

Carttrend

Présentation de Laurent BRUNET du 7 avril 2025

Projet final réalisé en collaboration avec

Yanis BIDAk, Ghazi EL AYDI et Isis ZAKY

Qu'est-ce que Carttrend ?

INTRODUCTION

L'analyse de données du projet se situe dans les quatre domaines suivants :







- Ventes
- Satisfaction client
- Marketing
- Logistique

CONSTAT/PROBLEMATIQUES

- Données massives (Google Sheets, fichiers CSV).
- Manque de clarté sur les performances des produits
- Efficacité marketing limitée
- Saturation logistique
- Insatisfaction client croissante
- Mauvaise anticipation des tendances

Extraction de KPIs potentiels

Extraction de KPIs potentiels

 Carttrend_Campaigns
 Carttrend_Clients
 Carttrend_Commandes
 Carttrend_Details_Commandes
 Carttrend_Entrepots
 Carttrend_Entrepots_Machines
 Carttrend_Posts
 Carttrend_Produits
 Carttrend_Promotions
 Carttrend_Satisfaction

Carttrend

Extraction de KPIs potentiels

Les fichiers peuvent
compter jusqu'à 80 000
lignes de données (le
fichier Commandes).

Carttrend_Promotions

Fichier Édition Affichage Insertion Format Données Outils Extensions Aide

100% Lecture seule

A1 id_promotion

	A	B	C	D	E	F	G
1	id_promotion	id_produit	type_promotion	valeur_promotion	date_début	date_fin	responsable_promotion
2	PROM001	P00429	Pourcentage	5%	2024-12-02	2024-12-15	Responsable Marketing
3	PROM002	P00380	Pourcentage	48%	2024-02-26	2024-02-25	Responsable Marketing
4	PROM003	P00899	Pourcentage	23%	2024-03-11	2024-04-05	Responsable Marketing
5	PROM004	P00243	Remise fixe	€ 8,00	2024-06-17	2024-07-11	Responsable Marketing
6	PROM005	P00953	Remise fixe	€ 12,00	2023-12-13	2023-12-27	Responsable Marketing
7	PROM006	P00480	Pourcentage	46%	2024-03-28	2024-04-15	Responsable Marketing
8	PROM007	P00913	Remise fixe	€ 5,00	2024-10-14	2024-10-25	Responsable Marketing
9	PROM008	P00304	Pourcentage	44%	2024-07-17	2024-07-27	Responsable Marketing

Carttrend_Commandes

FichierÉditionAffichageInsertionFormatDonnéesOutilsExtensionsAide

100%Lecture seule

172023-12-14

	A	B	C	D	E	F	G	H	I	
1	id_commande	id_client	id_entrepôt	départ	date_commande	statut_commande	id_promotion_appliquée	mode_de_paiement	numéro_tracking	date_livraison_estimée
2	CMD000001	C02081	E004		2023-12-03	En transit		Virement bancaire	11e09318-c960-4694-a8d0-8c032138c	2023-12-28
3	CMD000002	C06655	E006		2023-12-03	Livrée		Virement bancaire	fa528e98-9177-4665-885f-5b8348b805	2023-12-28
4	CMD000003	C00466	E009		2023-12-03	Annulée		Carte de crédit	c3e61a34-8e8a-45ca-bc47-49e2bef90	2023-12-04
5	CMD000004	C05069	E005		2023-12-03	Livrée		PayPal	badc54cc-134b-490c-bc26-2a744898e	2023-12-16
6	CMD000005	C03517	E003		2023-12-03	Annulée		Virement bancaire	ebfb75e8-478a-43ce-a38e-bd3f9a32c5	2023-12-15
7	CMD000006	C01068	E009		2023-12-03	Validée		Carte de crédit	5b1a0f3d-c1d3-4f30-ab81-641e355b1f	2023-12-14
8	CMD000007	C04333	E005		2023-12-03	Livrée		Carte de crédit	a9b579dd-b2b2-428a-98f9-91d35088c	2023-12-25
9	CMD000008	C08345	E005		2023-12-03	En transit		Carte de crédit	15b90c0e-1a8e-45bb-bd97-65b66eaf5	2023-12-28

Extraction de KPIs potentiels

Quelques exemples :

Fichier concerné		KPI		Intérêt opérationnel
Carttrend_Commandes	→	Nombre total de commandes	→	Evolution chiffre d'affaires
Carttrend_Commandes	→	Taux de commandes livrées à temps	→	Réduction des plaintes
Carttrend_Satisfaction	→	Note des clients	→	Impact réseaux sociaux
Carttrend_Entrepots_Machines	→	Taux de disponibilité des machines (%)	→	Productivité
Carttrend_Promotions	→	Durée moyenne des promotions	→	Effort financier
Carttrend_Clients	→	Nombre total de clients actifs	→	Fidélité clients

OBJECTIF

Quels sont les objectifs du projet ?

- Améliorer les ventes
- Optimiser les campagnes Marketing
- Améliorer la satisfaction client
- Optimiser ses performances logistiques

CHOIX DU PIPELINE

Qu'est-ce qu'un pipeline ?

CHOIX DU PIPELINE

Nous avons le choix entre 2 types de pipelines :

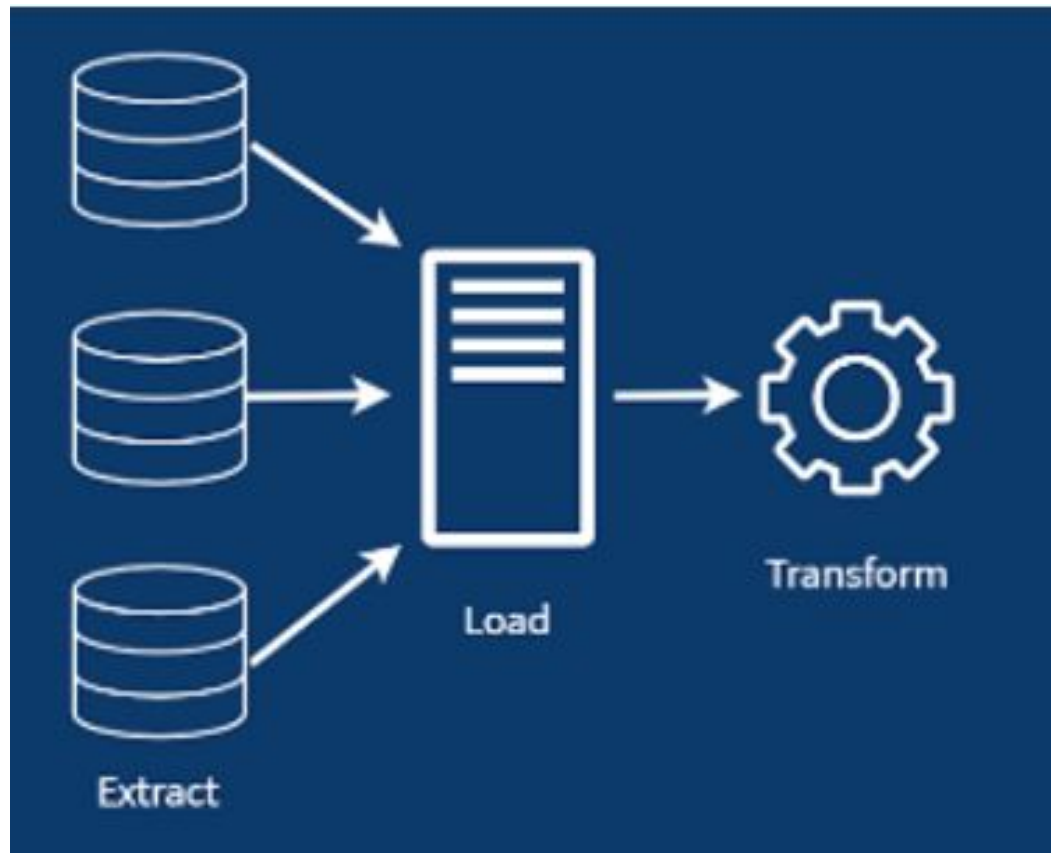
- Pipeline ELT (Extract-Load-Transform)
- Pipeline ETL (Extract-Transform-Load)

CHOIX DU PIPELINE

Les raisons de notre choix pour le pipeline ELT :

- Traitement efficace des données massives.
- Flexibilité et rapidité d'analyse.
- Infrastructure évolutive pour répondre aux besoins futurs.

Synthèse : Le pipeline ELT



CHOIX DU PIPELINE

Les outils principaux sélectionnés pour le pipeline ELT :



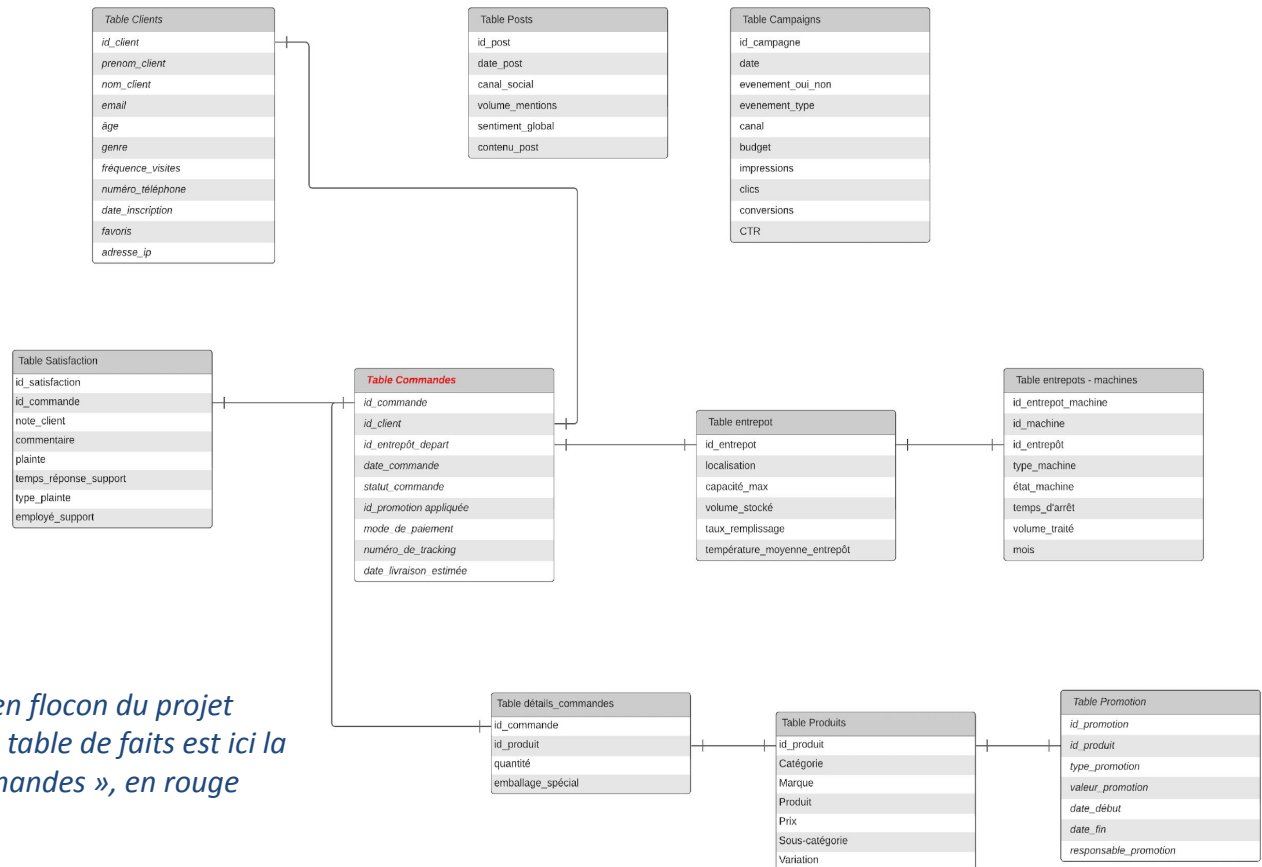
MODELE EN FLOCON

MODELE EN FLOCON



MODELE EN FLOCON

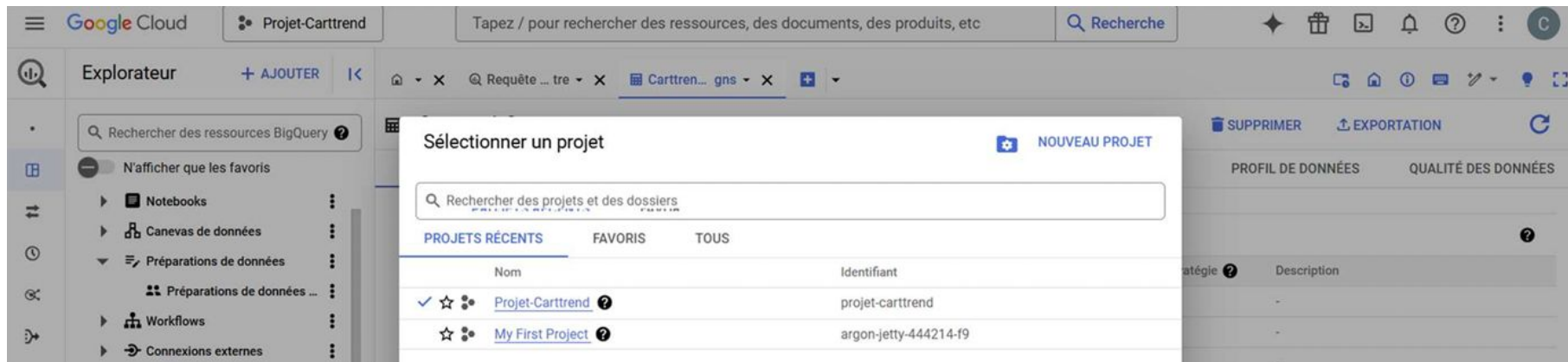
- Schéma de modélisation : étoile, flocon, constellation ?
- Nous avons conclu au modèle en flocon.



CP : Modèle en flocon du projet Carttrend ; la table de faits est ici la table « Commandes », en rouge

ETAPES E-L-T

- Préparation de notre projet sur Google Cloud.
- Création du projet, du dataset et ajout des tables dans BigQuery.





Apache
Airflow



ETAPES E-L-T

carttrend	
analyses	
macros	
models	M
marts	M
analyse_vente	M
logistique	M
marketing	M
satisfaction_clients	M
staging_models	M
schema.yml	A
stg_campaigns_data.sql	
stg_clients_data.sql	
stg_commandes_data.sql	
stg_details_commandes_data.sql	
stg_entrepots_data.sql	
stg_machines_data.sql	
stg_posts_data.sql	
stg_produits.sql	
stg_promotions.sql	
stg_satisfaction_data.sql	
source.yml	

- Un exemple de script SQL sur DBT pour un modèle staging.

```
stg_commandes_data.sql  ×

models / staging_models / stg_commandes_data.sql  ⓘ

1
2  WITH commandes_data AS (
3      -- Sélection des commandes livrées et leurs détails associés
4      SELECT
5          id_commande,
6          `id_client`,
7          `id_entrepôt_départ`,
8          date_commande,
9          statut_commande,
10         `id_promotion_appliquée`,
11         mode_de_paiement,
12         `numéro_tracking`,
13         `date_livraison_estimée`,
14     FROM
15         {{ source('data_carttrend', 'Carttrend_Commandes') }}
16 )
```

- Un exemple de script SQL d'un modèle analytique marts

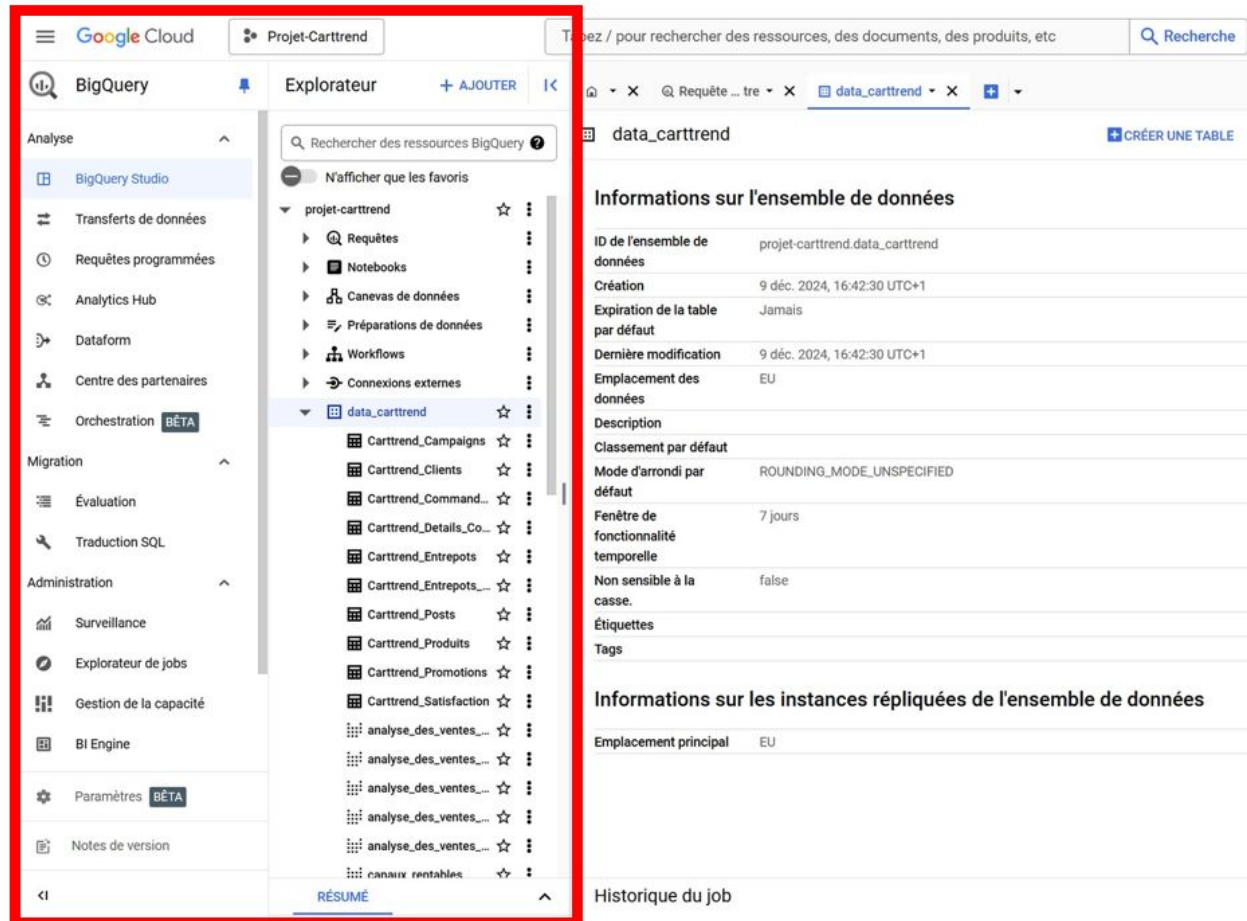
analyse_des_ventes_pro_f... X

models / marts / analyse_vente / analyse_des_ventes_pro_fre.sql



```
1 WITH Paires_Produits AS (  
2     SELECT  
3         dc1.id_produit AS produit_1,  
4         dc2.id_produit AS produit_2,  
5         COUNT(*) AS nombre_occurrences  
6     FROM {{ref('stg_commandes_data')}} c1  
7     JOIN {{ref('stg_commandes_data')}} c2  
8         ON c1.id_client = c2.id_client  
9         AND c1.date_commande = c2.date_commande  
10    JOIN {{ref('stg_details_commandes_data')}} dc1  
11        ON dc1.id_commande = c1.id_commande  
12    JOIN {{ref('stg_details_commandes_data')}} dc2  
13        ON dc1.id_produit < dc2.id_produit -- Évite les doublons symétriques  
14        AND dc2.id_commande = c2.id_commande  
15    GROUP BY dc1.id_produit, dc2.id_produit  
16 )
```

- Vues sur DBT



The screenshot displays the Google Cloud BigQuery interface. On the left, a sidebar lists various BigQuery features under categories like 'Analyse', 'Migration', and 'Administration'. The main area is divided into two panes. The left pane, titled 'Explorateur', shows a hierarchical view of the 'projet-carttrend' project, with the 'data_carttrend' dataset selected. The right pane, titled 'Informations sur l'ensemble de données', provides details about the 'data_carttrend' dataset, including its ID, creation and expiration dates, location, and description. Below this, there is a section for 'Informations sur les instances répliquées de l'ensemble de données'.

Informations sur l'ensemble de données

ID de l'ensemble de données	projet-carttrend.data_carttrend
Création	9 déc. 2024, 16:42:30 UTC+1
Expiration de la table par défaut	Jamais
Dernière modification	9 déc. 2024, 16:42:30 UTC+1
Emplacement des données	EU
Description	
Classement par défaut	
Mode d'arrondi par défaut	ROUNDING_MODE_UNSPECIFIED
Fenêtre de fonctionnalité temporelle	7 jours
Non sensible à la casse.	false
Étiquettes	
Tags	

Informations sur les instances répliquées de l'ensemble de données

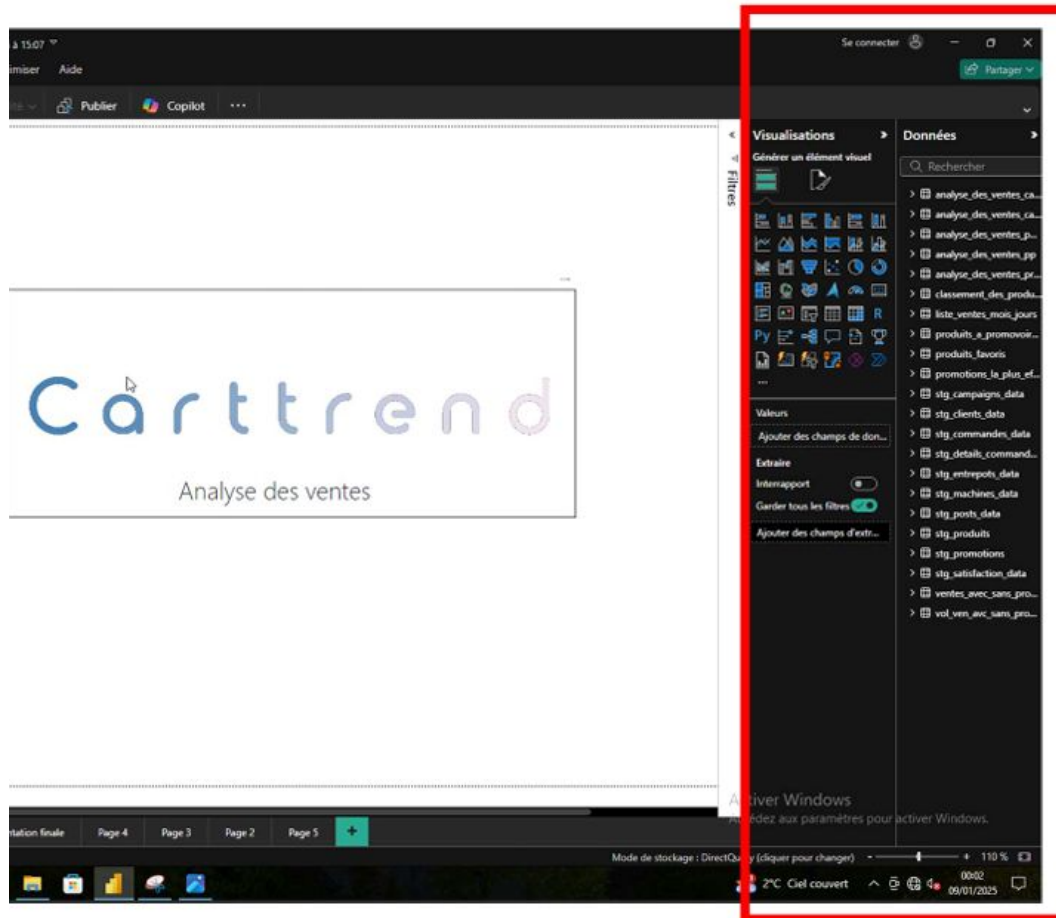
Emplacement principal	EU
-----------------------	----

Historique du job

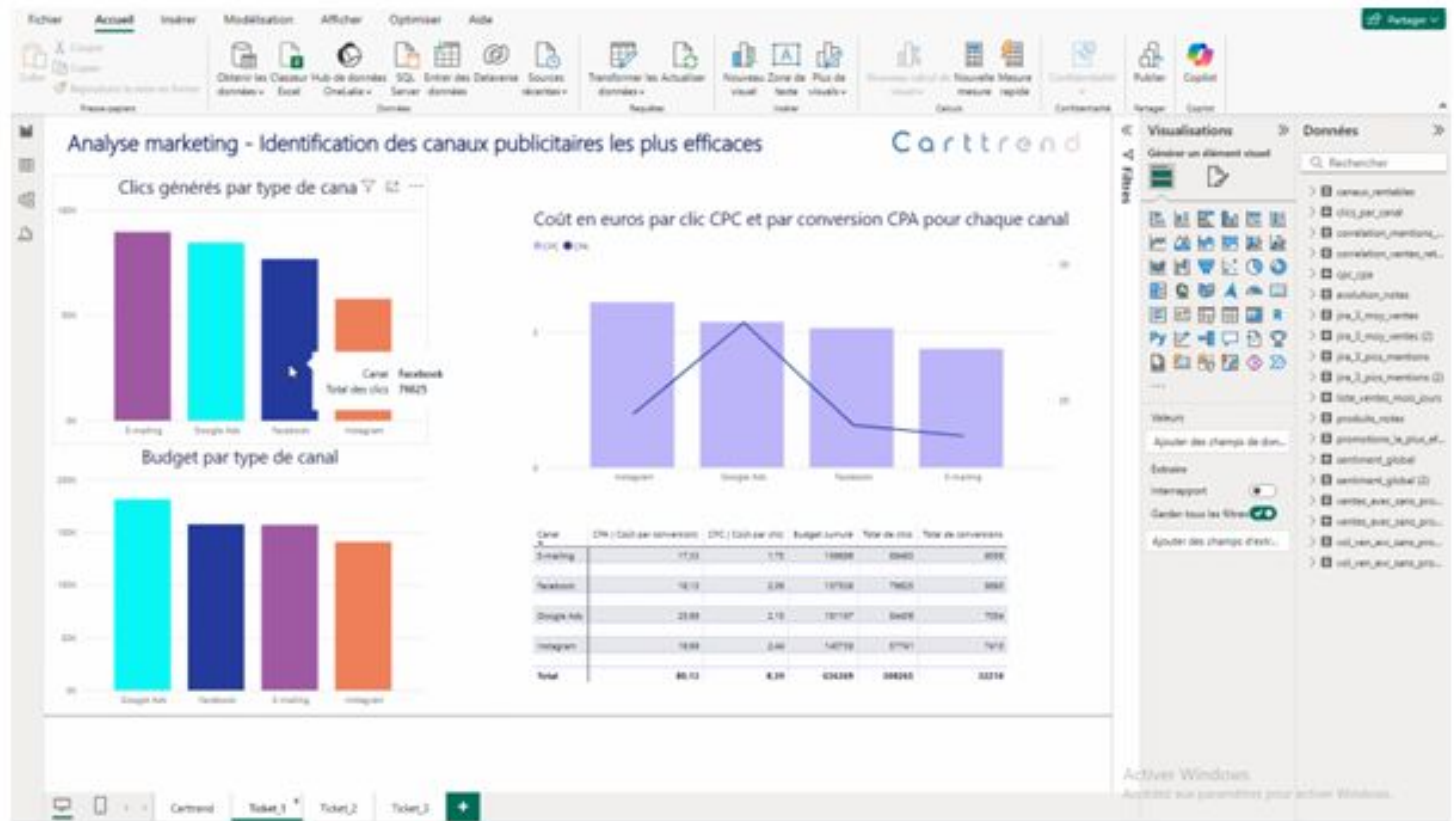
VISUALISATION

VISUALISATION

- Voici une capture d'écran de l'interface de PowerBI avec les onglets de visualisations et de données à droite.

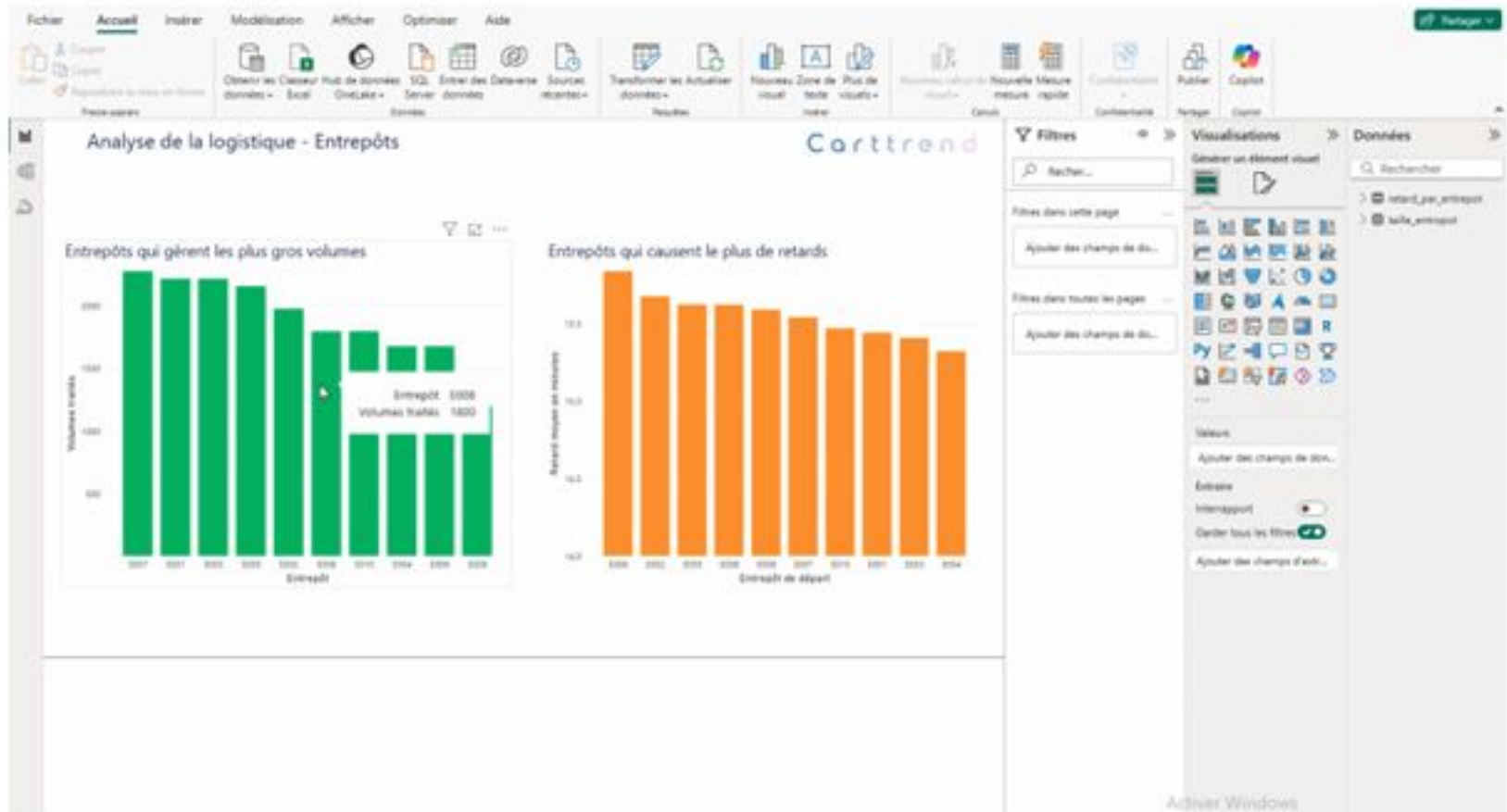


- Mise en place de dashboards interactifs sur PowerBI pour visualiser les insights clés.

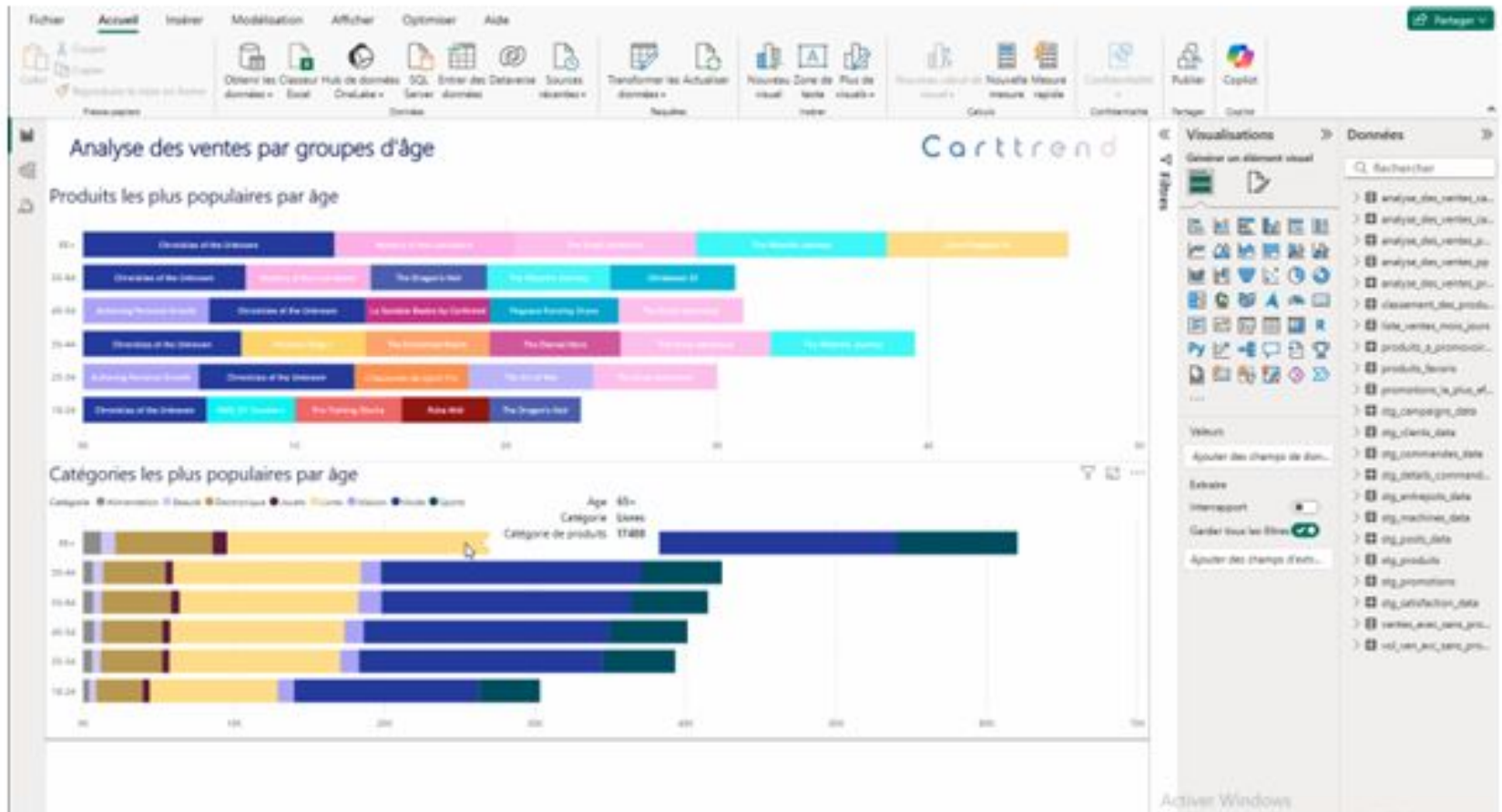


Dashboards interactif PowerBI : Marketing

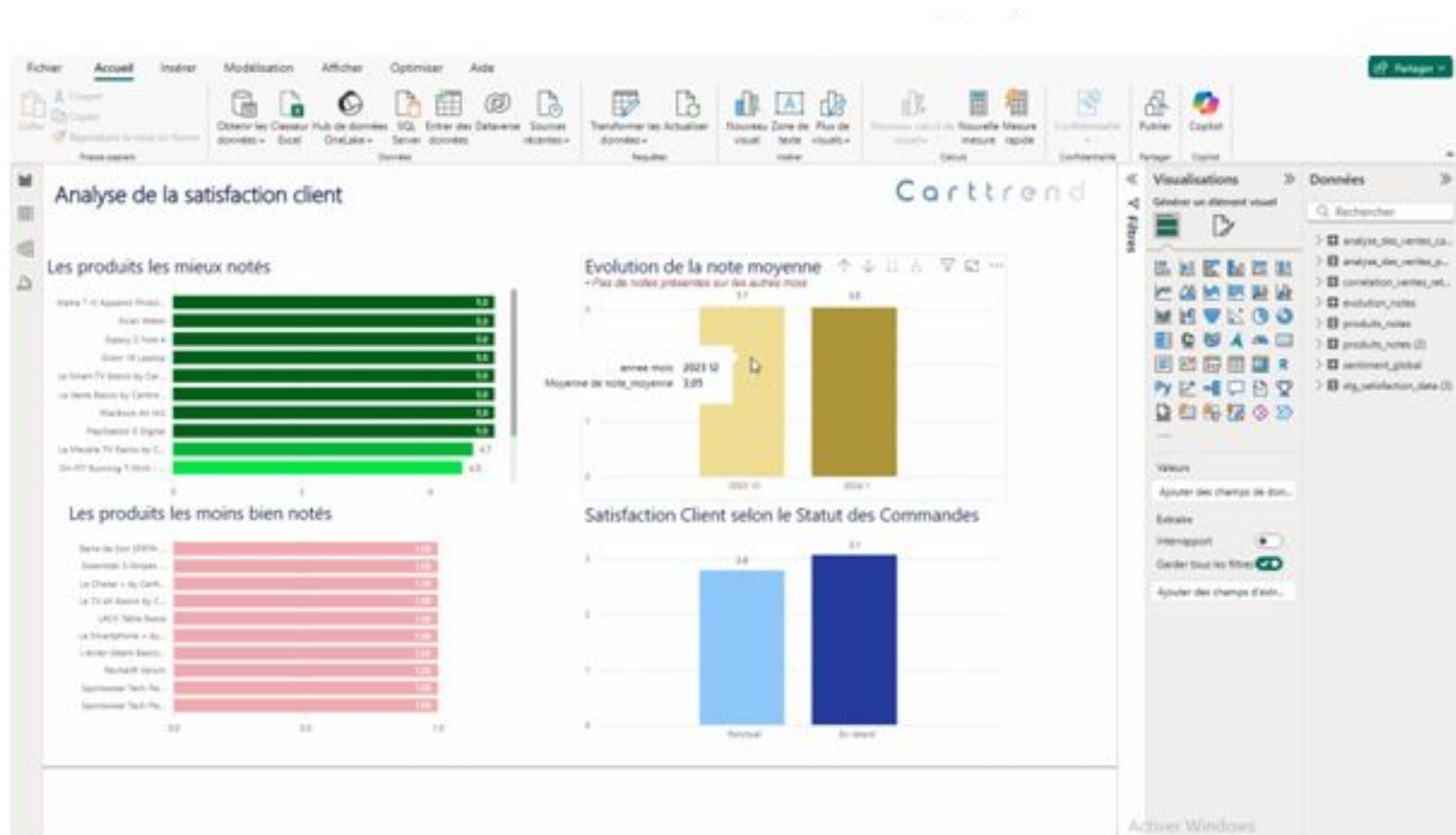
Carttrend



Dashboards interactif PowerBI : Logistique



Dashboards interactif PowerBI : Analyse des ventes



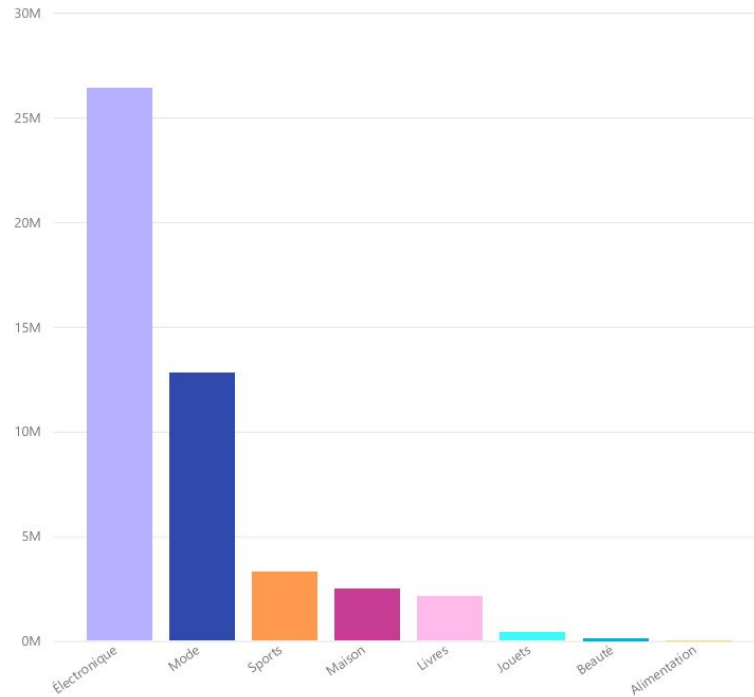
Dashboards interactif PowerBI : Satisfaction client

ANALYSE DES RESULTATS

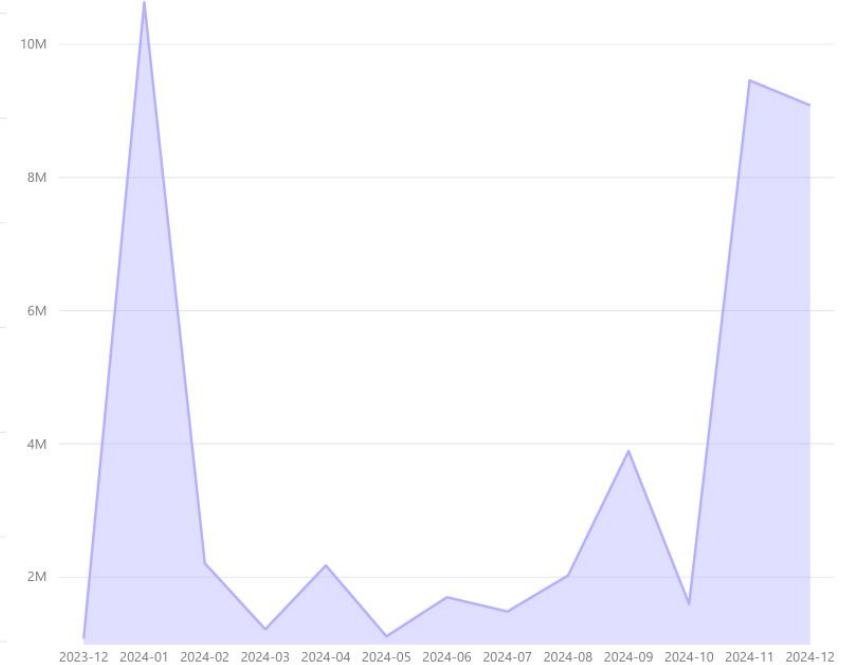
Analyse des ventes - Revenus totaux

Catégories ayant générées le plus de revenus

Catégorie ● Électronique ● Mode ● Sports ● Maison ● Livres ● Jouets ● Beauté ● Alimentation



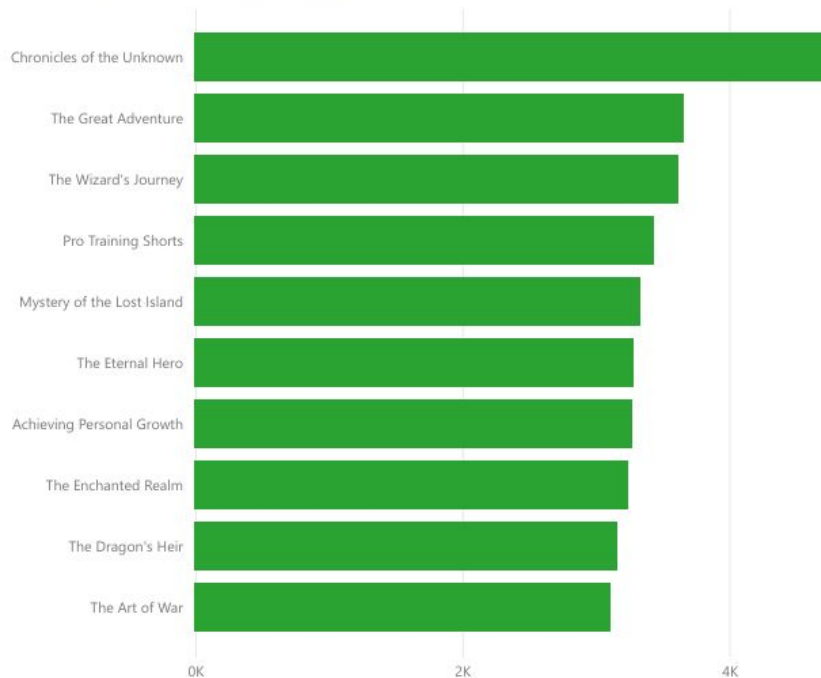
Revenus totaux de Carttrend par mois



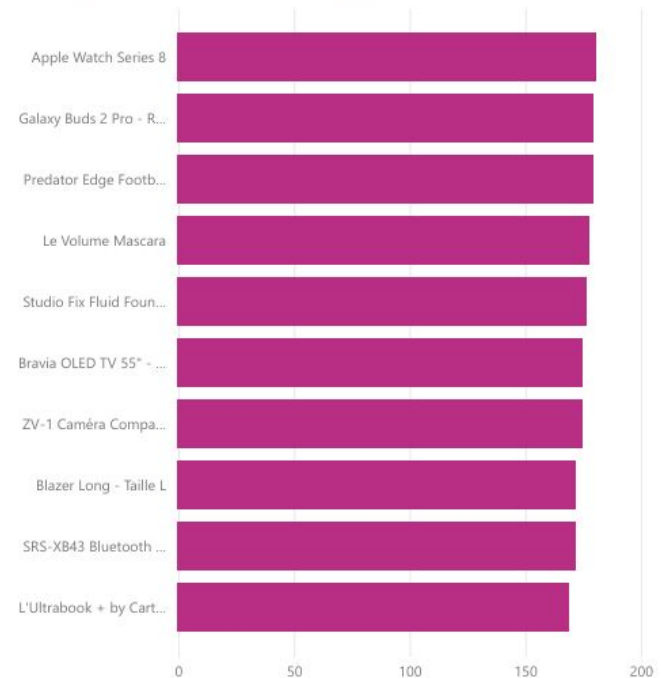
Analyse des ventes - Produits populaires/ Impopulaires

Carttrend

Les 10 produits les plus populaires



Les 10 produits les moins populaires



Analyse des ventes par groupes d'âge

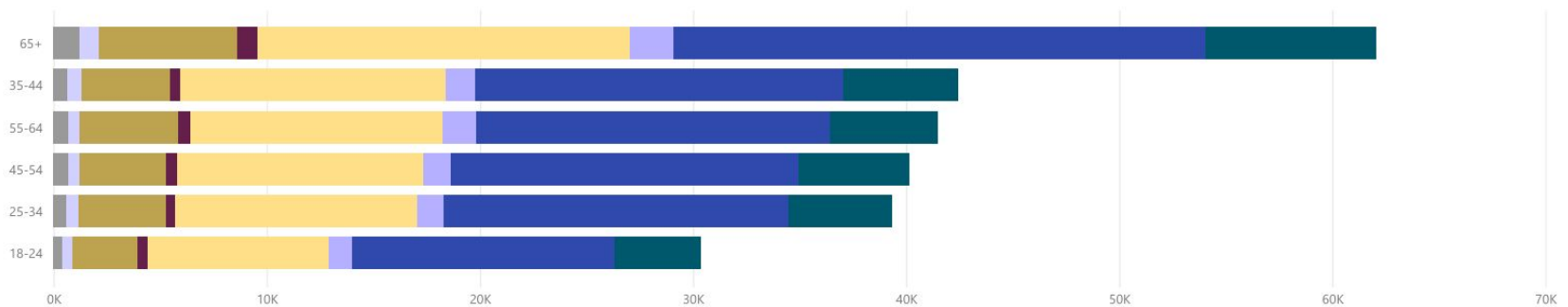
Carttrend

Produits les plus populaires par âge



Catégories les plus populaires par âge

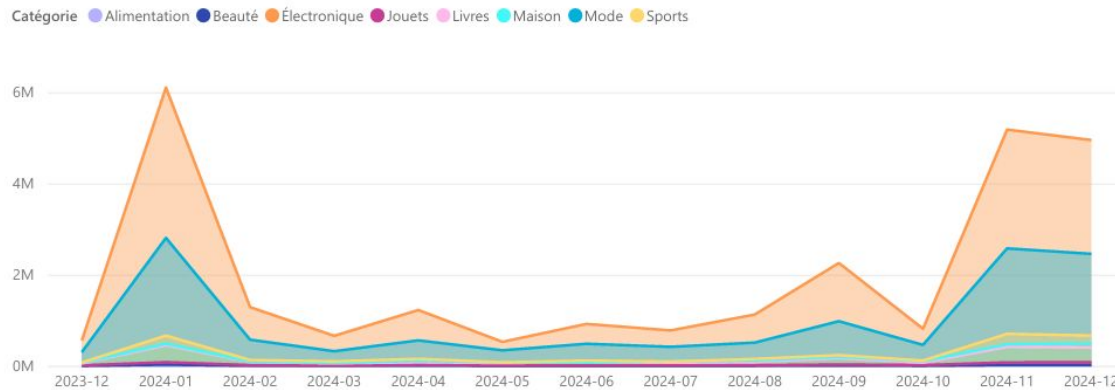
Catégorie ● Alimentation ● Beauté ● Électronique ● Jouets ● Livres ● Maison ● Mode ● Sports



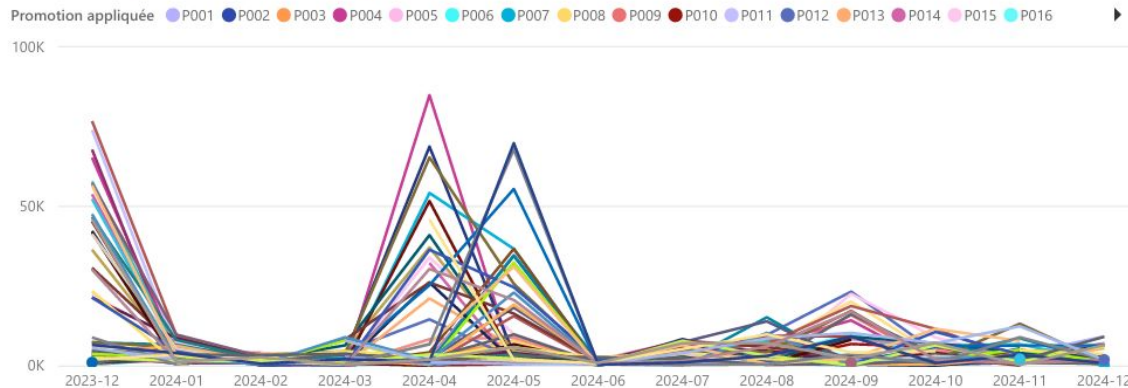
Analyse des ventes - Revenus générés par catégories et impact des promotions

Carttrend

Revenus par mois et par catégories



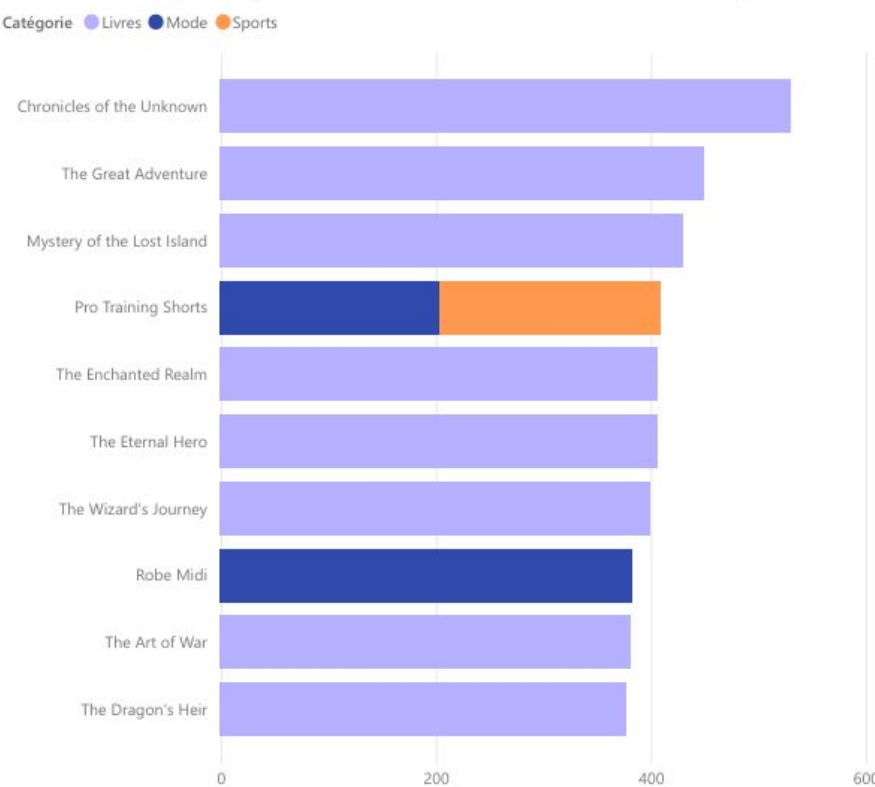
Revenus générés par mois par réponses aux promotions



Carttrend

Analyse des ventes - Produits plébiscités dans les favoris - Produits achetés ensemble

Produits les plus ajoutés dans les favoris et leur catégorie



Produits le plus souvent achetés ensemble (combinaisons)

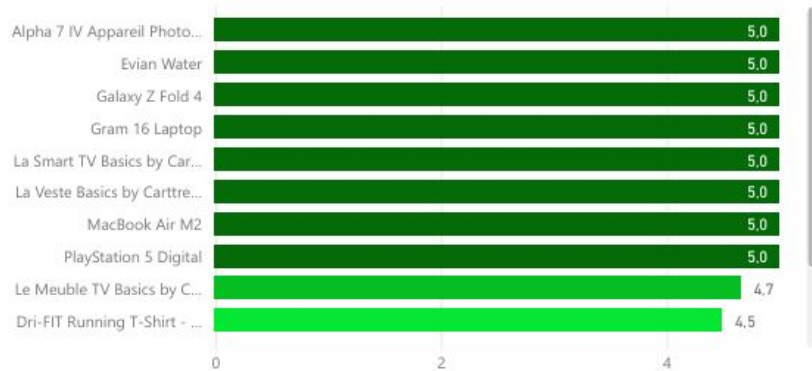
Nombre de combinaisons	Produit 1	Produit 2
4	Achieving Personal Growth	The Enchanted Realm
4	La Chaussure de Sport + by Carttrend	Pro Training Shorts
4	La Sneaker + by Carttrend	The History of Technology
4	Pro Training Shorts	The Power of Habits
4	The Eternal Hero	Tiro 21 Training Pants
3	Adicolor Classics Tee	NMD_R1 Sneakers
3	Chronicles of the Unknown	La Sneaker Basics by Carttrend
3	Chronicles of the Unknown	The Eternal Hero
3	Chronicles of the Unknown	The Story of Civilization
3	Famous Historical Figures	Predator Edge.1
3	La Chaussure de Sport + by Carttrend	La Sandale + by Carttrend
3	La Chaussure de Sport Basics by Carttrend	NMD_R1 Primeblue
3	La Sandale Pro by Carttrend	Zoom Pegasus 39
3	Le T-shirt Basics by Carttrend	The Art of War
3	NMD_R1 Primeblue	Pegasus Running Shoes
3	NMD_R1 Sneakers	The Enchanted Realm
3	Pull en Cachemire	The Wizard's Journey
3	Quantum Mechanics Simplified	The History of Technology
3	The Eternal Hero	Training Shorts
3	The Great Adventure	The Power of Habits
3	The Great Adventure	The Wonders of Biology
2	Achieving Personal Growth	Chemisier Satiné
2	Achieving Personal Growth	Chronicles of the Unknown
2	Achieving Personal Growth	La Chaussure de Sport Basics by Carttrend
2	Achieving Personal Growth	La Sneaker + by Carttrend
2	Achieving Personal Growth	La Sneaker Pro by Carttrend

2002

Analyse de la satisfaction client

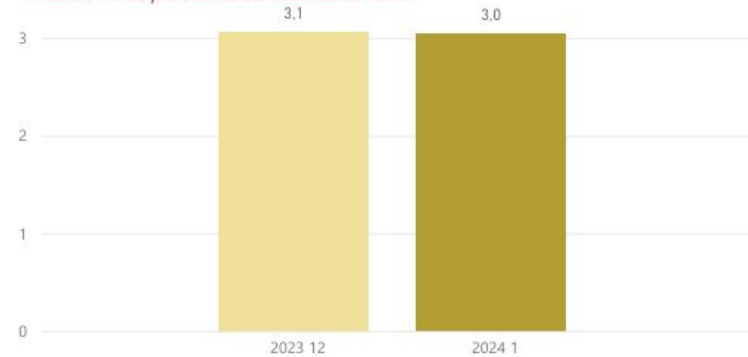
Carttrend

Les produits les mieux notés

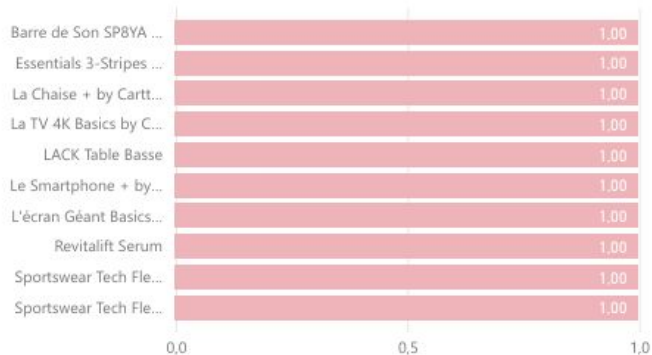


Evolution de la note moyenne

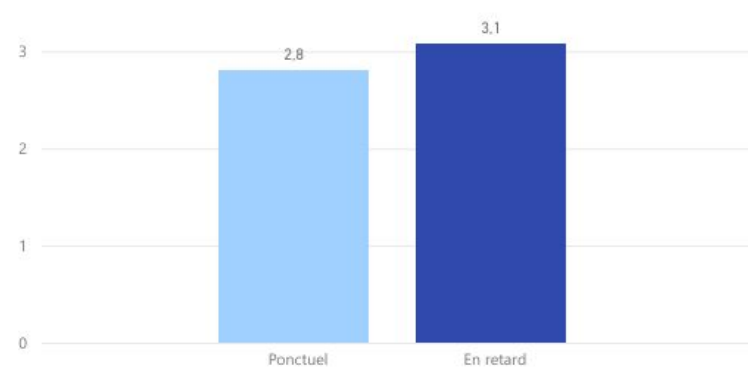
• Pas de notes présentes sur les autres mois



Les produits les moins bien notés



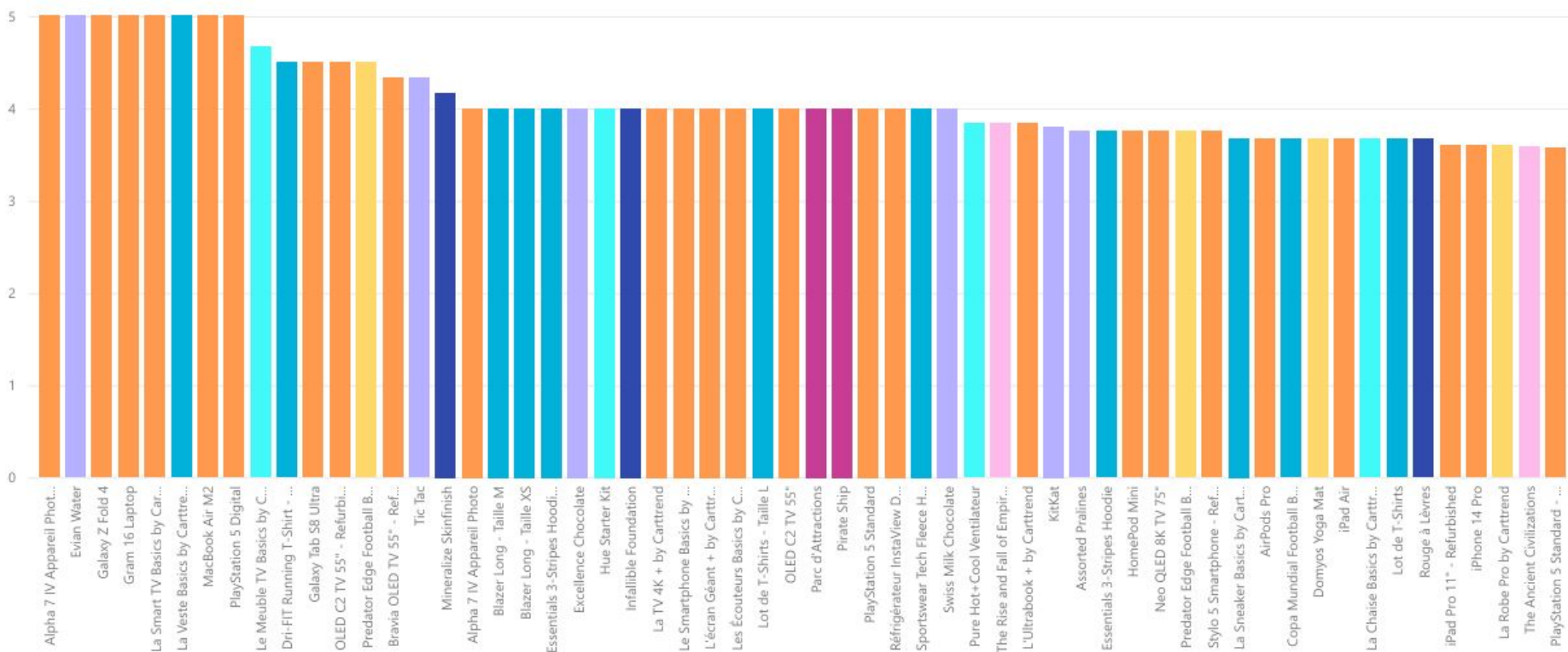
Satisfaction Client selon le Statut des Commandes



Analyse de la satisfaction client - Impact du type de catégorie sur les notes

Carttrend

Catégorie Alimentation Beauté Électronique Jouets Livres Maison Mode Sports



Carttrend



Analyse de la satisfaction client

Carttrend

Nuage de Mots Récurrents dans les Commentaires :

Nuage de mots - Commentaires Positifs:



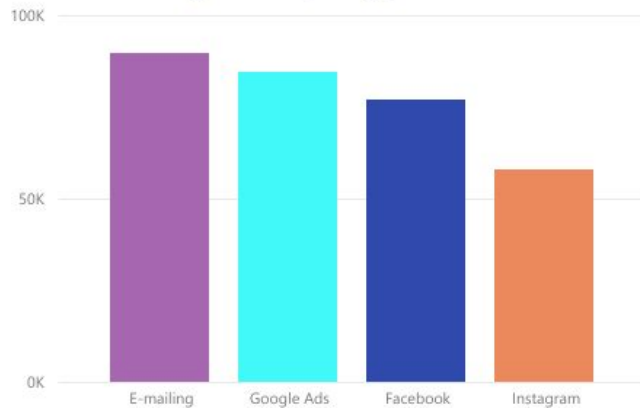
Nuage de mots - Commentaires Négatifs:



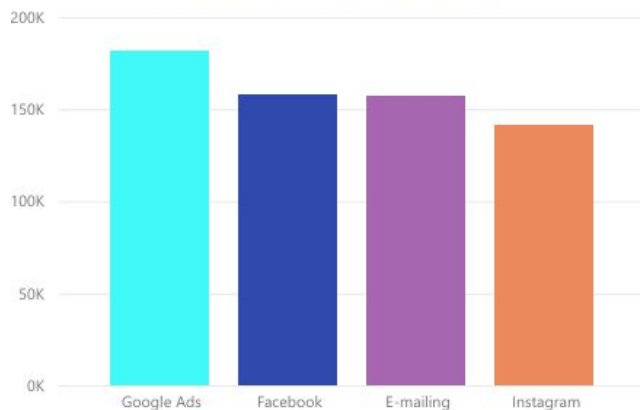
Analyse marketing - Identification des canaux publicitaires les plus efficaces

Carttrend

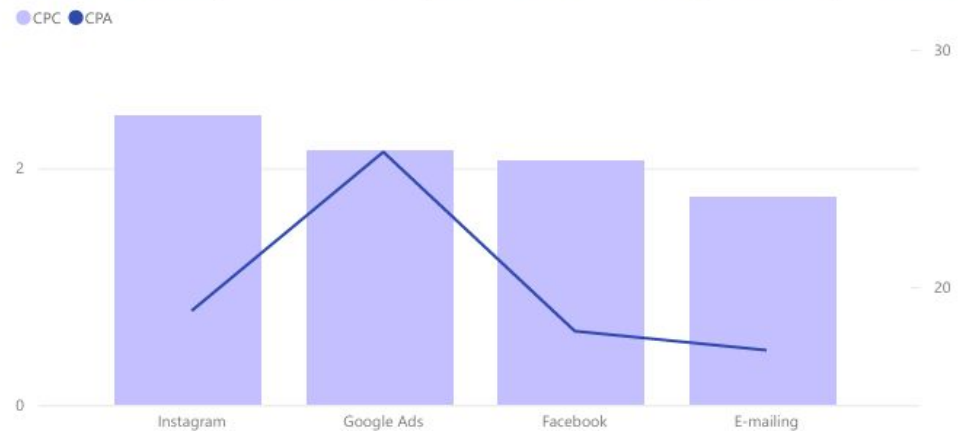
Clics générés par type de canal



Budget par type de canal



Coût en euros par clic CPC et par conversion CPA pour chaque canal



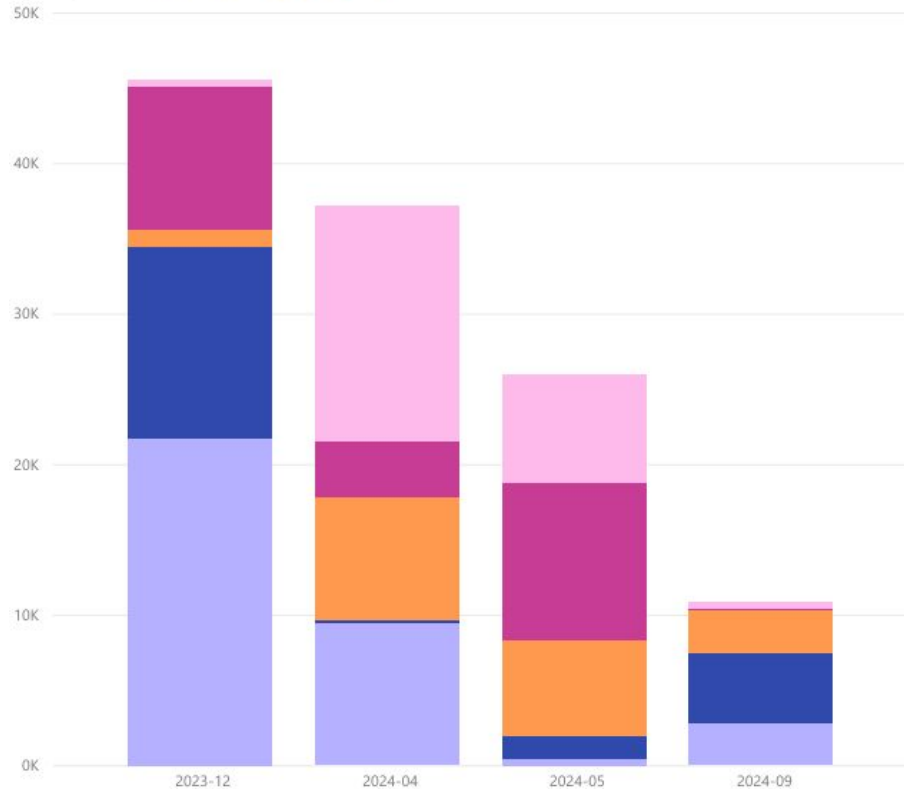
Canal	CPA (Coût par conversion)	CPC (Coût par clic)	Budget cumulé	Total de clics	Total de conversions
E-mailing	17,33	1,75	156896	89493	9056
Facebook	18,13	2,06	157538	76625	8690
Google Ads	25,69	2,15	181197	84406	7054
Instagram	18,99	2,44	140738	57741	7410
Total	80,13	8,39	636369	308265	32210

Analyse marketing - Impact des promotions sur les ventes et les marges

Carttrend

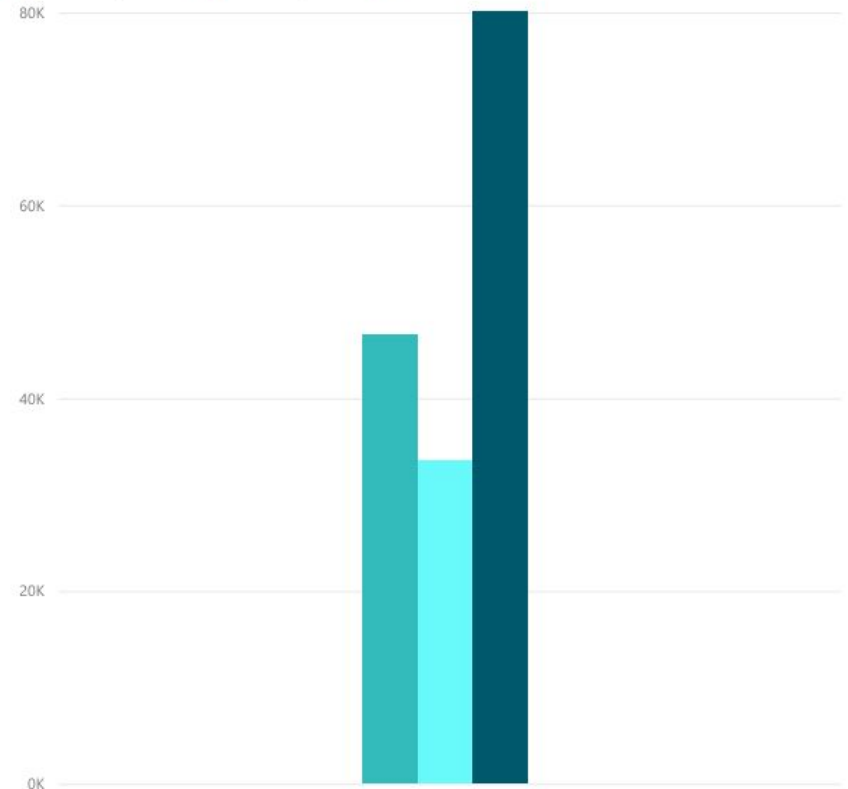
Promotions les plus efficaces par mois les plus performants

Type de promotion ● P004 ● P029 ● P038 ● P092 ● P096



Ventes avec ou sans promotions par rapport au volume total

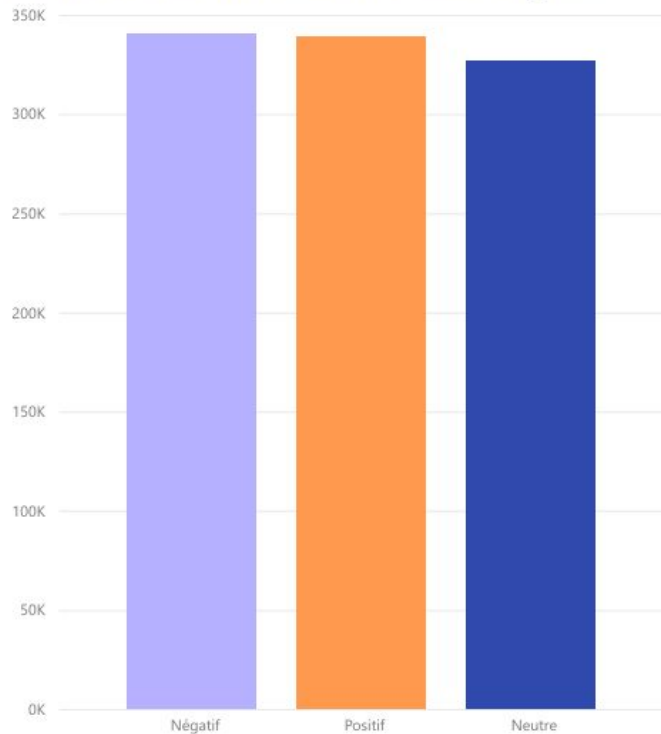
● Ventes avec promotions ● Ventes sans promotion ● Volume de ventes



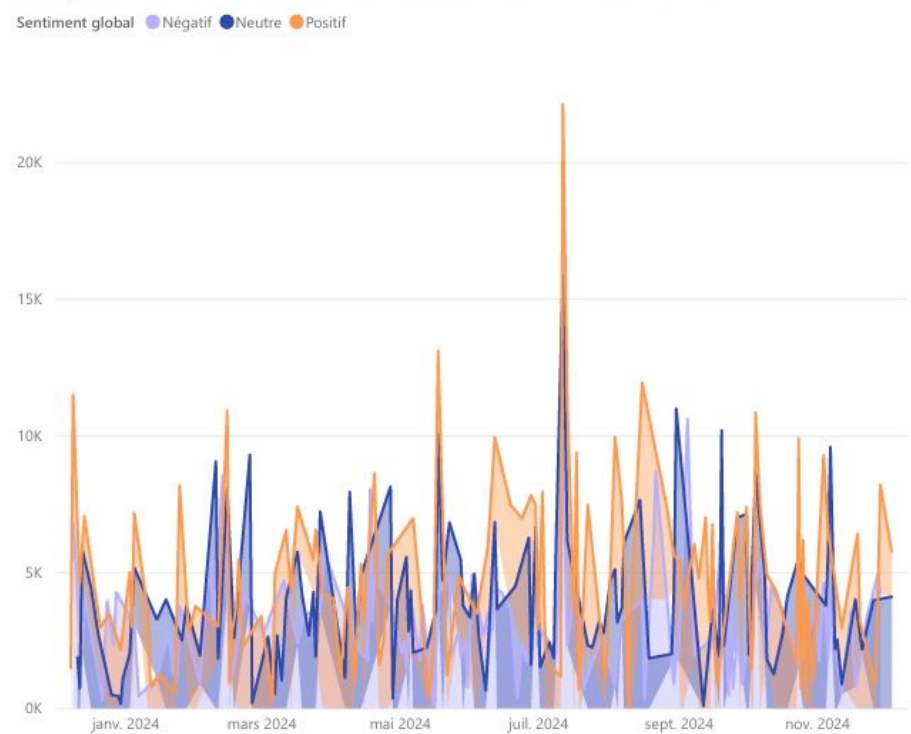
Analyse marketing - Impact des réseaux sociaux sur les ventes et la satisfaction

Carttrend

Analyse des volumes de mentions et sentiment global



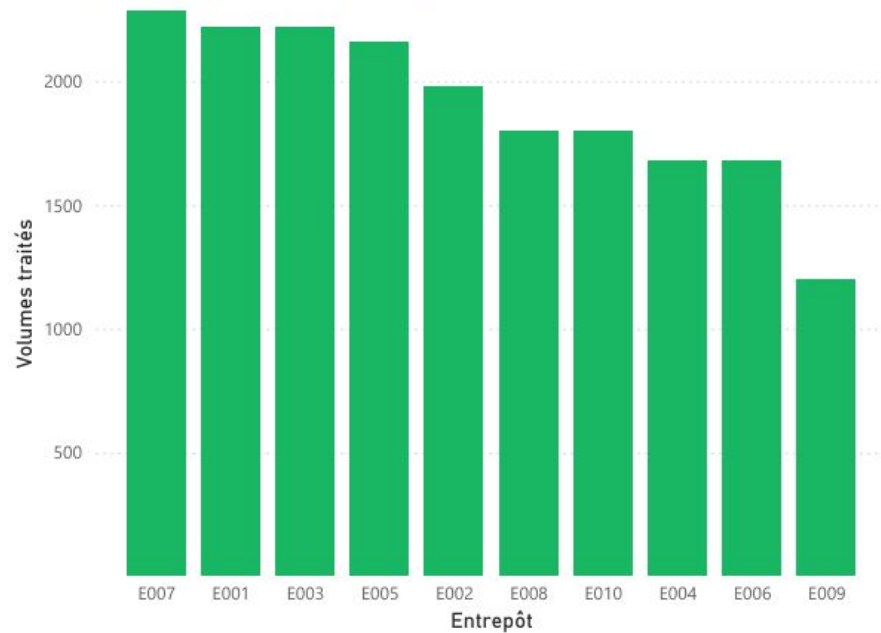
Les pics de mentions de satisfaction selon les ventes



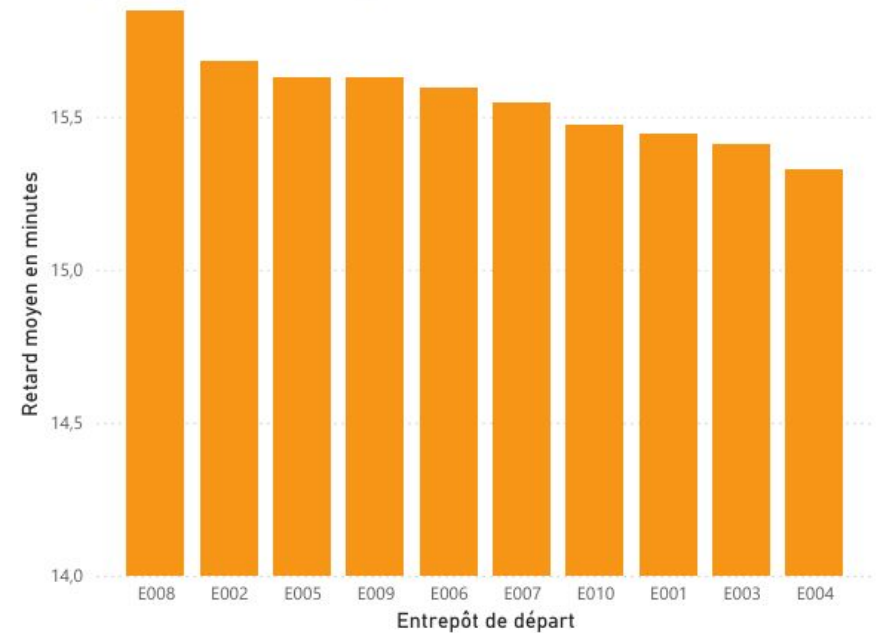
Analyse de la logistique - Entrepôts

Carttrend

Entrepôts qui gèrent les plus gros volumes



Entrepôts qui causent le plus de retards



MACHINE LEARNING

Rôle du machine learning :

- Le ML est généralement adapté aux requêtes prédictive, basée sur des données avec un certain nombre de critères.

Exemple de questions pouvant être résolues par le ML :

- Ventes :
Quelle sera la demande pour chaque produit/catégorie le mois prochain ?
- Satisfaction client :
Quels aspects précis impactent négativement la satisfaction des clients ?
- Marketing :
Quel est le meilleur moment et/ou canal pour envoyer une offre promotionnelle ?
- Logistique :
Quels entrepôts ou transporteurs sont susceptibles d'engendrer des retards ?

MACHINE LEARNING

- Forêts aléatoires (Random Forest) :

```
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import mean_absolute_error, mean_squared_error

# Initialiser le modèle
model = RandomForestRegressor(n_estimators=100, random_state=42)

# Entraîner le modèle
model.fit(X_train, y_train)

# Prédire les valeurs sur l'ensemble de test
y_pred = model.predict(X_test)

# Évaluer les performances
mae = mean_absolute_error(y_test, y_pred)
rmse = mean_squared_error(y_test, y_pred, squared=False)

print(f"MAE: {mae}")
print(f"RMSE: {rmse}")

MAE: 0.08101068343021665
RMSE: 4.057429730599401
```

- Régression linéaire :

```
[64]: from sklearn.linear_model import LinearRegression
      from sklearn.metrics import mean_absolute_error, mean_squared_error

      # Initialiser le modèle de régression linéaire
      linear_model = LinearRegression()

      # Entraîner le modèle
      linear_model.fit(X_train, y_train)

      # Prédire sur l'ensemble de test
      y_pred_linear = linear_model.predict(X_test)

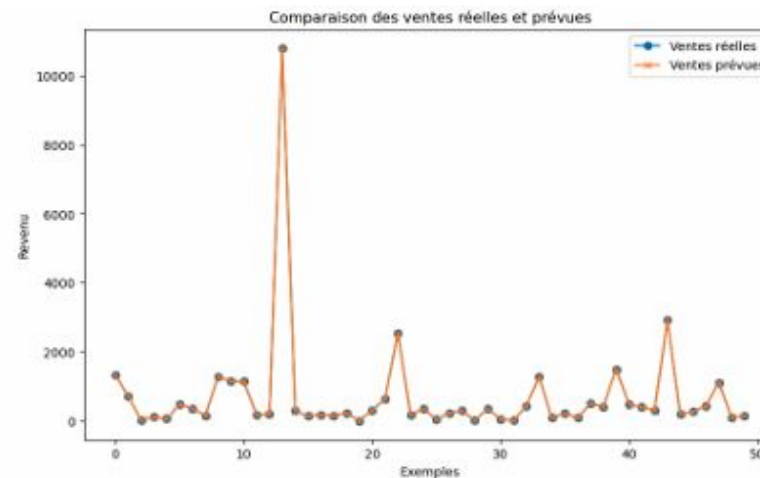
      # Calcul des métriques
      mae_linear = mean_absolute_error(y_test, y_pred_linear)
      rmse_linear = mean_squared_error(y_test, y_pred_linear, squared=False)

      print(f"Régression Linéaire - MAE: {mae_linear}")
      print(f"Régression Linéaire - RMSE: {rmse_linear}")

      Régression Linéaire - MAE: 459.12654395080267
      Régression Linéaire - RMSE: 1185.9037993429984
```

Analyse prédictive pour analyser les ventes futures du 20/12/2024 :

```
# Comparaison des valeurs réelles et prévues
plt.figure(figsize=(10, 6))
plt.plot(y_test.values[:50], label='Ventes réelles', marker='o')
plt.plot(y_pred[:50], label='Ventes prévues', marker='x')
plt.title("Comparaison des ventes réelles et prévues")
plt.xlabel("Exemples")
plt.ylabel("Revenu")
plt.legend()
plt.show()
```



Comparer les ventes réelles et prévues

CONCLUSION

Améliorations à prendre en compte pour de futurs projets :

- Améliorer l'analyse prédictive
- Meilleure planification des visualisations
- Réaliser des tableaux de bord PowerBI plus interactifs

- Compréhension des besoins clients
- Mise en place d'une stratégie efficace
- Analyse approfondie des données
- Visualisation claire et impactante
- Utilisation pertinente et suivi continu des KPI