimLocation) :



```

Fichier Édition Afficher Requête Git Projet Outils Fenêtre Aide | Rechercher | Solution1
Nouvelle requête Exécuter Copier
MyDataBase | mynewdatabase...surface (69)* | SQLQuery19.sql...Surface (52)* | nidal-PC\OC...c - Diagram_1* | SQLQuery18.sql...Surface (60)
Explorateur d'objets | SQLQuery20.sql...Surface (62) |
Connexion à (localdb)\MSSQLLocalDB (SQL Server 15.0.4382 - NIDAL-PC) | Bases de données | Tables | Objets serveur | Liste d'erreurs
Bases de données
  Bases de données système
  Instantanés de base de données
  MyDataBase
    Diagrammes de base de données
    Tables
      Tables système
      Filetables
      Tables externes
      Tables de graphe
      dbo.DimCampaign
      dbo.DimCompany
      dbo.DimCustomerSegment
      dbo.DimDate
      dbo.DimLocation
      dbo.DimTargetAudience
      dbo.FactCampaignPerformance
      dbo.FactCampaign_Performance_dataset1
    Vues
    Ressources externes
    Synonymes
    Programmabilité
    Service Broker
    Stockage
    Sécurité
    Sécurité
    Objets serveur
    Réplication
    Polybase
    Gestion
    XEvent Profiler
Tables
  Tables système
  Filetables
  Tables externes
  Tables de graphe
  dbo.DimCampaign
  dbo.DimCompany
  dbo.DimCustomerSegment
  dbo.DimDate
  dbo.DimLocation
  dbo.DimTargetAudience
  dbo.FactCampaignPerformance
  dbo.FactCampaign_Performance_dataset1
  Vues
  Ressources externes
  Synonymes
  Programmabilité
  Service Broker
  Stockage
  Sécurité
  Sécurité
  Objets serveur
  Réplication
  Polybase
  Gestion
  XEvent Profiler
  Messages
  Résultats
    FactKey CompanyName CampaignDate LocationName
    1 989 Alpha Innovations 2021-01-09 Chicago
    2 991 TechCorp 2021-01-10 Chicago
    3 1000 Innovate Industries 2021-02-17 Chicago
    4 1088 Innovate Industries 2021-04-16 Chicago
    5 1059 Alpha Innovations 2021-03-13 Chicago
    6 1084 DataTech Solutions 2021-04-14 Chicago
    7 1080 Innovate Industries 2021-04-20 Chicago
    8 1158 Innovate Industries 2021-06-25 Chicago
    9 1219 Innovate Industries 2021-08-27 Chicago
    10 1220 Innovate Industries 2021-09-01 Chicago
    11 1228 DataTech Solutions 2021-09-05 Chicago
    12 1164 DataTech Solutions 2021-07-03 Chicago
    13 1171 DataTech Solutions 2021-07-10 Chicago
    14 1142 DataTech Solutions 2021-06-11 Chicago
    15 1298 Innovate Industries 2021-11-14 Chicago
    16 1658 Innovate Industries 2021-11-09 Chicago
    17 1660 NeuGen Systems 2021-11-10 Chicago
    18 1659 NeuGen Systems 2021-11-12 Chicago
    19 1654 NeuGen Systems 2021-11-05 Chicago
    20 1344 NeuGen Systems 2021-12-30 Chicago
    21 1357 NeuGen Systems 2021-01-12 Chicago
    22 1552 NeuGen Systems 2021-07-26 Chicago
    23 1431 Alpha Innovations 2021-03-27 Chicago
    24 1323 DataTech Solutions 2021-12-09 Chicago
    25 1561 DataTech Solutions 2021-08-01 Chicago
  Résultats
  Exécution de requête réussie
  (localdb)\MSSQLLocalDB (15... NIDAL-PC\Surface (69) MyDataBase 00:00:00 39 817 lignes
  Liste d'erreurs

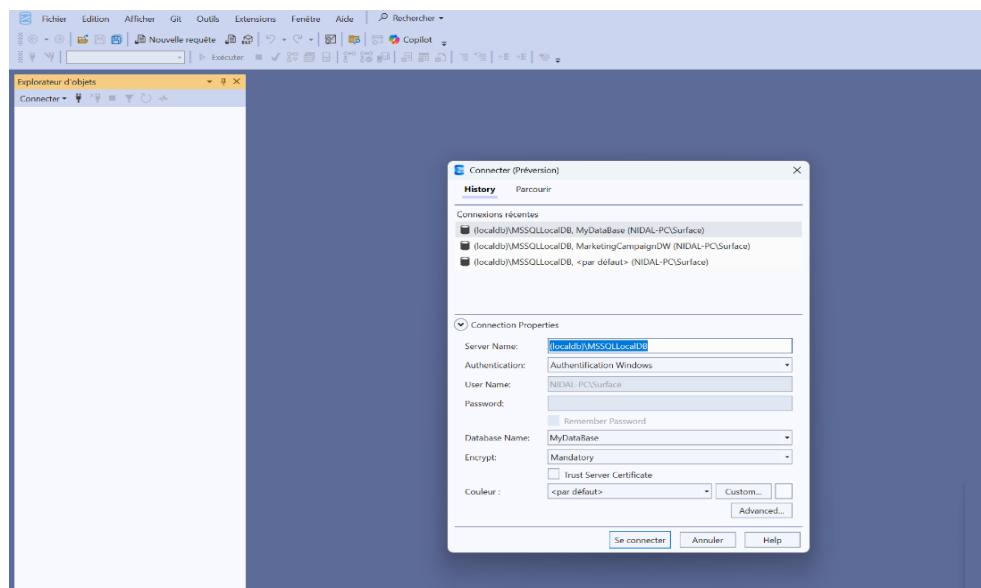
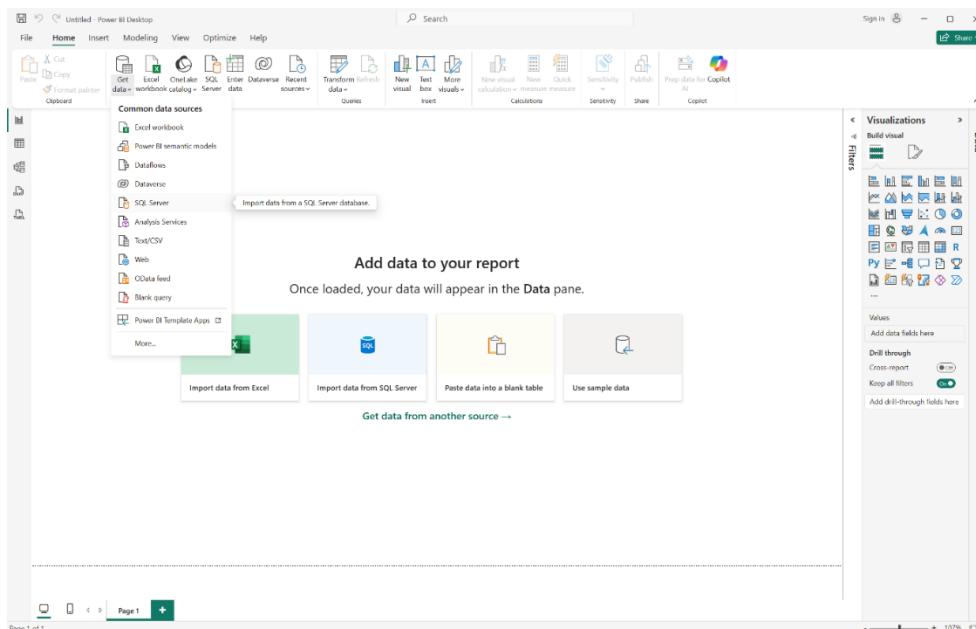
```

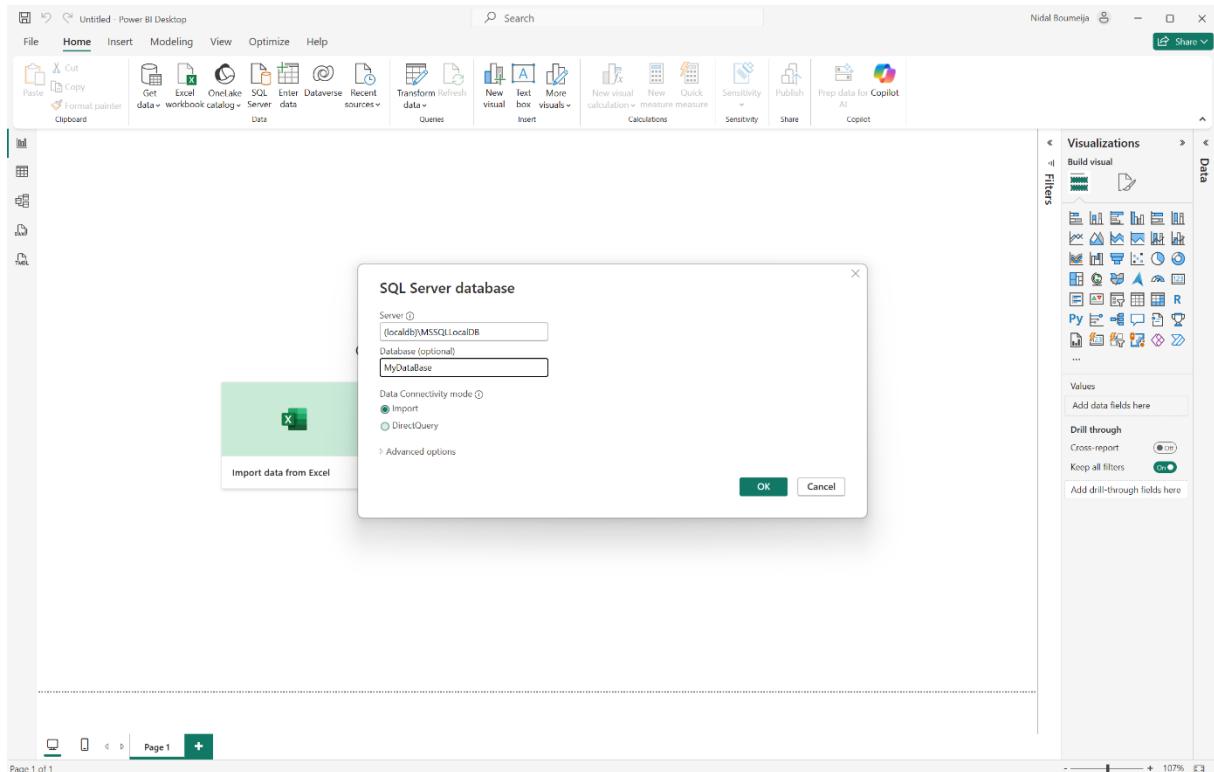
Cette requête te donne des informations détaillées sur chaque FactKey, y compris le nom de la compagnie, la date de la campagne et la localisation associée à chaque enregistrement dans FactCampaignPerformance.

Intégration et visualisation dans Power BI

Cette première étape a pour objectif de préparer l'environnement Power BI et d'importer les jeux de données issus du Data Warehouse (SSIS / SSMS) afin d'obtenir un modèle propre et exploitable. Les opérations réalisées ici garantissent la qualité des données et l'organisation en schéma en étoile nécessaires pour la construction ultérieure des mesures DAX et des visualisations analytiques.

Sources et tables importées





Les tables suivantes sont importées depuis la base SQL Server du projet :

CampaignKey	CampaignID	CampaignType	Duration	ChannelUsed	FactCampaignPerformance
1 1	Email	30 days	Google Ads	Table	
2 2	Email	60 days	Google Ads	Table	
3 3	Influencer	30 days	YouTube	Table	
4 4	Display	60 days	YouTube	Table	
5 5	Email	15 days	YouTube	Table	
6 6	Display	15 days	Instagram	Table	
7 7	Email	60 days	Website	Table	
8 8	Search	45 days	Google Ads	Table	
9 9	Social Media	15 days	Facebook	Table	
10 10	Email	15 days	Instagram	Table	
11 11	Display	45 days	Email	Table	
12 12	Influencer	60 days	Google Ads	Table	
13 13	Social Media	60 days	Facebook	Table	
14 14	Email	45 days	Instagram	Table	
15 15	Display	45 days	Website	Table	
16 16	Social Media	60 days	YouTube	Table	
17 17	Display	45 days	Website	Table	
18 18	Influencer	15 days	Instagram	Table	
19 19	Social Media	60 days	Google Ads	Table	
20 20	Influencer	15 days	Google Ads	Table	
21 21	Search	15 days	Email	Table	
22 22	Influencer	30 days	Facebook	Table	

Transformations de Données

Power Query — nettoyage : pour chaque table appliquer systématiquement :

- Corriger les **types** de colonnes (Date, Decimal, Whole Number, Text).

The screenshot shows the Power Query Editor interface with the 'DimCompany' query selected. In the 'CompanyNames' column, the data type is set to 'Text'. A black oval highlights this 'Text' entry in the 'Data Type' dropdown.

The screenshot shows the Power Query Editor interface with the 'DimDate' query selected. In the 'Date' column, the data type is set to 'Date'. A black oval highlights this 'Date' entry in the 'Data Type' dropdown.

rappt

Transform

Home Add Column View Tools Help

Data Type: Decimal Number ▾ **Decimal Number** **Replace Values** **Unpivot Columns** **Pivot Column** **Convert to List**

Group Use First Row By as Headers Count Rows Table

Queries [7]

DimCampaign DimCompany DimCustomerSegment DimDate DimLocation DimTargetAudience FactCampaignPerformance

19 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON WEDNESDAY

Py Run R script Run Python script Scripts

Properties Name FactCampaignPerformance All Properties

Applied Steps Source Navigation Reordered Columns

ConversionRate DateKey AcquisitionCost ROI Clicks Impressions EngagementsScore

1 20210101 0.04 18174 6.29 508 1922
2 20210102 0.12 21566 5.61 116 1523
3 20210103 0.07 20200 7.18 588 7698
4 20210104 0.11 22724 5.55 217 1820
5 20210105 0.05 18452 6.5 379 4201
6 20210106 0.07 9716 4.36 200 1643
7 20210107 0.13 11067 2.88 817 8749
8 20210108 0.08 13280 5.55 624 7854
9 20210109 0.09 18088 6.73 882 1754
10 20210110 0.09 13766 3.78 642 3856
11 20210111 0.12 8580 3.49 322 6628
12 20210112 0.05 17502 3.59 659 8948
13 20210113 0.09 17188 4.91 677 8817
14 20210114 0.14 5975 7.06 994 2201
15 20210115 0.08 11398 5.28 482 8470
16 20210116 0.12 9407 2.91 299 1512
17 20210117 0.08 5478 4.53 932 2488
18 20210118 0.14 9485 4.48 218 9264
19 20210119 0.07 19224 6.08 182 5788
20 20210120 0.09 10258 3.83 293 6916
21 20210121 0.09 16580 7.99 975 1561
22 20210122 0.02 12824 6.21 319 8586
23 20210123 0.13 8699 2.12 646 3811
24 20210124 0.04 17608 7.31 764 2157
25 20210125 0.07 8773 6.03 527 1733
26 20210126 0.04 14756 6.23 809 8236
27 20210127 0.09 9182 6.03 993 6916
28 20210128 0.07 6601 3.29 604 2420
29 20210129 0.09 11552 7.12 384 7620
30 20210130 0.09 11608 3.61 952 4055
31 20210131 0.07 13124 2.77 512 4837
32 20210201 0.07 13245 6.83 309 2885
33 20210202 0.06 5796 4.42 367 3218
34 20210203 0.14 15082 2.71 243 5724
35 20210204 0.07 18332 7.81 360 5349
36 20210205 0.09 13256 2.51 250 7373
37

rappt

Transform

Home Add Column View Tools Help

Data Type: Decimal Number ▾ **Whole Number** **Replace Values** **Unpivot Columns** **Pivot Column** **Convert to List**

Group Use First Row By as Headers Count Rows Table

Queries [7]

DimCampaign DimCompany DimCustomerSegment DimDate DimLocation DimTargetAudience FactCampaignPerformance

19 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON WEDNESDAY

Py Run R script Run Python script Scripts

Properties Name FactCampaignPerformance All Properties

Applied Steps Source Navigation Reordered Columns

ConversionRate DateKey AcquisitionCost ROI Clicks Impressions EngagementsScore

1 20210101 0.04 18174 6.29 508 1922
2 20210102 0.12 21566 5.61 116 1523
3 20210103 0.07 20200 7.18 588 7698
4 20210104 0.11 22724 5.55 217 1820
5 20210105 0.05 18452 6.5 379 4201
6 20210106 0.07 9716 4.36 200 1643
7 20210107 0.13 11067 2.88 817 8749
8 20210108 0.08 13280 5.55 624 7854
9 20210109 0.09 18088 6.73 882 1754
10 20210110 0.09 13766 3.78 642 3856
11 20210111 0.12 8580 3.49 322 6628
12 20210112 0.05 17502 3.59 659 8948
13 20210113 0.09 17188 4.91 677 8817
14 20210114 0.14 5975 7.06 994 2201
15 20210115 0.08 11398 5.28 482 8470
16 20210116 0.12 9407 2.91 299 1512
17 20210117 0.08 5478 4.53 932 2488
18 20210118 0.14 9485 4.48 218 9264
19 20210119 0.07 19224 6.08 182 5788
20 20210120 0.09 10258 3.83 293 6916
21 20210121 0.09 16580 7.99 975 1561
22 20210122 0.02 12824 6.21 319 8586
23 20210123 0.13 8699 2.12 646 3811
24 20210124 0.04 17608 7.31 764 2157
25 20210125 0.07 8773 6.03 527 1733
26 20210126 0.04 14756 6.23 809 8236
27 20210127 0.09 9182 6.03 993 6916
28 20210128 0.07 6601 3.29 604 2420
29 20210129 0.09 11552 7.12 384 7620
30 20210130 0.09 11608 3.61 952 4055
31 20210131 0.07 13124 2.77 512 4837
32 20210201 0.07 13245 6.83 309 2885
33 20210202 0.06 5796 4.42 367 3218
34 20210203 0.14 15082 2.71 243 5724
35 20210204 0.07 18332 7.81 360 5349
36 20210205 0.09 13256 2.51 250 7373
37

Supprimer les doublons sur les clés (si présents).

The screenshot shows the Power BI Editor interface. The 'Transform' ribbon tab is selected. A query step named 'Table.Distinct(..., "FactKey")' is visible in the main area. The 'Transform' ribbon has several buttons, and one of them, 'Remove Duplicates', is highlighted with a callout line pointing to it.

Renommer les colonnes pour cohérence (ex. LocationName).

The screenshot shows the Power BI Editor interface. The 'Transform' ribbon tab is selected. A query step named 'Table.TransformColumnTypes(..., type text)' is visible in the main area. The 'Transform' ribbon has several buttons, and one of them, 'Changed Type', is highlighted with a callout line pointing to it in the 'APPLIED STEPS' list.

Power Query — transformations métier :

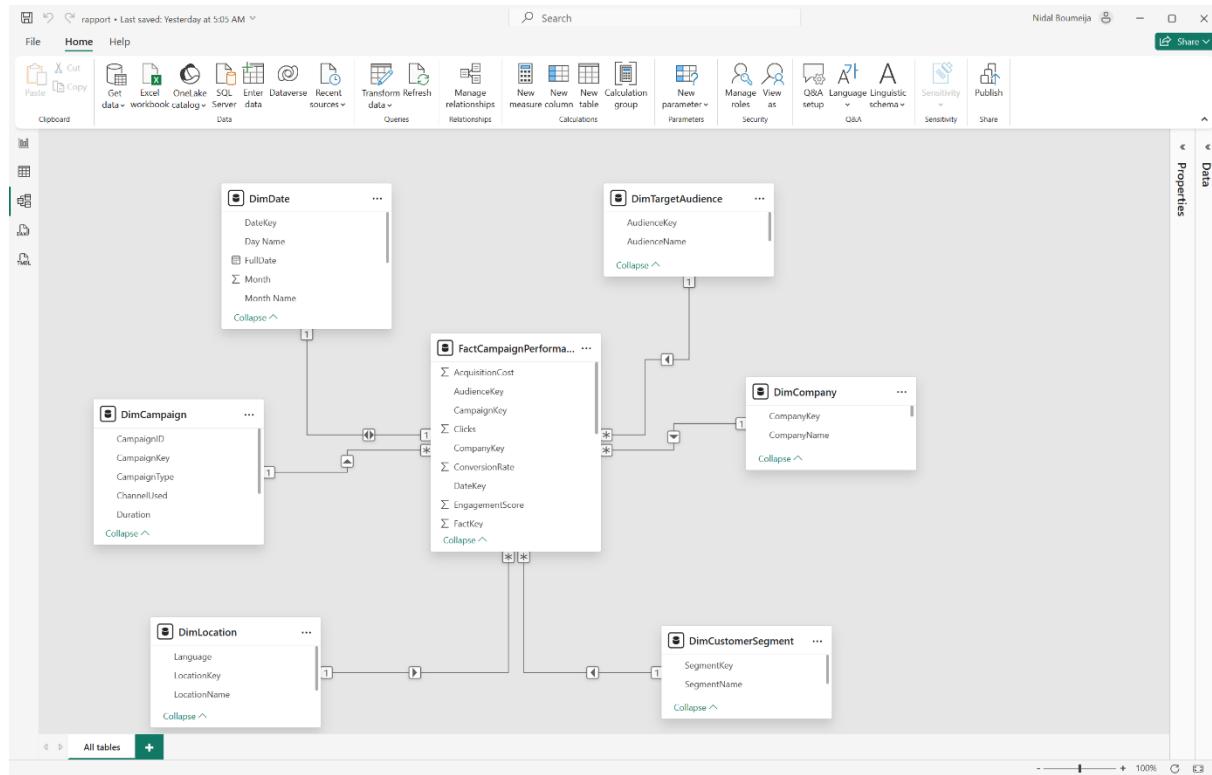
- Dans DimDate : créer MonthNumber (1..12), YearMonth (format YYYY-MM), DayOfWeekNumber et DayName.

The screenshot shows the Microsoft Power Query Editor interface. In the center, there is a table named 'DimDate' with columns: DateKey, FullDate, Month Name, and Day Name. A context menu is open over the 'Month Name' column, with the 'Name of Month' option highlighted by a red circle. The 'APPLIED STEPS' pane on the right shows various steps taken, including 'Reordered Columns'. The 'PROPERTIES' pane shows the name 'DimDate'.

Close & Apply : appliquer les transformations et charger dans le modèle Power BI.

The screenshot shows the Microsoft Power BI Desktop interface. At the top, it says 'rapprt - Last saved: 9/5/2025 at 4:34 AM'. The ribbon has tabs like Home, Insert, Modeling, View, Optimize, Help, and a search bar. On the left, there's a 'Queries' pane with icons for various data sources. The main area shows a 'Load' dialog box with a warning: 'There are pending changes in your queries that haven't been applied.' It lists tables being loaded: DimCampaign, DimCompany, DimCustomerSegment, DimDate, and DimLocation. To the right, there are panes for 'Visualizations' and 'Data'.

Modélisation : vérifier et ajuster les relations pour former un schéma en étoile



Tri des périodes : appliquer **Sort by Column** pour MonthName par MonthNumber et DayName par DayOfWeekNumber afin d'éviter un tri alphabétique incorrect.

The screenshot shows the **DimDate** table in the Power BI Data Model view. The columns are listed as follows:

DateKey	FullDate	Year	Day Name	Month Name	Month
20210101	Friday, January 1, 2021	2021	Friday	January	1
20210102	Saturday, January 2, 2021	2021	Saturday	January	1
20210103	Sunday, January 3, 2021	2021	Sunday	January	1
20210104	Monday, January 4, 2021	2021	Monday	January	1
20210105	Tuesday, January 5, 2021	2021	Tuesday	January	1
20210106	Wednesday, January 6, 2021	2021	Wednesday	January	1
20210107	Thursday, January 7, 2021	2021	Thursday	January	1
20210108	Friday, January 8, 2021	2021	Friday	January	1
20210109	Saturday, January 9, 2021	2021	Saturday	January	1
20210110	Sunday, January 10, 2021	2021	Sunday	January	1
20210111	Monday, January 11, 2021	2021	Monday	January	1
20210112	Tuesday, January 12, 2021	2021	Tuesday	January	1
20210113	Wednesday, January 13, 2021	2021	Wednesday	January	1
20210114	Thursday, January 14, 2021	2021	Thursday	January	1
20210115	Friday, January 15, 2021	2021	Friday	January	1
20210116	Saturday, January 16, 2021	2021	Saturday	January	1
20210117	Sunday, January 17, 2021	2021	Sunday	January	1
20210118	Monday, January 18, 2021	2021	Monday	January	1
20210119	Tuesday, January 19, 2021	2021	Tuesday	January	1
20210120	Wednesday, January 20, 2021	2021	Wednesday	January	1
20210121	Thursday, January 21, 2021	2021	Thursday	January	1
20210122	Friday, January 22, 2021	2021	Friday	January	1
20210123	Saturday, January 23, 2021	2021	Saturday	January	1
20210124	Sunday, January 24, 2021	2021	Sunday	January	1
20210125	Monday, January 25, 2021	2021	Monday	January	1
20210126	Tuesday, January 26, 2021	2021	Tuesday	January	1
20210127	Wednesday, January 27, 2021	2021	Wednesday	January	1
20210128	Thursday, January 28, 2021	2021	Thursday	January	1
20210129	Friday, January 29, 2021	2021	Friday	January	1
20210130	Saturday, January 30, 2021	2021	Saturday	January	1
20210131	Sunday, January 31, 2021	2021	Sunday	January	1
20210201	Monday, February 1, 2021	2021	Monday	February	2
20210202	Tuesday, February 2, 2021	2021	Tuesday	February	2
20210203	Wednesday, February 3, 2021	2021	Wednesday	February	2
20210204	Thursday, February 4, 2021	2021	Thursday	February	2
20210205	Friday, February 5, 2021	2021	Friday	February	2
20210206	Saturday, February 6, 2021	2021	Saturday	February	2
20210207	Sunday, February 7, 2021	2021	Sunday	February	2
20210208	Monday, February 8, 2021	2021	Monday	February	2

The screenshot also shows the Power BI ribbon at the top and the Data Model pane on the right side.

Mesures et KPIs calculés

Avant de développer les mesures DAX et les visualisations dans Power BI, il est essentiel de définir clairement les questions analytiques auxquelles notre projet doit répondre.

Ces questions ne sont pas choisies au hasard : elles représentent les enjeux réels d'une entreprise qui souhaite analyser l'impact de ses campagnes marketing sur les ventes en ligne.

Elles permettent :

- de structurer l'analyse et de s'assurer que chaque mesure créée a un objectif clair,
- de justifier les choix de KPI retenus,
- et de garantir que les visualisations produites répondent à des problématiques de gestion, de pilotage et de prise de décision.

Ainsi, les questions suivantes orienteront notre analyse :

1. Quelle est la performance globale de chaque compagnie ?
(en termes de clics, impressions, ROI et taux de conversion)
2. Quelles sont les compagnies les plus rentables ?
(celles qui affichent le ROI le plus élevé)
3. Quelle est la compagnie avec le coût d'acquisition le plus faible ?
(et inversement, celle qui investit le plus)
4. Quel segment de clients génère le meilleur ROI ?
(identifier les segments les plus porteurs)
5. Quelle audience réagit le mieux aux campagnes marketing ?
(évaluer l'engagement moyen par type d'audience)
6. Quelle localisation apporte le plus de conversions ?
(analyser l'impact géographique des campagnes)
7. Quelle est la tendance des conversions dans le temps (mois/année) ?
(identifier une évolution positive ou négative)
8. Est-ce que le ROI est en amélioration ou en baisse au fil des mois ?
(suivi de la rentabilité dans le temps)
9. Quels sont les meilleurs mois et jours de la semaine pour les performances marketing ?
(optimisation des périodes de campagne)

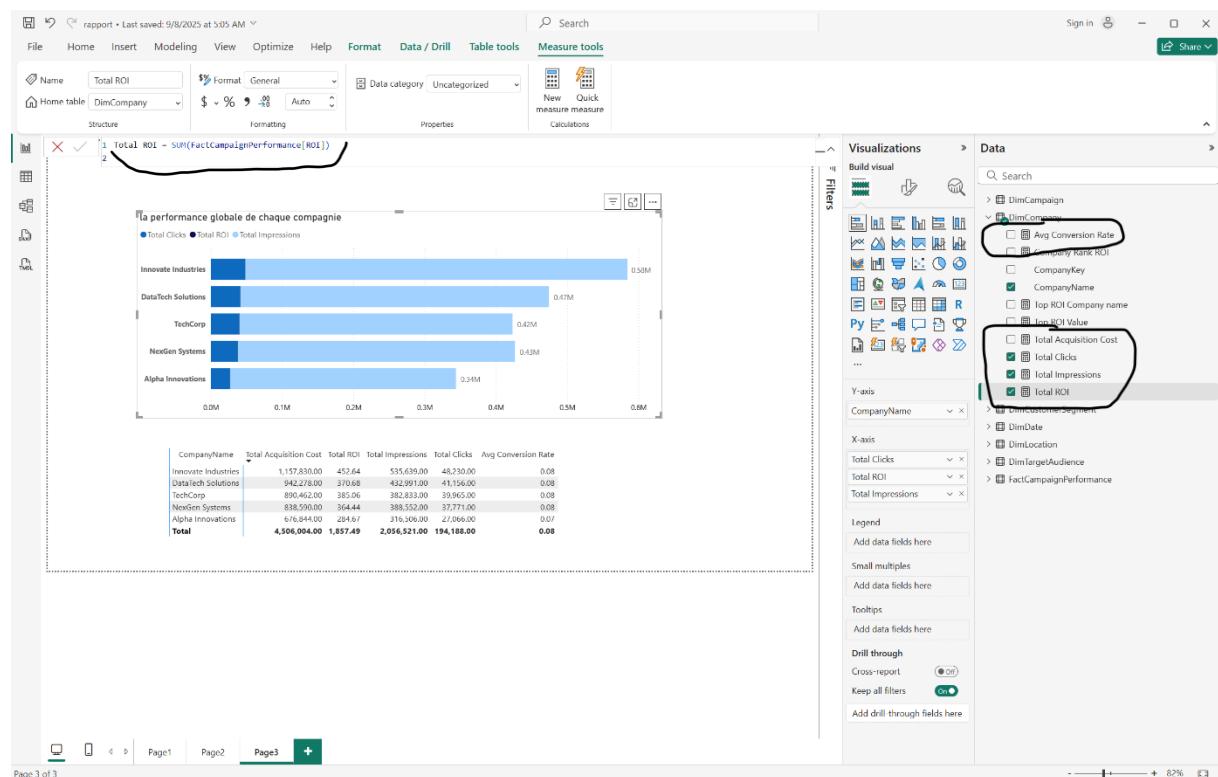
10. Quel est le coût moyen par conversion et que révèle-t-il sur l'efficacité des campagnes ?

11. Quel canal marketing génère le meilleur retour sur investissement (ROI) ?

(comparer Email, Réseaux sociaux, Display, etc.)

Analyse et visualisations Power BI

Question 1 : Quelle est la performance globale de chaque compagnie ?



Justification du KPI

L'objectif de cette analyse est d'évaluer la performance marketing des compagnies dans leur globalité. Pour cela, plusieurs indicateurs ont été définis afin d'avoir une vision équilibrée de l'efficacité des campagnes :

- Total Acquisition Cost : mesure les dépenses engagées pour acquérir des clients.
- Total ROI : indique le retour sur investissement global.
- Total Impressions : reflète la portée et la visibilité des campagnes.
- Total Clicks : mesure l'engagement des utilisateurs.
- Average Conversion Rate : représente la capacité des campagnes à transformer les clicks en conversions.

Ces KPIs sont complémentaires et permettent d'identifier à la fois l'efficacité économique (ROI, Acquisition Cost), la visibilité (Impressions) et l'engagement (Clicks, Conversion Rate).

Visualisations choisies

Deux visualisations complémentaires ont été mises en place : Tableau Matrix et Bar Chart

Cette combinaison (Matrix + Bar Chart) permet d'obtenir à la fois une vue détaillée chiffrée et une représentation visuelle synthétique des résultats.

Question 2 : Quelles sont les compagnies les plus rentables ?

The screenshot shows the Power BI Desktop interface. On the left, the DAX code for the 'Top ROI Value' measure is displayed:

```

1 Top ROI Value =
2 VAR L =
3   TOPN(
4     1,
5     ADDCOLUMNS(
6       VALUES(DimCompany[CompanyName]),
7       "_roi", [total ROI]
8     ),
9     [_roi], DESC
10 )
11 RETURN
12 MAXX(t, [_roi])
  
```

In the center, a card visual titled 'Innovate Industries' displays the value '452.64' under the heading 'Top ROI Value'. Below the card, it says 'Top ROI Company name'.

On the right, the 'Data' pane is open, showing the data model structure. A red circle highlights the 'top ROI Company name' and 'top ROI Value' measures under the 'DimCompany' table. A hand-drawn arrow points from this highlighted area towards the card visual.

Pour identifier les compagnies les plus rentables, nous avons développé deux mesures principales en DAX :

- **Top ROI Company** : cette mesure permet d'isoler la compagnie présentant le meilleur retour sur investissement (ROI).
- **Top ROI Value** : cette mesure calcule la valeur exacte du ROI correspondant à cette compagnie.

Ces deux mesures ont été intégrées dans des **cartes visuelles (Card Visuals)** afin de mettre en évidence, de façon claire et intuitive, la compagnie la plus performante en termes de rentabilité et la valeur de son ROI.

Ce choix de visualisation permet une lecture immédiate des résultats sans nécessiter de comparaison manuelle entre toutes les compagnies. Le lecteur peut ainsi identifier rapidement l'acteur le plus rentable et évaluer l'ampleur de son retour sur investissement.

Question 3 : Quelle est la compagnie avec le coût d'acquisition le plus faible ?

The screenshot shows a Power BI report titled 'rapport'. In the top ribbon, 'Format' is selected. The left pane shows DAX code:

```

1 Lowest Acquisition Company =
2 VAR t =
3   TOPN(
4     1,
5       ADDCOLUMNS(
6         VALUES(DimCompany[CompanyName]),
7         "_cost", [Avg Acquisition Cost]
8       ),
9       [_cost], ASC
10 )
11 RETURN
12 CONCATENATEX(t, DimCompany[CompanyName], ", ")
13

```

The main area displays a card visual with the following data:

- NexGen Systems**
- Lowest Acquisition Company**
- 11.98K**
- Lowest Acquisition Value**

The right pane shows the 'Data' view with filters applied to 'Lowest Acquisition Company' and 'Lowest Acquisition Value'.

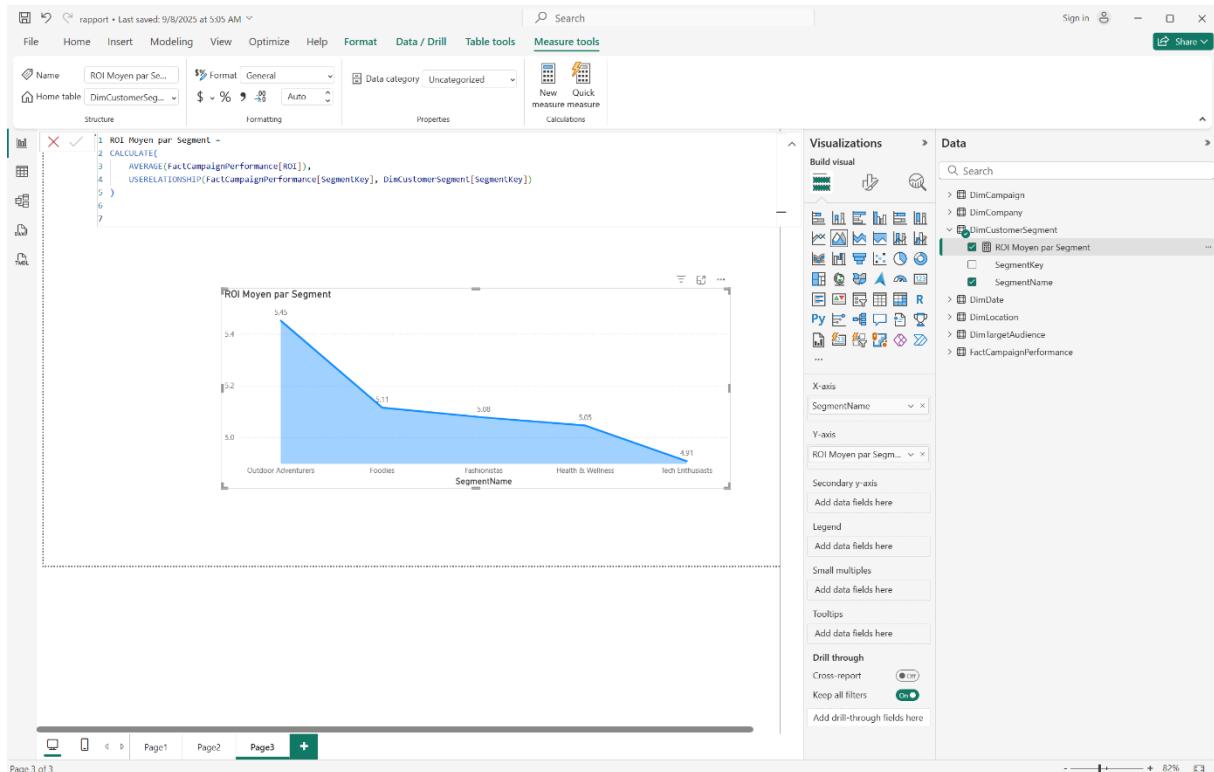
Afin de déterminer quelle compagnie présente le **coût d'acquisition le plus faible**, nous avons défini deux mesures DAX :

- **Lowest Acquisition Company** : identifie la compagnie ayant le coût d'acquisition le plus bas.
- **Lowest Acquisition Value** : calcule la valeur correspondante de ce coût d'acquisition minimal.

Ces mesures ont été représentées à travers des **cartes visuelles (Card Visuals)**. Cette approche permet de mettre en avant, de manière claire et immédiate, la compagnie la plus efficiente en matière de coûts d'acquisition, ainsi que la valeur chiffrée de cette performance.

L'utilisation des cartes assure une lisibilité rapide, ce qui facilite la comparaison avec les résultats précédents liés à la performance globale et à la rentabilité.

Question 4 : Quel segment de clients génère le meilleur ROI ?



Pour analyser l'efficacité des segments de clients, nous avons défini une mesure en DAX :

- **Roi Moyen par Segment** : cette mesure calcule le **retour sur investissement moyen** pour chaque segment de clientèle.

Objectif de cette mesure

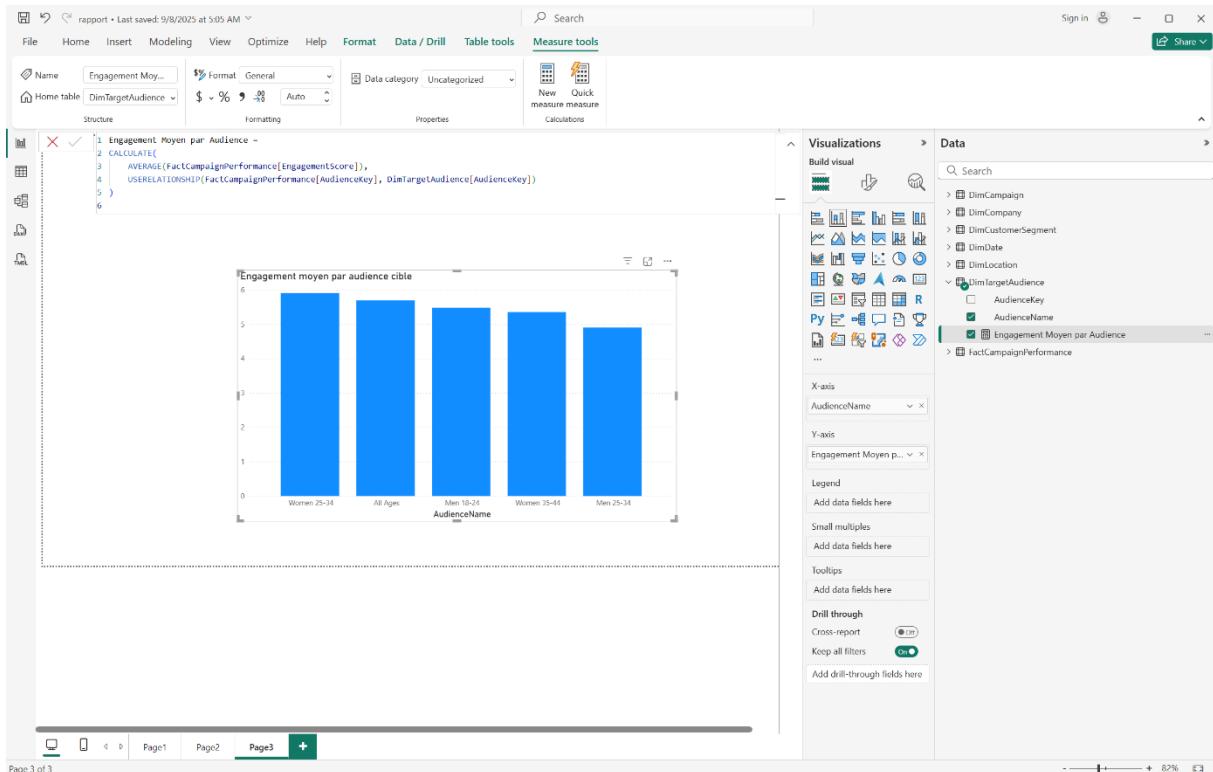
L'objectif est d'identifier le ou les segments de clients qui apportent le meilleur rendement par rapport aux investissements marketing réalisés. En comparant le ROI moyen entre segments, il devient possible d'orienter les efforts marketing vers les audiences les plus rentables.

Visualisation choisie

Les résultats de cette mesure ont été représentés sous forme de **Area Chart** comparant les différents segments. Ce choix permet une comparaison directe et visuelle entre segments, mettant en évidence ceux qui génèrent la meilleure performance.

Cette visualisation fournit une information essentielle pour la prise de décision stratégique : elle aide à déterminer sur quels segments concentrer les ressources afin d'optimiser le retour global sur investissement.

Question 5 : Quelle audience réagit le mieux ?



Pour répondre à cette question, nous avons créé une mesure en DAX permettant d'évaluer la performance des différentes audiences en fonction de leur **taux de réaction** (engagement ou conversion).

- **Average Conversion Rate by Audience** : cette mesure calcule le **taux de conversion moyen** pour chaque audience ciblée.

Objectif de cette mesure

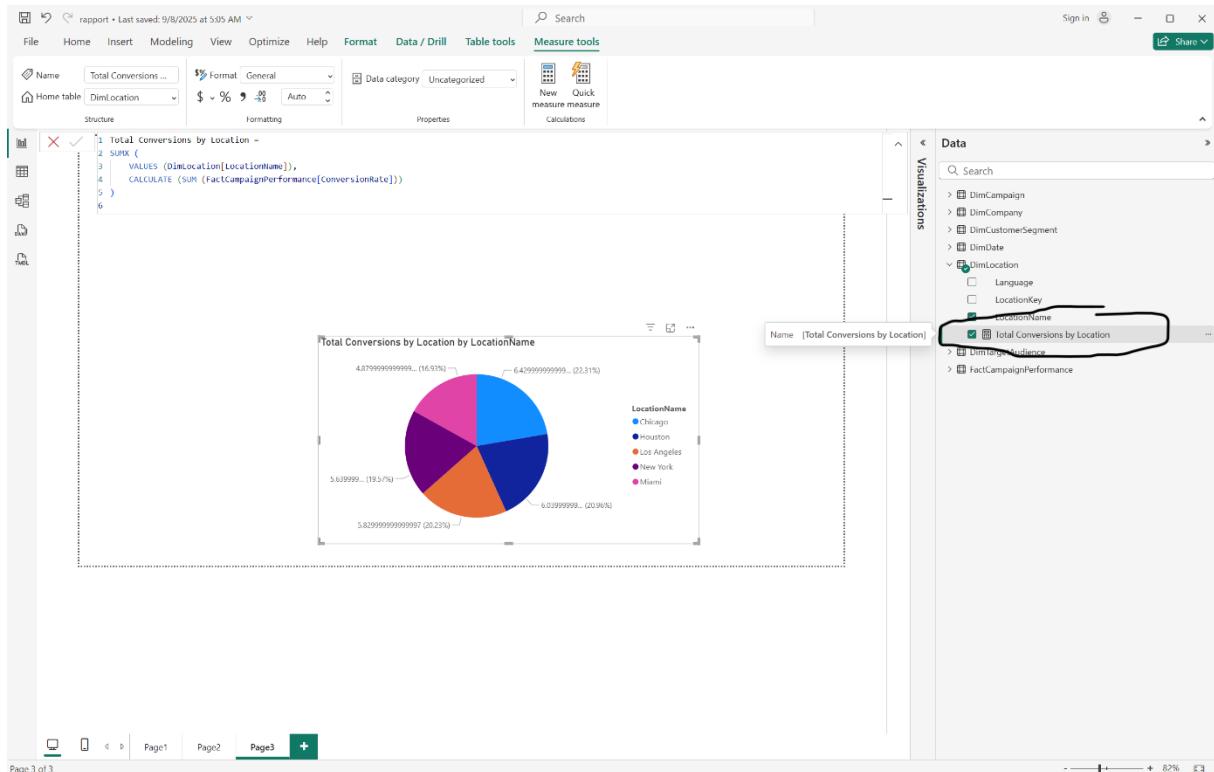
L'objectif est d'identifier quelle audience présente le **meilleur engagement** vis-à-vis des campagnes marketing. En mesurant le taux de conversion moyen, il est possible de comparer l'efficacité relative des différentes audiences et de comprendre où les investissements produisent le plus d'impact.

Visualisation choisie

Les résultats ont été présentés dans un **column Chart** permettant de comparer visuellement les audiences. Ce type de graphique met clairement en évidence les différences de performance et permet de repérer rapidement l'audience la plus réceptive.

Cette analyse constitue un levier stratégique pour l'optimisation des campagnes, car elle indique où concentrer les efforts pour maximiser l'impact des actions marketing.

Question 6 : Quelle localisation apporte le plus de conversions ?



Objectif du KPI :

L'objectif de cette mesure est d'identifier les zones géographiques (locations) qui génèrent le plus de conversions. Cet indicateur est essentiel pour comprendre où les campagnes marketing performent le mieux et permet d'orienter les investissements futurs vers les marchés les plus porteurs.

Choix des mesures :

Pour répondre à cette question, nous avons utilisé la mesure **Total Conversions**, qui représente le nombre total de conversions enregistrées par chaque localisation.

Visualisation utilisée :

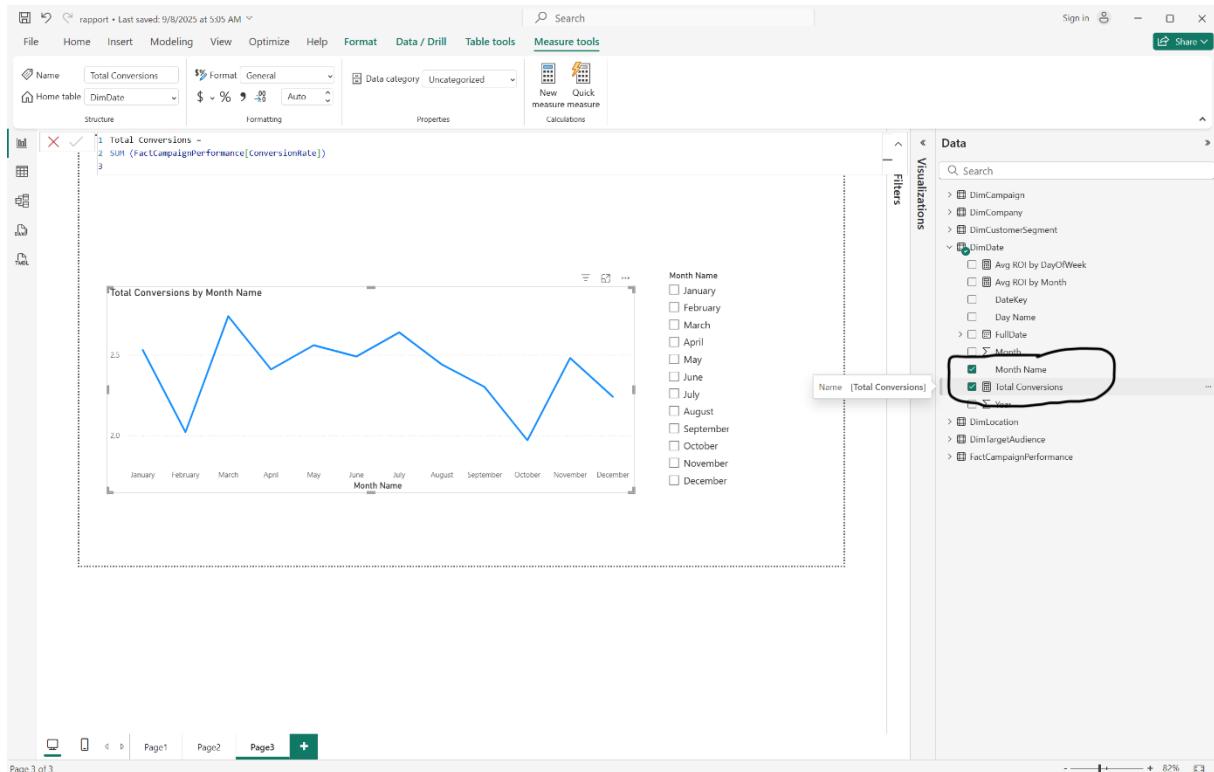
Afin de représenter cette répartition de manière claire et visuelle, un **Pie Chart** a été utilisé. Ce graphique permet de comparer les différentes localisations et de voir rapidement celles qui contribuent le plus aux conversions globales.

- Chaque part du graphique correspond à une localisation.
- La taille de la part reflète le volume de conversions réalisées.
- Les pourcentages affichés donnent un aperçu immédiat du poids relatif de chaque localisation dans la performance globale.

Analyse obtenue :

Grâce à cette visualisation, il est possible d'identifier les localisations les plus performantes et de comparer leurs résultats. Cela permet de cibler plus efficacement les investissements marketing par zone géographique et d'ajuster la stratégie en fonction de la performance constatée.

Question 7 : Quelle est la tendance des conversions dans le temps ?



Objectif du KPI :

L'objectif de cette analyse est de comprendre l'évolution du nombre de conversions dans le temps. Suivre la tendance mensuelle permet de détecter des variations saisonnières, d'identifier des périodes de forte ou faible performance, et d'évaluer l'efficacité des campagnes sur la durée.

Choix des mesures :

Nous avons utilisé la mesure **Total Conversions**, représentée en fonction de la dimension temporelle (**Month Name**).

Visualisation utilisée :

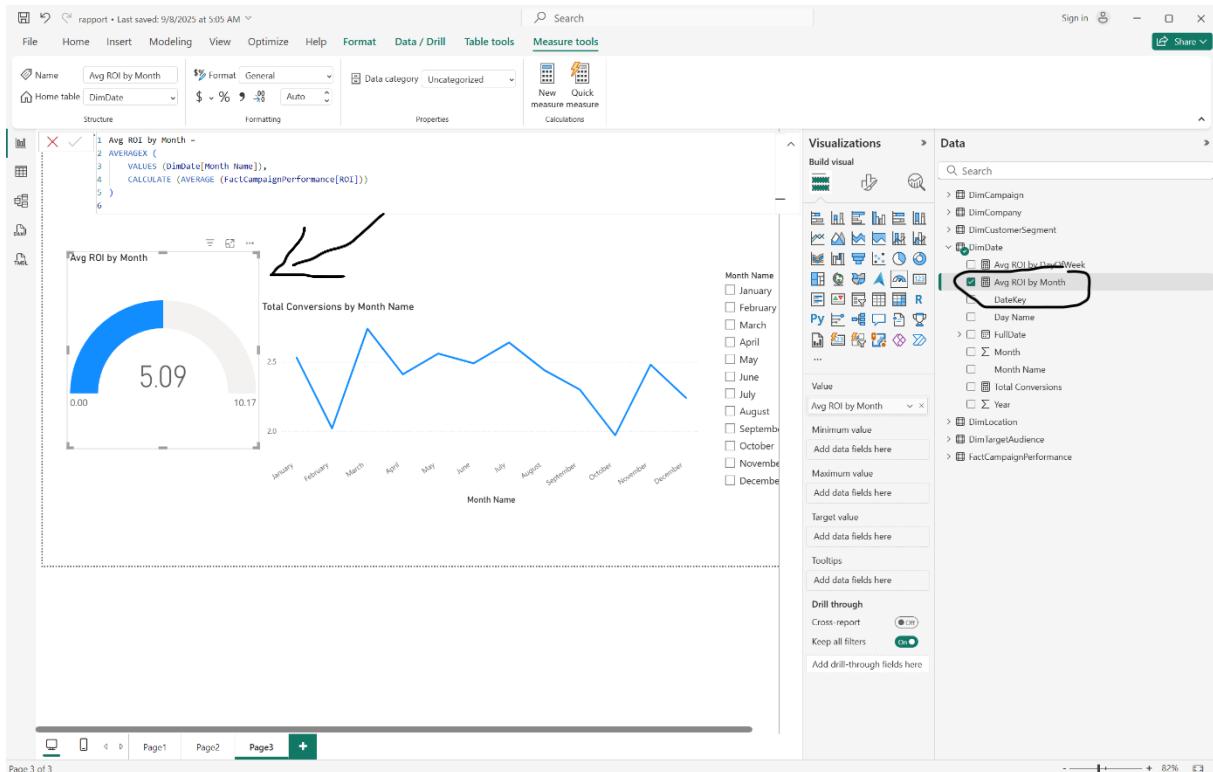
Un **Line Chart (Graphique en courbes)** a été choisi afin de représenter la tendance des conversions au fil des mois.

- L'axe horizontal (X) correspond aux **mois de l'année**.
- L'axe vertical (Y) représente le **nombre total de conversions**.
- La courbe montre clairement les variations d'un mois à l'autre et facilite la détection de pics ou de baisses.

Analyse obtenue :

Cette visualisation met en évidence la tendance générale des conversions au cours de l'année. Elle permet de repérer les mois où les campagnes ont eu le plus de succès et ceux qui nécessitent une amélioration. Les pics observés indiquent les périodes où les actions marketing ont été les plus efficaces, tandis que les baisses signalent un besoin d'optimisation.

Question 8 : Est-ce que le ROI est en amélioration ou en baisse au fil des mois ?



Objectif du KPI :

L'objectif est d'évaluer la tendance du **Retour sur Investissement (ROI)** au fil du temps. Cette analyse permet de déterminer si les campagnes marketing gagnent en efficacité ou, au contraire, si leur rentabilité diminue au cours de l'année.

Choix des mesures :

Nous avons utilisé la mesure **Average ROI**, calculée par mois. Cette mesure permet d'obtenir une vision claire et standardisée de la performance moyenne mensuelle, indépendamment des volumes.

Visualisation utilisée :

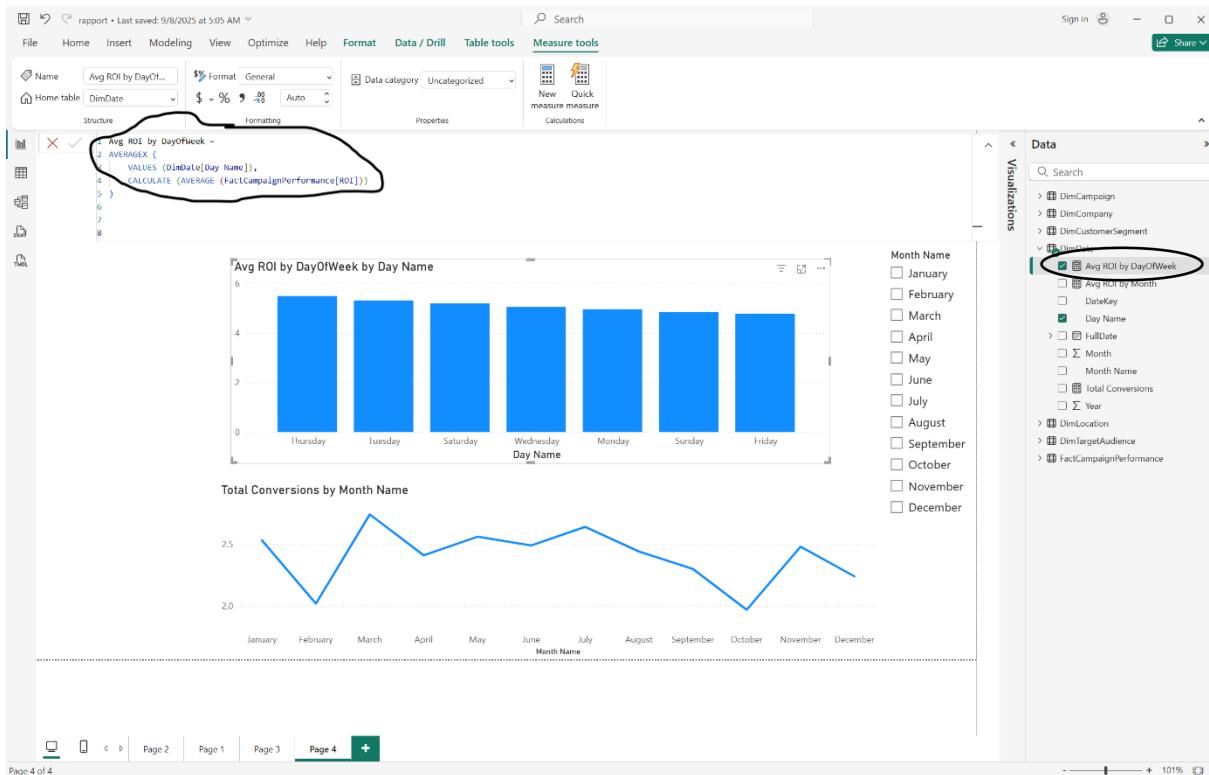
Un **Gauge Chart (Jauge)** et un **Line Chart (Graphique en courbes)** ont été utilisés conjointement :

- Le **Gauge Chart** présente la valeur moyenne du ROI de manière globale, facilitant une lecture rapide du niveau atteint.
- Le **Line Chart** montre la tendance du ROI par **Month Name**, permettant de comparer les mois et d'identifier les hausses et les baisses.

Analyse obtenue :

La combinaison de ces deux visualisations permet d'évaluer la rentabilité globale et son évolution mensuelle. On peut ainsi voir si le ROI suit une progression positive, reste stable ou connaît des périodes de baisse. Cette information est essentielle pour ajuster les stratégies marketing, renforcer les canaux les plus performants et corriger ceux qui génèrent une rentabilité plus faible.

Question 9 : Quels sont les meilleurs mois et jours de la semaine pour les performances marketing ?



Objectif du KPI :

L'objectif est d'identifier les périodes les plus performantes afin d'optimiser la planification et la diffusion des campagnes marketing. Connaître les **meilleurs mois** et les **jours les plus rentables** permet d'orienter les efforts publicitaires vers les moments offrant le plus de conversions et de ROI.

Choix des mesures :

Deux mesures principales ont été utilisées :

- **Total Conversions by Month Name**, pour analyser le volume de conversions par mois.
- **Average ROI by Day Name**, pour comparer le retour sur investissement moyen selon le jour de la semaine.

Visualisations utilisées :

1. Bar Chart — Avg ROI by Day Name (positionné en haut)

- Axe X : Day Name (nom du jour)
- Valeur : Avg ROI by DayOfWeek
- Raison : permet de comparer visuellement le ROI moyen obtenu chaque jour de la semaine ; la hauteur des barres identifie immédiatement les jours les plus performants.

2. Line Chart — Total Conversions by Month Name (positionné en bas)

- Axe X : Month Name (mois de l'année) — trié chronologiquement
- Valeur : Total Conversions
- Raison : la courbe montre la dynamique mensuelle des conversions et facilite la détection de pics ou de creux saisonniers.

Analyse obtenue :

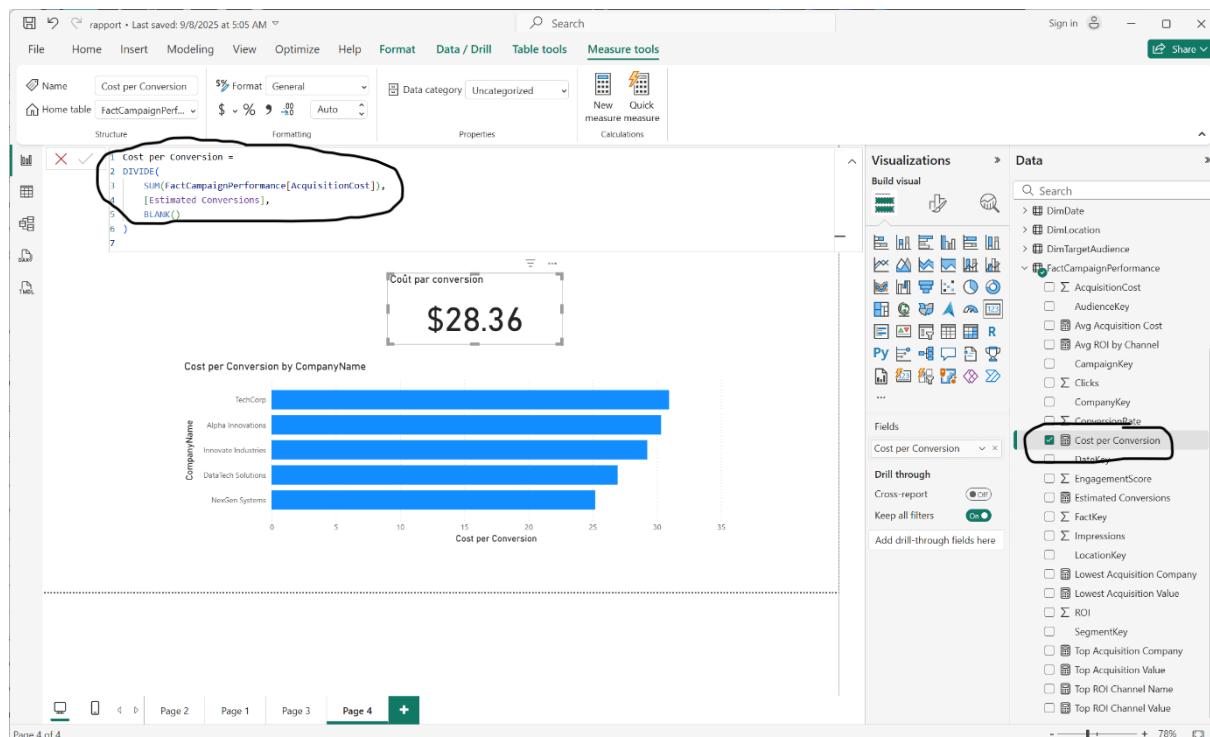
Grâce à cette double approche temporelle (mois + jours), il est possible d'identifier :

- Les mois où les conversions atteignent un pic, révélant les périodes stratégiques pour investir davantage.
- Les jours de la semaine offrant un ROI supérieur, ce qui aide à mieux répartir les campagnes pour maximiser leur efficacité.

Question 10 : Quel est le coût moyen par conversion et que révèle-t-il sur l'efficacité des campagnes ?

Objectif du KPI :

Mesurer combien coûte en moyenne chaque conversion pour évaluer l'efficacité des dépenses marketing.



Cette mesure divise le **coût total d'acquisition** par le **nombre total de conversions**.

Visualisation choisie :

- Une carte KPI affiche la valeur globale : **28.36 \$** par conversion.

- Un **bar chart** présente le coût par conversion détaillé par **CompanyName** (TechCorp, Alpha Innovations, etc.).

Interprétation :

- La valeur globale de **28.36 \$** signifie que, pour générer une conversion, il faut en moyenne investir ce montant.
- Les variations entre compagnies (par ex. TechCorp vs NexGen Systems) permettent d'identifier celles qui gèrent mieux leurs budgets marketing.
- Plus le coût par conversion est bas, plus la campagne est efficace.
- Ici, certaines entreprises parviennent à rester proches de la moyenne tandis que d'autres dépassent légèrement ce seuil, révélant des marges d'amélioration possibles.

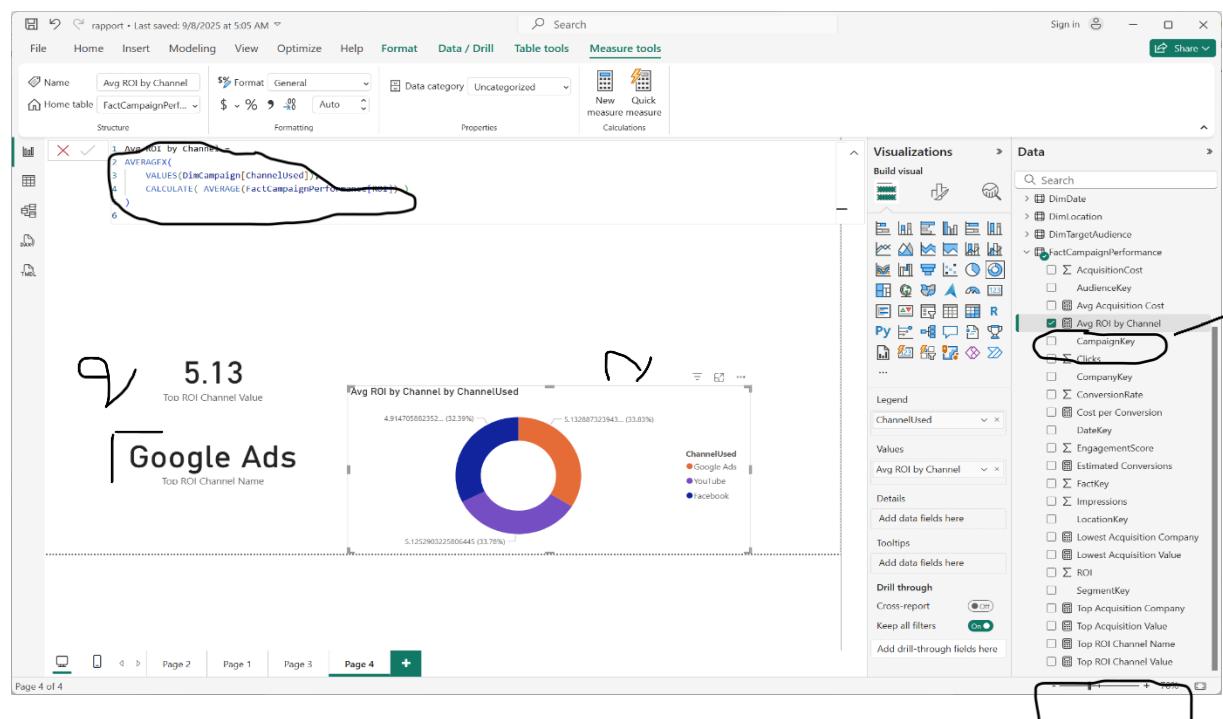
Analyse obtenue :

L'analyse du KPI *Coût par conversion* révèle une moyenne de 28.36 \$, avec des différences notables entre entreprises. Ce résultat permet d'identifier les sociétés les plus performantes en matière de gestion budgétaire et de cibler celles nécessitant une optimisation pour améliorer l'efficacité de leurs campagnes.

Question 11 : Quel canal marketing génère le meilleur retour sur investissement (ROI) ?

Objectif du KPI :

Comparer les canaux marketing (Google Ads, YouTube, Facebook) afin d'identifier celui qui maximise la rentabilité.



Visualisation choisie :

- Une carte **KPI** affiche le **Top ROI Channel Value : 5.13.**
- Une autre carte affiche le **Top ROI Channel Name : Google Ads.**
- Un **donut chart** illustre la comparaison entre les canaux (Google Ads, YouTube, Facebook).

Analyse & Interprétation :

- **Google Ads** ressort comme le canal ayant le ROI le plus élevé (5.13).
- YouTube et Facebook sont proches en performance (autour de 5.12 et 4.91), mais légèrement inférieurs.
- Cela montre que Google Ads est actuellement le canal le plus rentable, mais que les autres canaux restent compétitifs.
- La décision stratégique pourrait être de maintenir l'investissement sur Google Ads tout en optimisant YouTube et Facebook.

L'évaluation des canaux marketing met en évidence que Google Ads génère le meilleur retour sur investissement (5.13), suivi de près par YouTube et Facebook. Cela confirme que Google Ads est le canal le plus rentable, mais que les autres restent des leviers stratégiques intéressants à optimiser.

Visualisations Finales du Dashboard Power BI

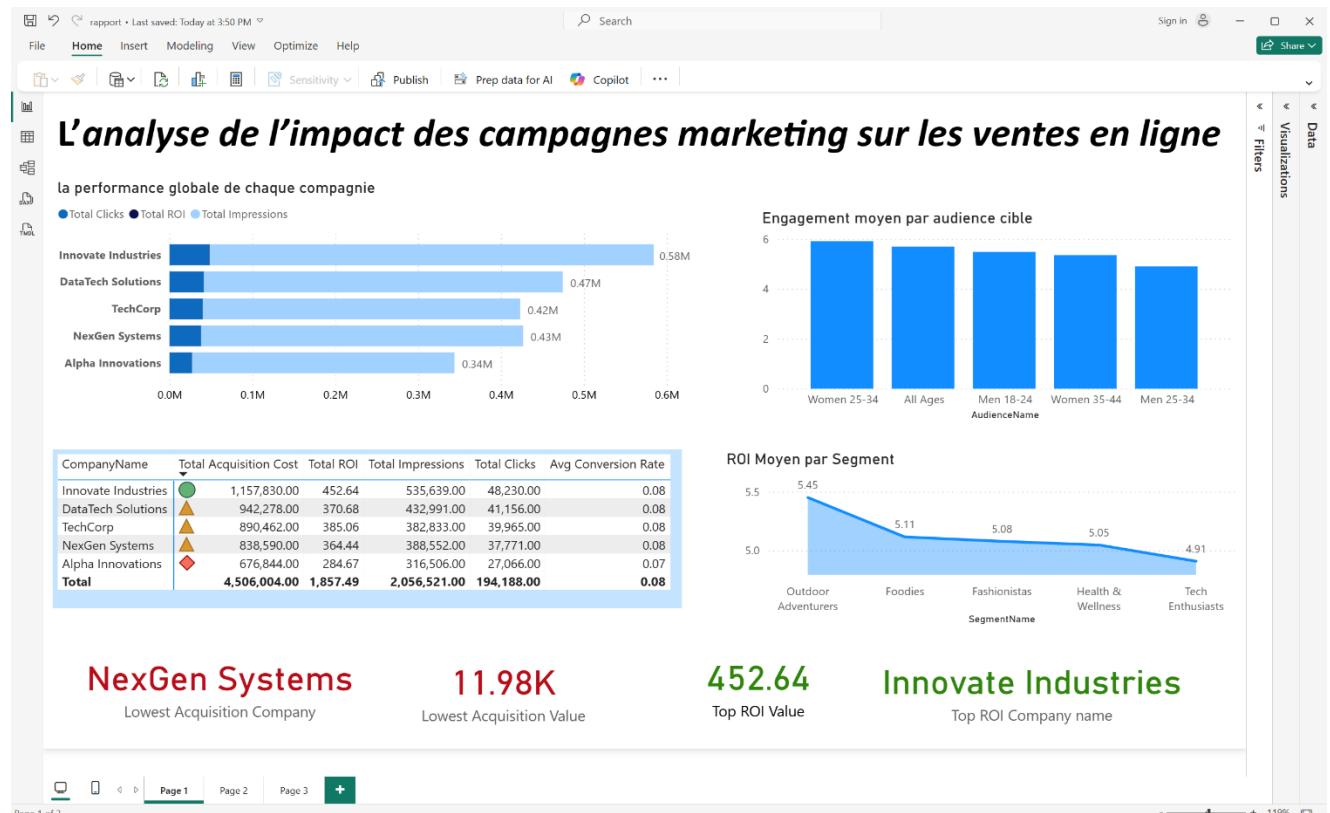
Après avoir construit et analysé individuellement chaque KPI et chaque question de recherche, nous avons consolidé l'ensemble des mesures et visualisations dans des **dashboards complets et interactifs** sous Power BI.

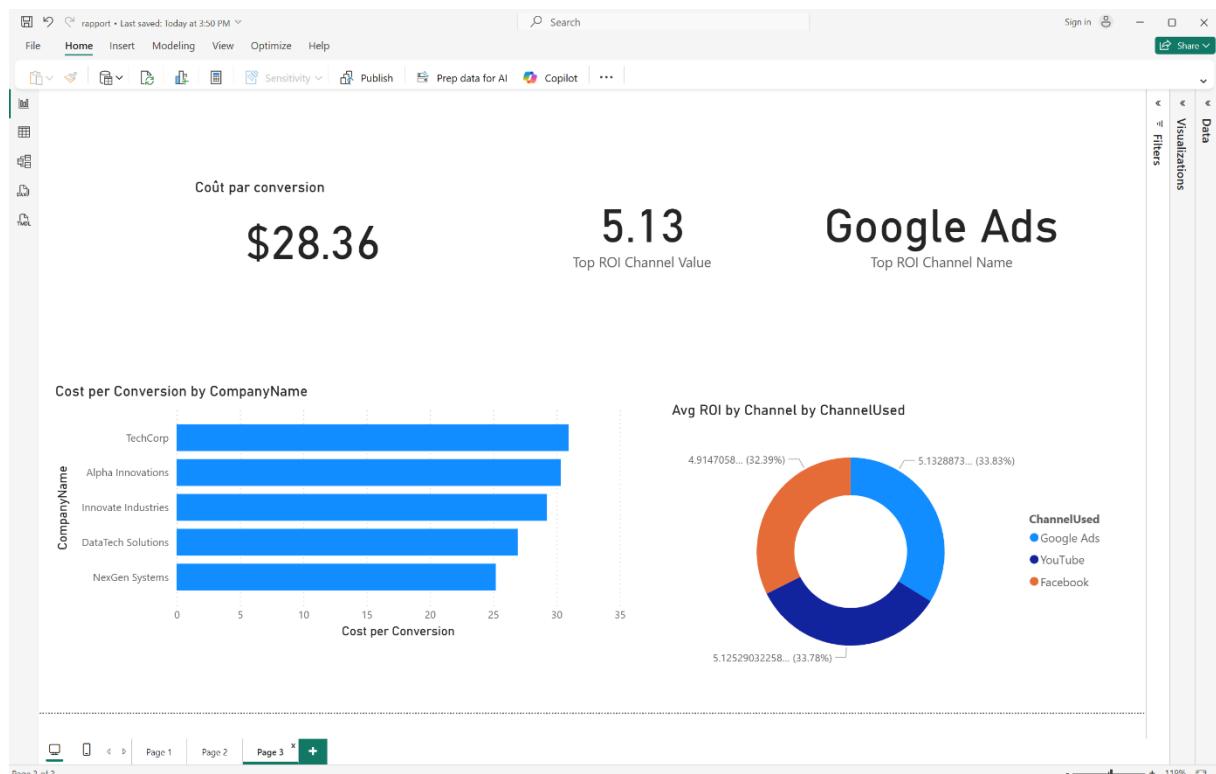
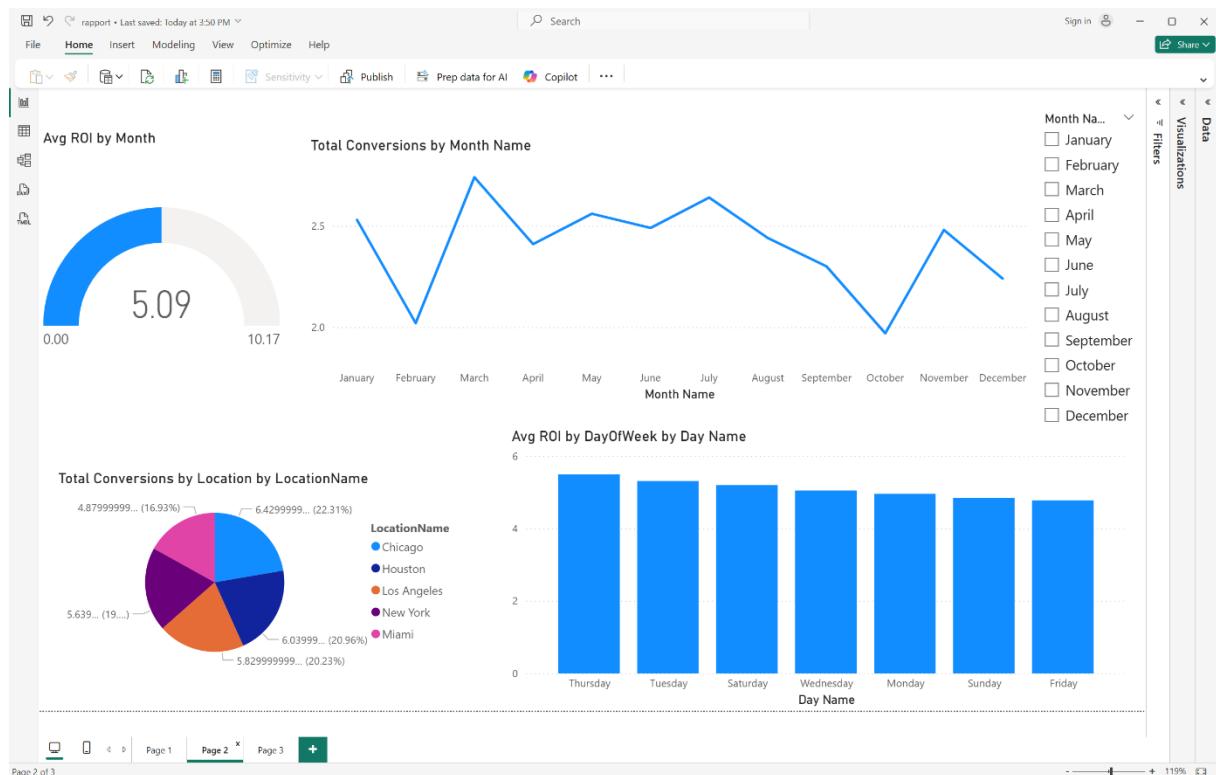
Ces dashboards regroupent :

- Les indicateurs clés de performance (KPI) tels que le ROI, le coût par conversion, les impressions et les clics.
- Des visualisations comparatives (par compagnie, canal marketing, localisation, période).
- Des représentations dynamiques (tableaux, bar charts, line charts, pie charts, donut charts, gauges) permettant une lecture claire et intuitive des résultats.

L'objectif de cette étape est de fournir une **vue globale et synthétique** de la performance marketing, facilitant la prise de décision grâce à une représentation visuelle claire et interactive des données.

Les captures d'écran ci-dessous présentent le résultat final des dashboards, intégrant l'ensemble des visualisations construites tout au long du projet.





Conclusion Générale

Le projet intitulé « **L’analyse de l’impact des compagnies marketing sur les ventes en ligne** » s’inscrit dans le cadre du Master en **Intelligence Artificielle pour l’Économie Numérique**. Il a permis de mettre en pratique des compétences essentielles en ingénierie et analyse de données, en mobilisant un ensemble cohérent d’outils professionnels : **SSIS**, **SSMS** (Data Warehouse) et **Power BI**.

Dans une première étape, l’utilisation de **SSIS** a rendu possible l’extraction, la transformation et le chargement (ETL) des données depuis les sources brutes vers l’entrepôt de données. Ensuite, la modélisation sous **SSMS** a permis de structurer ces données selon une architecture en étoile, facilitant leur exploitation analytique.

Enfin, l’intégration de ces données dans **Power BI** a rendu possible la création de tableaux de bord dynamiques et interactifs, répondant à un ensemble de questions stratégiques définies en amont (Q1 à Q11). Les visualisations conçues — qu’il s’agisse d’indicateurs globaux, de comparaisons par compagnie, de performances par canal marketing, ou encore d’analyses temporelles et géographiques — ont permis d’obtenir une vision claire et exploitable de la performance des campagnes marketing.

Ce projet a ainsi permis de :

- Développer une compréhension pratique du processus complet d’analyse de données, de l’ETL à la visualisation.
- Mettre en évidence les apports d’un entrepôt de données pour consolider et fiabiliser l’information.
- Dégager des **insights pertinents** pour évaluer l’efficacité des investissements marketing et optimiser les décisions futures.

En somme, ce travail a constitué une expérience formatrice, rapprochant la théorie académique des pratiques professionnelles d’un **data analyst** et préparant à l’intégration future dans des environnements réels de l’économie numérique.