



Rapport d'Analyse

Projet : Bottleneck – Analyse d'un site e-commerce de vins

Outil de visualisation : Tableau

Date : Mai 2025

Réalisé par : Giulia Governatori



Table des matières

1. **Introduction**
 - 1.1 Objectif du rapport
 - 1.2 Contexte de l'analyse
2. **État des lieux et évaluation de la pertinence**
 - 2.1 Description de la situation actuelle
 - 2.2 Sélection des données exploitables
 - 2.3 Problèmes identifiés
3. **Besoins en outils**
 - 3.1 Besoins pour l'accès et l'extraction des données
 - 3.2 Besoins pour le traitement des données
 - 3.3 Besoins pour la visualisation des données
4. **Solutions envisageables pour l'extraction, le traitement et la visualisation**
 - 4.1 Outils d'extraction et leur utilité
 - 4.2 Outils de traitement de données
 - 4.3 Outils de visualisation
5. **Justification des choix**
6. **Conclusion**
 - 6.1 Synthèse
 - 6.2 Recommandations
7. **Data Cleaning**
 - 7.1 Correction des types de données
 - 7.2 Nettoyage des valeurs invalides ou incohérentes
 - 7.3 Gestion des valeurs manquantes
 - 7.4 Création de nouvelles colonnes pour les analyses
 - 7.5 Reproductibilité garantie
8. **User Stories par profil**
 - 8.1 Chef de produit
 - 8.2 Responsable des achats / Supply Chain
 - 8.3 Directeur financier / Contrôleur de gestion
 - 8.4 Comité exécutif
 - 8.5 Analyste marketing / CRM
9. **Indicateurs clés de performance (KPI)**
 - 9.1 KPI financiers et commerciaux
 - 9.2 KPI de promotion et ROI
 - 9.3 KPI de satisfaction et d'engagement client
 - 9.4 KPI de suivi des prix et de l'inflation
 - 9.5 KPI par segment et classement produit
10. **Présentation des pages du tableau de bord**
 - 10.1 Vue d'ensemble
 - 10.2 Analyse de l'inflation et de la marge brute
 - 10.3 Promotions et ROI
 - 10.4 Meilleurs produits
 - 10.5 Préférences des clients

- 10.6 Analyse par segment
- 10.7 Rotation stock
- 10.8 Propositions business

11. Les Recommandations Business

- 11.1 Cibler les produits à faible rotation
- 11.2 Ne pas promouvoir les meilleures ventes
- 11.3 Analyser les commentaires pour améliorer la qualité perçue
- 11.4 Synchroniser les campagnes avec les périodes creuses
- 11.5 Adapter les objectifs de stock par segment
- 11.6 Suivre les performances par segment pour guider la croissance

12. Méthodologie et utilisation des fichiers

13. Annexes inclus:

- le dictionnaire des données,
- le diagramme UML de la base de données,
- des captures d'écran annotées.

1. Introduction

1.1 Objectif du rapport

Ce rapport présente les étapes de gestion des données menées dans le cadre du projet *Bottleneck*, visant à créer un tableau de bord interactif pour un site e-commerce spécialisé dans la vente de vins. L'objectif de cette visualisation est d'identifier les tendances commerciales, détecter les éventuels freins à la performance et formuler des recommandations stratégiques.

1.2 Contexte de l'analyse

Les données disponibles proviennent d'un fichier SQLite (.db) contenant plusieurs tables relatives à l'activité du site (produits, ventes, promotions, etc.). Une première étape de nettoyage et de structuration a été nécessaire pour préparer les données à la phase de visualisation dans Tableau.

2. État des lieux et évaluation de la pertinence

2.1 Description de la situation actuelle

Le fichier SQLite contient plusieurs tables aux structures hétérogènes. Certaines données étaient incomplètes, mal typées (nombres avec virgules, dates en texte, etc.), ou peu pertinentes pour l'analyse.

2.2 Sélection des données exploitables

Quatre tables sont disponibles pour l'analyse :

- **Web** : informations produits (titres, type, notes clients...)
- **Finance** : informations sur les transactions
- **Promo** : détails des offres promotionnelles
- **Sales** : volumes de vente par produit

2.3 Problèmes identifiés

Les principaux enjeux liés aux données étaient :

- Formats non uniformes (texte au lieu de numérique, dates mal formatées)
- Données manquantes ou incohérentes
- Clés primaires non explicites ou mal nommées
- Difficulté d'importation directe dans Tableau sans traitement préalable

3. Besoins en outils

L'analyse du fichier source a révélé plusieurs besoins à adresser tout au long du processus, de l'extraction à la visualisation. Ces besoins impliquent de faire des choix d'outils et de méthodes adaptés à la nature relationnelle des données et à leur utilisation stratégique.

3.1 Besoins pour l'accès et l'extraction des données

L'accès aux données peut se faire de plusieurs manières, chacune ayant ses avantages et limites. Trois approches génériques ont été identifiées :

Approche	Avantages	Inconvénients
Connexion directe à la base SQLite	<ul style="list-style-type: none">• Données toujours à jour• Pas d'étape d'export supplémentaire	<ul style="list-style-type: none">• Tableau/Power BI ne supportent pas nativement SQLite• Ralentissements possibles si la base est volumineuse
Extraction CSV depuis la base	<ul style="list-style-type: none">• Compatible avec tous les outils• Processus simple à comprendre	<ul style="list-style-type: none">• Export manuel ou script séparé• Nécessite une automatisation pour être robuste
Utilisation d'un ETL (KNIME, PowerQuery, etc.)	<ul style="list-style-type: none">• Workflow visuel et documenté• Automatisation possible• Traitements avancés	<ul style="list-style-type: none">• Courbe d'apprentissage (ex : KNIME)• Dépendance à un logiciel tiers

➔ Besoins associés :

- Accès structuré à plusieurs tables relationnelles via SQLite
- Automatisation de l'importation et de la mise à jour
- Filtrage des colonnes pertinentes
- Gestion des formats hétérogènes (texte, nombres, dates)

3.2 Besoins pour le traitement des données

- Nettoyage des erreurs et gestion des doublons
- Normalisation des formats (dates, devises, booléens)
- Création de nouvelles colonnes KPI (ex : marges, volume par segment)
- Reproductibilité des étapes de traitement

3.3 Besoins pour la visualisation des données

- Interface dynamique et intuitive adaptée à plusieurs profils utilisateurs
- Suivi des KPI : CA, marges, effets des promotions
- Simplicité de mise à jour (fichiers CSV automatisés ou connecteurs)
- Compatibilité avec des sources de données évolutives

4. Solutions envisageables pour l'extraction, le traitement et la visualisation

4.1 Outils d'extraction et leur utilité

Plusieurs outils ont été considérés :

Outil	Avantages	Inconvénients
Python (SQLite3)	Flexible, automatisable, adapté aux structures complexes	Nécessite des bases en programmation
Excel / LibreOffice	Facile à utiliser, bonne lisibilité	Limité sur les gros volumes, pas adapté aux bases relationnelles
Power BI / Tableau (connexion directe)	Visualisation rapide	Risque de lenteur ou bug si les données sont mal nettoyées

➔ **Choix final:** Python avec SQLite3, pour un contrôle complet sur l'extraction et un traitement ciblé des données.

4.2 Outils de traitement de données

Outil	Avantages	Inconvénients
Python (Pandas)	Puissant, automatisable, adapté aux mises à jour régulières	Demande un minimum de compétences en code
KNIME	Interface intuitive sans code	Moins souple, plus lent avec des structures complexes, pas populaire.
Excel	Simple et accessible	Peu fiable sur le long terme, pas adapté aux bases volumineuses

➔ **Choix final:** Python dans Google Colab, pour un pipeline automatisé, versionné, exécutable dans le cloud.

4.3 Outils de visualisation

Outil	Avantages	Inconvénients
Power BI	Intégration facile avec Excel et Microsoft	Moins flexible pour la mise en page, version gratuite limitée
Tableau	Visualisations esthétiques, filtres puissants	Version complète payante
Looker Studio (ex Google Data Studio)	Gratuit, accessible en ligne	Moins performant pour les gros volumes

➔ **Choix final :** Tableau, pour sa puissance graphique et la richesse de ses interactions.

5. Justification des choix

Les outils sélectionnés ont été choisis pour leur capacité à répondre efficacement aux besoins du projet, tout en garantissant flexibilité, automatisation et pérennité technique.

- **Python (via Google Colab)** assure un traitement des données automatisé, reproductible et évolutif. Grâce à l'extraction depuis une base .db SQLite, les scripts centralisent toutes les étapes critiques : import des données brutes, nettoyage, structuration, calcul des KPI. L'exécution dans le cloud (via Colab) évite toute contrainte d'installation locale et permet une maintenance facilitée avec des mises à jour régulières.

+ **Avantage stratégique** : comme toutes les transformations sont codées dans un environnement open-source, le système reste totalement indépendant de l'outil de visualisation. Il suffit de connecter un autre logiciel (Power BI, Looker Studio, etc.) aux fichiers CSV générés, sans toucher au pipeline de préparation des données.

+ **Gain de performance** : le calcul des KPI directement dans Python — plutôt que dynamiquement dans Tableau — permet un affichage plus rapide et fluide des tableaux de bord. En évitant les calculs complexes en temps réel, Tableau se concentre uniquement sur la visualisation, ce qui améliore sensiblement les performances, en particulier sur des volumes moyens.

- **Tableau vs Power BI: Tableau Public c'est un outil de visualisation accessible et intuitif.**
- Mon objectif était de rendre le tableau de bord facilement modifiable par toute personne de l'entreprise, même sans compétences techniques poussées. Contrairement à d'autres outils comme Power BI, qui nécessitent parfois une connaissance plus approfondie du langage DAX ou un environnement Windows, Tableau permet une prise en main rapide grâce à son interface en glisser-déposer.
- De plus, Tableau est conçu pour démocratiser l'accès à la donnée : il facilite la lecture, l'interprétation et la prise de décision. C'est pour cette raison que je l'ai estimé particulièrement adapté à un contexte comme celui d'un site e-commerce de vins, où les équipes marketing, commerciales ou logistiques pourraient toutes avoir besoin de consulter ou d'ajuster les visualisations.
- Les fichiers CSV produits par Python sont directement exploitables dans Tableau, évitant toute connexion directe instable avec SQLite et assurant de bonnes performances d'affichage même sur des volumétries moyennes.

✅ **Conclusion** : la combinaison **extraction automatisée** → **traitement en Python** → **visualisation dans Tableau** constitue un workflow robuste, modulaire et évolutif, garantissant la fiabilité de l'analyse tout en laissant la porte ouverte à d'éventuels changements d'outils à l'avenir.

6. Conclusion

6.1 Synthèse

- Les données d'origine (la base SQLite) présentaient plusieurs problèmes de typage, complétude et structure.
- Un processus automatisé en Python a permis de nettoyer, restructurer et enrichir les données en cinq tables prêtes à être visualisées.
- Tableau offrira une visualisation claire, utile à la prise de décisions stratégiques.

6.2 Recommandations

- Intégrer le notebook Python dans un processus d'automatisation régulier afin de garantir la mise à jour continue des données (par exemple via une planification hebdomadaire dans Google Colab).
- Déployer le tableau de bord Tableau auprès des équipes métiers, accompagné d'une documentation simple et d'un mini-guide de lecture.
- Mettre en place un contrôle qualité en amont (sur les fichiers ou la base SQLite) pour garantir la fiabilité des données sources et éviter les erreurs d'import ou de typage à l'avenir.
- Envisager une évolution future vers une base de données plus robuste (MySQL, PostgreSQL) si les volumes de données ou la fréquence d'analyse augmentent.

7. DATA CLEANING

Actions de nettoyage et traitement appliquées:

7.1. Correction des types de données:

- Conversion des chaînes de caractères en int, float, datetime, bool selon les colonnes
- Normalisation des formats décimaux (virgule → point)
- Transformation explicite des champs booléens (0/1, true/false, yes/no → bool)
- Conversion des dates avec gestion des erreurs (format européen dayfirst=True)

7.2. Nettoyage des valeurs invalides ou incohérentes:

- Suppression des lignes avec des valeurs hors norme (ex. : Onsale_Web ≠ 0/1)
- Vérification et suppression des doublons (via fonctions auxiliaires)
- Renommage des clés primaires (Key_* → Key_Finance) pour uniformiser les jointures

7.3. Gestion des valeurs manquantes:

- Imputation aléatoire pour préserver la distribution normale (Average_rating)
- Récupération de valeurs cachées dans d'autres colonnes (ex. : Contenance extraite des titres produits si vide)

7.4. Création de nouvelles colonnes pour les analyses:

- Création automatique de colonnes manquantes (SKU, Date) à partir de Key_Finance.

7.5. Reproductibilité garantie:

- Encapsulation des conversions dans des fonctions robustes et réutilisables
- Nettoyage systématique appliqué à chaque table (Web, Finance, Promo, Sales)
- Affichage des erreurs et rapports des colonnes traitées

8. User Stories par profil

8.1 Chef de produit (Vin, Spiritueux, Sans-alcool)

- **En tant que** chef de produit,
je veux accéder rapidement au chiffre d'affaires et aux volumes de vente de ma gamme,
afin de suivre la performance de mon portefeuille et ajuster mes offres.
 - **En tant que** chef de produit,
je veux visualiser les tops et flops produits de mon segment,
afin de décider quels références promouvoir ou retirer.
 - **En tant que** chef de produit,
je veux comparer les ventes avant, pendant et après promotions,
afin de mesurer l'efficacité des offres et leur ROI.
-

8.2 Responsable des achats / Supply Chain

- **En tant que** Responsable des achats,
je veux suivre la rotation des stocks et les niveaux de stock fin de mois,
afin de anticiper les ruptures et planifier les réapprovisionnements.
 - **En tant que** Responsable des achats,
je veux connaître l'évolution mensuelle du prix d'achat,
afin de négocier avec les fournisseurs et maîtriser les coûts.
-

8.3 Directeur financier / Contrôleur de gestion

- **En tant que** Directeur financier,
je veux suivre le chiffre d'affaires mensuel et la marge brute,
afin de évaluer la rentabilité globale et réagir aux fluctuations.
 - **En tant que** Directeur financier,
je veux analyser l'impact de l'inflation sur les prix d'achat et de vente,
afin de ajuster la politique tarifaire et protéger la marge.
-

8.4 Comité exécutif

- **En tant que** membre du Comité exécutif, je **veux** une « page de garde » avec les KPI clés, **afin de** détecter immédiatement les anomalies (baisse de marge, rupture de stock...).
 - **En tant que** membre du Comité exécutif, je **veux** disposer d'un accès global à toutes les analyses (ventes, promotions, satisfaction client...), **afin de** piloter la stratégie à long terme.
-

8.5 Analyste marketing / CRM

- **En tant que** analyste marketing, je **veux** visualiser les notes moyennes clients et le taux de commentaires, **afin de** mesurer la satisfaction et l'engagement pour adapter les actions de fidélisation.
 - **En tant que** analyste marketing, je **veux** filtrer par segment « sans-alcool », **afin de** suivre la croissance de ce nouveau marché.
-

9. Indicateurs clés de performance (KPI)

9.1 KPI financiers et commerciaux

1. **Chiffre d'affaires mensuel (CA)**
 - **Formule** : $\sum (\text{Price} \times \text{Sales_HP}) + \sum (\text{Price_Offer} \times \text{Sales_Offer})$
 - **Tables** : Finance (Price, Key_Finance), Sales (Sales_HP), Promo (Price_Offer, Sales_Offer)
 - **Objectif** : Suivre l'évolution du chiffre d'affaires global, promotions incluses.
 2. **Marge brute mensuelle**
 - **Formule** :
 $\sum [(\text{Price} - \text{Purchase_Price}) \times \text{Sales_HP}] + \sum [(\text{Price_Offer} - \text{Purchase_Price}) \times \text{Sales_Offer}]$
 - **Tables** : Finance (Price, Purchase_Price), Sales, Promo
 - **Objectif** : Mesurer la rentabilité globale, en tenant compte des promotions.
 3. **Rotation des stocks**
 - **Formule** : $(\text{Sales_HP} + \text{Sales_Offer}) / \text{Stock}$
 - **Tables** : Sales (Sales_HP, Stock), Promo (Sales_Offer)
 - **Objectif** : Identifier les produits à forte ou faible rotation pour ajuster les stocks.
 4. **Taux de conversion produit**
 - **Formule** : $\text{Orders} / (\text{Sales_HP} + \text{Sales_Offer})$
 - **Tables** : Sales (Orders, Sales_HP), Promo (Sales_Offer)
 - **Objectif** : Évaluer l'efficacité des offres et la transformation des intentions d'achat.
-

9.2 KPI de promotion et ROI

5. **Taux de promotion**
 - **Formule** : $\sum \text{Sales_Offer} / (\sum \text{Sales_HP} + \sum \text{Sales_Offer})$
 - **Tables** : Sales (Sales_HP), Promo (Sales_Offer)
 - **Objectif** : Suivre la part des ventes réalisées en promotion.
6. **ROI des promotions**
 - **Formule** : $(\text{CA_promo} - \text{Coût_promo}) / \text{Coût_promo}$
 - **Détail** :

- CA_promo = $\sum (\text{Price_Offer} \times \text{Sales_Offer})$
- Coût_promo = $\sum [(\text{Price} - \text{Price_Offer}) \times \text{Sales_Offer}]$
- **Tables** : Promo (Price_Offer, Sales_Offer), Finance (Price)
- **Objectif** : Mesurer la rentabilité réelle des campagnes promotionnelles.

9.3 KPI de satisfaction et d'engagement client

7. **Note moyenne client**
 - **Formule** : Moyenne(Average_rating)
 - **Table** : Web
 - **Objectif** : Suivre la perception des produits par les clients.
8. **Taux de commentaires**
 - **Formule** : $\sum \text{Comment_count} / \sum (\text{Sales_HP} + \text{Sales_Offer})$
 - **Tables** : Web (Comment_count), Sales + Promo (ventes)
 - **Objectif** : Mesurer l'engagement des clients pour chaque produit.

9.4 KPI de suivi des prix et de l'inflation

9. **Évolution mensuelle du prix d'achat**
 - **Formule** : $(\text{Price}_{\text{mois } n} - \text{Price}_{\text{mois } n-1}) / \text{Price}_{\text{mois } n-1}$
 - **Table** : Finance
 - **Objectif** : Suivre l'évolution de l'inflation sur les coûts d'approvisionnement.

9.5 KPI par segment et classement produit

10. **Chiffre d'affaires par segment**
 - **Formule** : $\sum (\text{Price} \times \text{Sales})$ groupé par Segment
 - **Tables** : Web (Segment), Finance (Price), Sales (Sales)
 - **b** : Identifier les segments les plus performants.
11. **Top/Flop par volume de ventes**
 - **Formule** : Classement des SKU selon $\sum \text{Sales}$
 - **Tables** : Sales (Sales), Web (SKU)
 - **Objectif** : Repérer les produits les plus ou moins populaires.
12. **Top/Flop par chiffre d'affaires**
 - **Formule** : Classement des SKU selon $\sum \text{CA}$
 - **Tables** : Sales (CA), Web (SKU)
 - **Objectif** : Identifier les produits qui génèrent le plus de revenu.
13. **Classement par note pondérée (score bayésien)**
 - **Formule** : $(\text{Average_rating} \times \text{Rating_count} + m \times C) / (\text{Rating_count} + m)$, avec $m = 30$
 - **Tables** : Web (Average_rating, Rating_count)
 - **Objectif** : Valoriser les produits bien notés avec un volume d'avis suffisant.
14. **Top/Flop rotation de stock**
 - **Formule** : Classement croissant des SKU selon Rotation_stock
 - **Tables** : Sales + Promo (ventes), Stock
 - **Objectif** : Identifier les produits à écoulement lent à promouvoir.

10. Présentation des pages du tableau de bord

Afin de répondre aux enjeux de compétitivité, d'organisation et de croissance de l'entreprise, un tableau de bord interactif a été conçu sous Tableau Public. Il est structuré en plusieurs pages thématiques, chacune dédiée à un aspect clé de la performance commerciale.

Chaque vue permet une lecture ciblée et une prise de décision rapide grâce à des indicateurs pertinents, des visualisations dynamiques et des filtres personnalisables.

10.1. Vue d'ensemble

Une synthèse visuelle des indicateurs clés de performance : chiffre d'affaires, marge brute, volumes de ventes et taux de commentaires. Cette vue permet de suivre l'évolution globale de l'activité et d'identifier rapidement les tendances mensuelles.

10.2. Analyse de l'inflation et de la marge brute

Cette page explore l'impact des variations de prix d'achat sur la marge brute. Elle met en lumière les périodes de baisse ou de hausse de rentabilité, aidant à anticiper les effets de l'inflation sur la performance.

10.3. Promotions et ROI

Une analyse comparative des performances des campagnes promotionnelles. On y mesure le retour sur investissement (ROI), en évaluant les ventes générées et leur rentabilité, pour optimiser les futures stratégies marketing.

10.4. Meilleurs produits

Un classement des produits les plus performants selon deux critères: ventes et chiffre d'affaires. L'objectif est de repérer les valeurs sûres et d'identifier les produits phares à mettre en avant.

10.5. Préférences des clients

Cette section analyse les préférences des clients à travers les notes et les commentaires. Elle permet de mieux comprendre les attentes des consommateurs et d'adapter l'offre en conséquence.

10.6. Analyse par segment

Une visualisation des performances par segment de marché: elle permet de comparer les chiffres d'affaires entre segments et de détecter les opportunités de développement. Il est également possible ici de comparer les produits selon la qualité des retours clients : notes et commentaires.

10.7. Rotation stock

Une mesure de la vitesse de vente des produits, essentielle pour une gestion efficace des stocks. Les produits à faible rotation sont identifiés pour des actions correctives, telles que des promotions ou une révision de l'assortiment.

10.8. Propositions business

Sur la base des analyses précédentes, cette page formule des recommandations concrètes pour améliorer la compétitivité, la gestion des stocks, la fidélisation client et la croissance de l'entreprise.

11. Les Recommandations Business

Objectif : Présenter clairement des propositions stratégiques issues des analyses, avec support visuel et justifications.

11.1. 📦 Cibler les produits à faible rotation pour les promotions

Insight

Certains produits restent longtemps en stock avec une rotation inférieure à 1.

Recommandation

Planifier des campagnes promotionnelles ciblées sur ces produits, en analysant leur marge et leur note moyenne pour adapter le prix. Un nouveau packaging ou une description retravaillée peut également relancer l'intérêt.

11.2. 💡 Ne pas promouvoir les meilleures ventes

Insight

Les top produits (SKU les plus vendus, meilleur CA, meilleures notes) se vendent bien sans promotion.

Recommandation

Éviter de réduire les marges sur ces références. Si promotion nécessaire, envisager une légère hausse de prix après la campagne pour maintenir la rentabilité.

11.3. 🗒️ Analyser les commentaires pour améliorer la qualité perçue

Insight

Un produit avec beaucoup de commentaires, une note Moyenne basse et une rotation élevée peut cacher un problème de qualité ou d'attentes non satisfaites.

Recommandation

Effectuer une analyse sémantique des commentaires pour comprendre les points faibles perçus par les clients et adapter l'offre ou la communication.

11.4. 🎯 Synchroniser les campagnes avec les périodes creuses

Insight

Le CA des ventes non promotionnelles augmente lors des campagnes (effet d'entraînement). En revanche, il stagne ou baisse sans animation marketing.

Recommandation

Organiser des campagnes à ROI élevé (ex : ROI de 5 lors de la deuxième campagne) pendant les périodes de faible activité pour booster globalement les ventes.

11.5. 💡 Adapter les objectifs de stock par segment

Insight

La rotation optimale varie selon le type de produit : spiritueux, huiles, vins, sans alcool. Une stratégie unique ne convient pas.

Recommandation

Définir des seuils de rotation adaptés à chaque segment et automatiser les alertes en cas de déviation (surstock

ou rupture).

11.6. 📁 Suivre les performances par segment pour guider la croissance

Insight

Le comportement des clients varie selon les segments (vins, huiles, sans alcool...). Certains segments montrent une forte croissance, d'autres stagnent malgré les promotions.

Recommandation

Mettre en place un suivi mensuel des KPIs par segment (CA, rotation, taux de commentaires, marges) afin d'identifier les relais de croissance et réallouer les ressources marketing vers les segments les plus porteurs.

12. Méthodologie et utilisation des fichiers

Ce projet repose sur une démarche rigoureuse, divisée en plusieurs étapes: extraction, nettoyage, modélisation, analyse, visualisation et recommandations stratégiques. Afin d'assurer la transparence, la reproductibilité et la maintenabilité du tableau de bord, plusieurs fichiers complémentaires sont mis à disposition:

- **Notebook Python pédagogique (Google Colab)**
Ce fichier contient l'analyse complète et détaillée du processus de préparation des données : nettoyage, transformation, création des indicateurs clés de performance (KPI), gestion des valeurs aberrantes, etc. Il est commenté de manière didactique pour faciliter la compréhension de la logique utilisée.
- **Notebook Python automatisé (Google Colab)**
Ce second fichier est un script simplifié destiné à un usage opérationnel. Il permet à l'utilisateur de **mettre à jour les données automatiquement**, en quelques clics ("Exécuter toutes les cellules"). Une fois l'exécution terminée, les nouveaux fichiers sont sauvegardés dans un **dossier cloud synchronisé** avec Tableau.

⚠️ **Il est nécessaire de recharger manuellement les fichiers (juste cliquer "refresh data") dans Tableau Desktop**, puis de **publier à nouveau le classeur sur Tableau Public** pour que le tableau de bord en ligne reflète les nouvelles données.
- **Fichier Tableau (.twb ou .twbx)**
Ce fichier contient toutes les visualisations, filtres, calculs et paramètres du tableau de bord. Il permet de **modifier la structure ou le design** du tableau de bord, si besoin, avant une nouvelle publication.
- **Ce rapport d'analyse**

Annexes inclus:

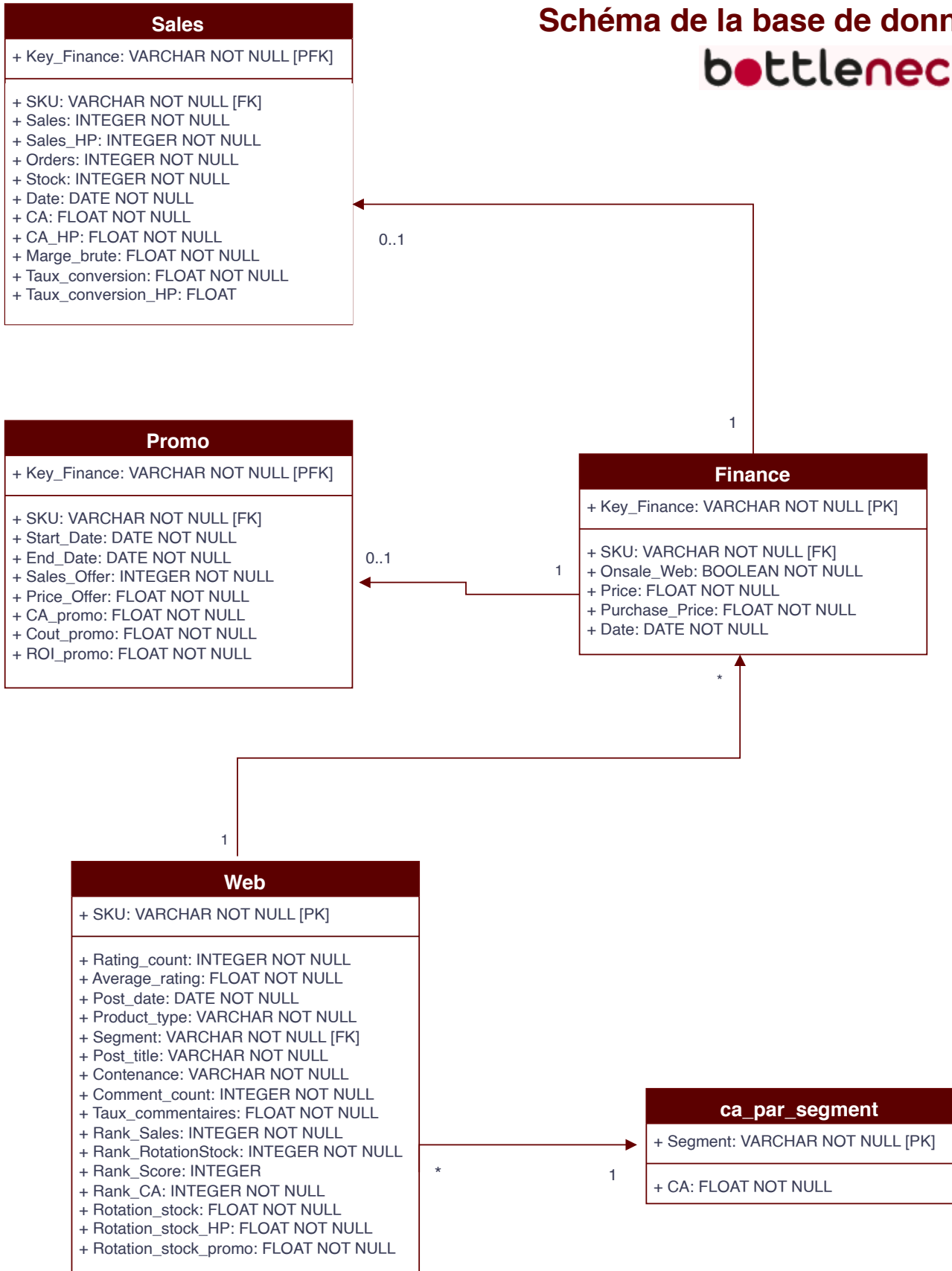
- le dictionnaire des données,
- le diagramme UML de la base de données,
- des captures d'écran annotées.

	Nom des colonnes	Type de données	Clé	Description
WEB	SKU	object	PK	Référence unique du produit (clé primaire)
	Rating_count	Int64		Nombre de notes clients reçues par le produit
	Average_rating	float64		Note moyenne des clients
	Post_date	datetime		Date de création de l'article sur le site
	Product_type	object		Type de produit (cognac, vin, champagne, etc.)
	Segment	object	FK	Segment de produit (spiritueux, vin, sans alcool, etc.)
	Post_title	object		Nom du vin
	Contenance	object		Contenance du produit (ex: 70 cl)
	Comment_count	Int64		Nombre de commentaires des clients
	Taux_commentaires	float64		Proportion de ventes ayant reçu un commentaire
	Rank_Sales	Int64		Classement du produit basé sur le volume de ventes
	Rank_RotationStock	Int64		Classement du produit basé sur la rotation stock moyenne du produit (SKU): ordre croissant
	Rank_Score	Int64		Classement global basé sur un score agrégé entre Average_rating et Rating_count (moyenne bayésienne)
	Rank_CA	Int64		Classement du produit basé sur le chiffre d'affaires
	Rotation_stock	float64		ratio entre les ventes et le stock moyen sur la période; mesure la fréquence de renouvellement du stock.
	Rotation_stock_HP	float64		rotation stock sans prendre en compte les promotions.
Rotation_stock_promo	float64		rotation stock seulement des ventes en promotion.	
FINANCE	Key_Finance	object	PK	Clé primaire : identifiant unique (concaténation de la date (année-mois) et de la référence produit)
	SKU	object	FK	Référence du produit
	Onsale_Web	bool		Présence du produit sur le site (1 = oui, 0 = non)
	Price	float64		Prix de vente unitaire hors promotion
	Purchase_Price	float64		Prix d'achat unitaire du produit
	Date	datetime		Date du mois (exemple : 01/10/2022 = données d'octobre 2022)
SALES	Key_Finance	object	PFK	Identifiant unique de la ligne de vente (année-mois_SKU)
	SKU	object	FK	Référence du produit
	Sales	int64		Nombre d'unités vendues sur le mois (TOTALE)
	Sales_HP	int64		Nombre d'unités vendues sur le mois (hors promotions)
	Orders	int64		Commandes fournisseurs passées sur le mois
	Stock	int64		Stock de l'entreprise au dernier jour du mois
	Date	datetime		Date du mois correspondant à la ligne de vente
	CA	float64		Chiffre d'affaires TOTALE (aussi les promotions)
	CA_HP	float64		Chiffre d'affaires réalisé hors périodes promotionnelles
	Marge_brute	float64		Marge brute réalisée sur les ventes hors période promotionnelle
	Taux_conversion	float64		Taux de conversion logistique (commandes fournisseurs / ventes réalisées)
	Taux_conversion_HP	float64		Taux de conversion logistique (sans considerer les promotions)
PROMO	Key_Finance	object	PFK	Identifiant unique de la ligne de promotion (année-mois_SKU)
	SKU	object	FK	Référence du produit
	Start_Date	datetime		Date de début de la promotion
	End_Date	datetime		Date de fin de la promotion
	Sales_Offer	int64		Nombre de produits vendus pendant la promotion
	Price_Offer	float64		Prix de vente appliqué pendant la promotion
	CA_promo	float64		Chiffre d'affaires généré pendant la promotion
	Cout_promo	float64		Coût total de la promotion (perte de revenu liée à la réduction du prix)
ROI_promo	float64		Retour sur investissement (ROI) de la promotion	

CPS

Segment	object	Segment de produit (spiritueux, vin, sans alcool, etc.)
CA	float64	Chiffre d'affaires par segment.

Schéma de la base de donnée



Analyse de l'Inflation & Marge



Type Produit: (All)

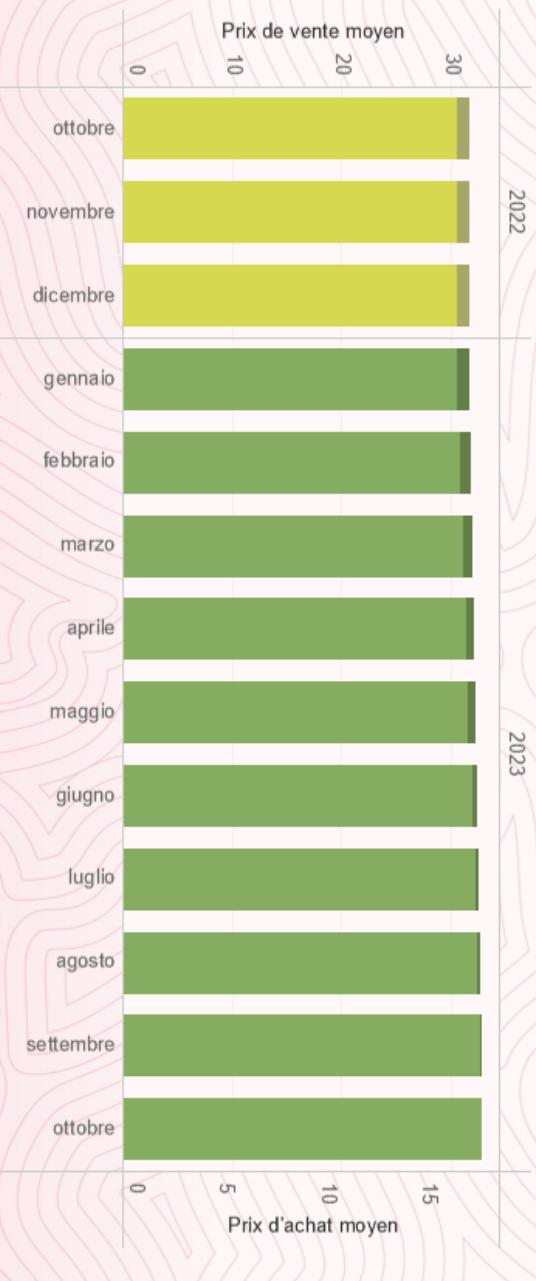
Date: (All)

Segment: (Multiple values)

Trend CA & Marge



Évolution des prix



Analyse du Delta des prix

La variation du « delta » est simplement la différence entre le prix le plus récent et le précédent: si le delta est positif, cela signifie que les prix ont augmenté, si le delta est négatif, cela signifie qu'ils ont diminué.

Month	%varPrixAchat	%varPrixVente
Null		
octobre 2022	0.7	0.4
novembre 2022	0.0	0.0
decembre 2022	0.0	0.0
januaire 2023	0.0	0.0
fevrier 2023	0.0	0.0
mars 2023	0.0	0.0
avril 2023	0.0	0.0
mai 2023	0.0	0.0
juin 2023	0.0	0.0
juillet 2023	0.0	0.0
août 2023	0.0	0.0
septembre 2023	0.0	0.0
octobre 2023	0.0	0.0

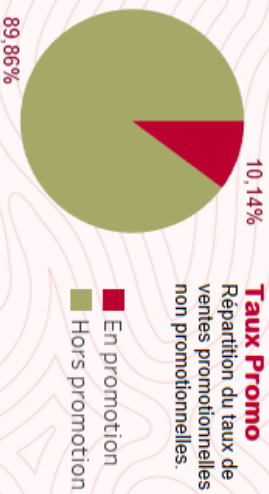
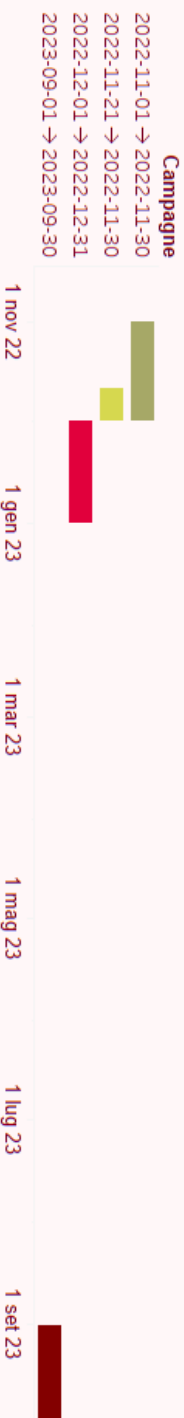
2 nulls

Promotions & ROI

Type Produit: (All) ▼ Date: (All) ▼ Segment: (Multiple values) ▼

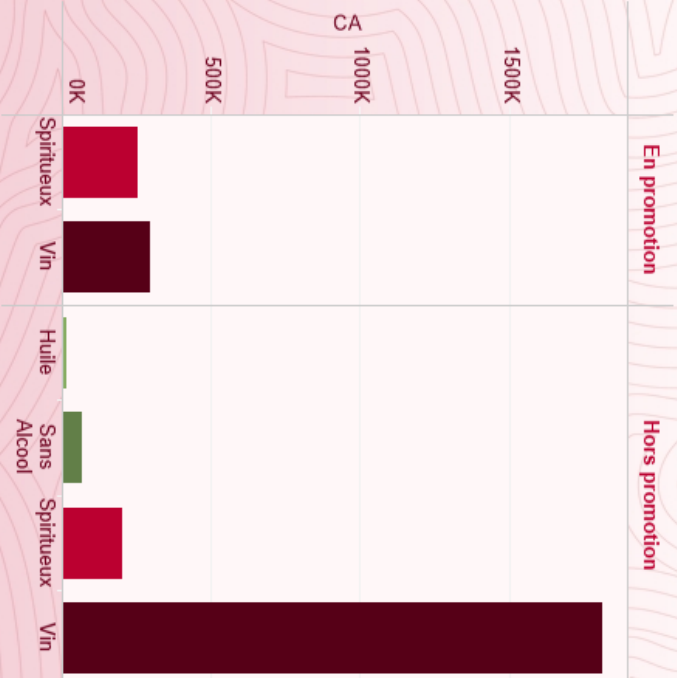
Gantt des promotions

Calendrier des campagnes promotionnelles.



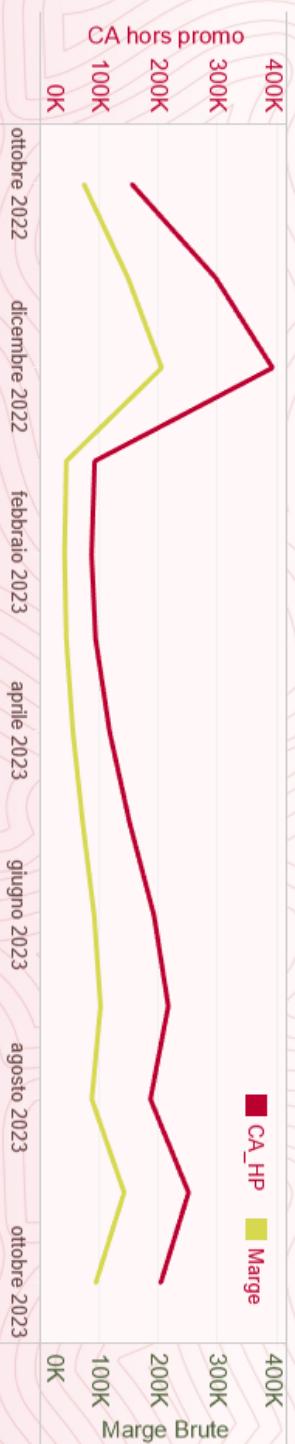
CA promo vs CA PAS promo

Comparaison du CA réalisé avec et sans promotions.



Trend CA (Hors Promo) & Marge

Visualisation de l'évolution du CA (hors promotions) et de la marge.

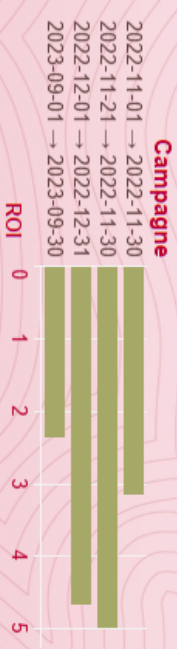


Campagne	CA_promo	Cout_promo	Price	Price_Offer	Sales_Offer
Null	0	0	289.637		
2022-11-01 → 2022-11-30	8.233	1.978	1.770	1.426	153
2022-11-21 → 2022-11-30	15.468	2.579	2.063	1.770	225
2022-12-01 → 2022-12-31	70.632	12.447	3.833	3.195	1.104
2023-09-01 → 2023-09-30	81.161	24.109	19.388	14.724	4.018

ROI Promos

Pour chaque euro sacrifié en réduction, le ROI indique combien d'euros de CA ont été générés.

ROI = Somme du CA en promotion - somme de la perte due à la réduction) / somme de la perte due à la réduction)

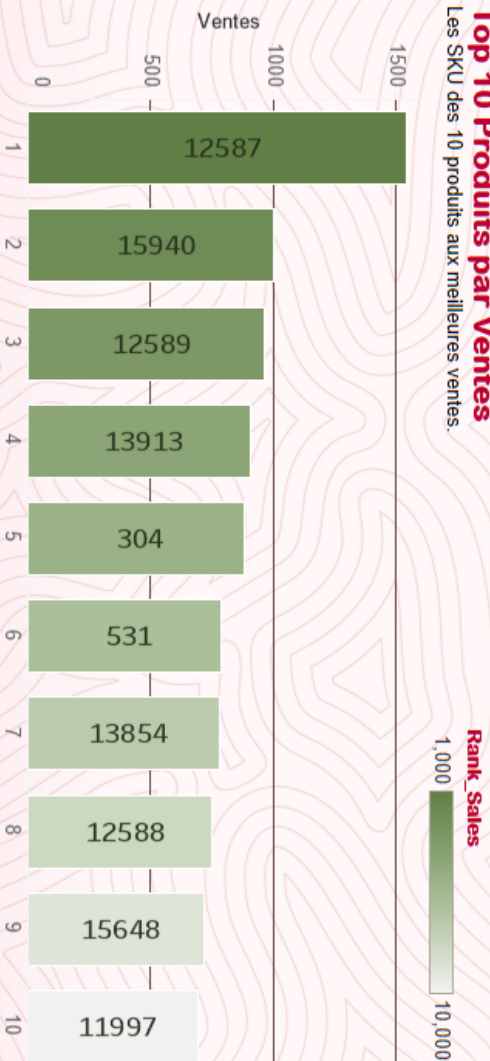


Meilleurs Produits

Type Produit: Date: Segment:

Top 10 Produits par Ventes

Les SKU des 10 produits aux meilleures ventes.

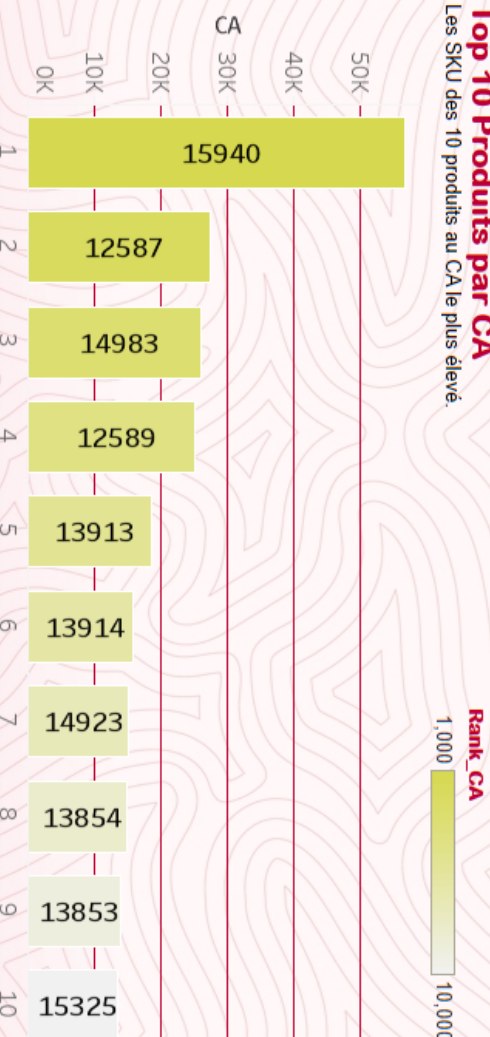


Top 10 Produits par Ventes

R..	Product_t.	Title	SKU	Avg..	Price	Ventes
1	Champagne	Champagne Egly-Ouriet Grand Cru Brut Rosé	12587	3,5	79,5	12587
2	Champagne	Champagne Egly-Ouriet Grand Cru Millésimé 2008	15940	3,8	225	15940
3	Champagne	Champagne Egly-Ouriet Grand Cru Blanc de Noirs	12589	4,7	126,5	12589
4	Champagne	Champagne Agrapat & Fils Minéral Extra Brut Blanc de Blancs Grand Cru 2012	13913	4,2	86,8	13913
5	Champagne	Champagne Gosset Grande Réserve	304	4,7	39	304
6	Champagne	Champagne Petit Lebrun & Fils Blanc de Blancs Grand Cru	531	3,2	22,5	531
7	Champagne	Champagne Lamandier-Bernier Grand Cru Vieilles Vignes du Levant 2008	13854	4	77	13854
8	Champagne	Champagne Egly-Ouriet Grand Cru Extra Brut V.P.	12588	3,4	79,5	12588
9	Champagne	Champagne Mailly Grand Cru Extra Brut Millésimé 2012	15648	4,5	33	15648
10	Champagne	Champagne Agrapat & Fils Terroirs Brut Blanc de Blancs Gr...	11997	4,7	62,1	11997

Top 10 Produits par CA

Les SKU des 10 produits au CA le plus élevé.



Top 10 Produits par CA

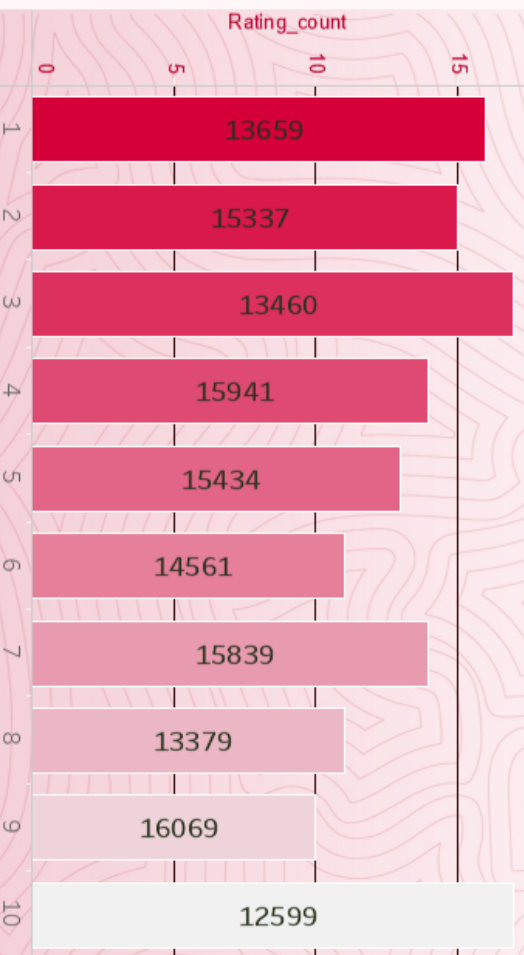
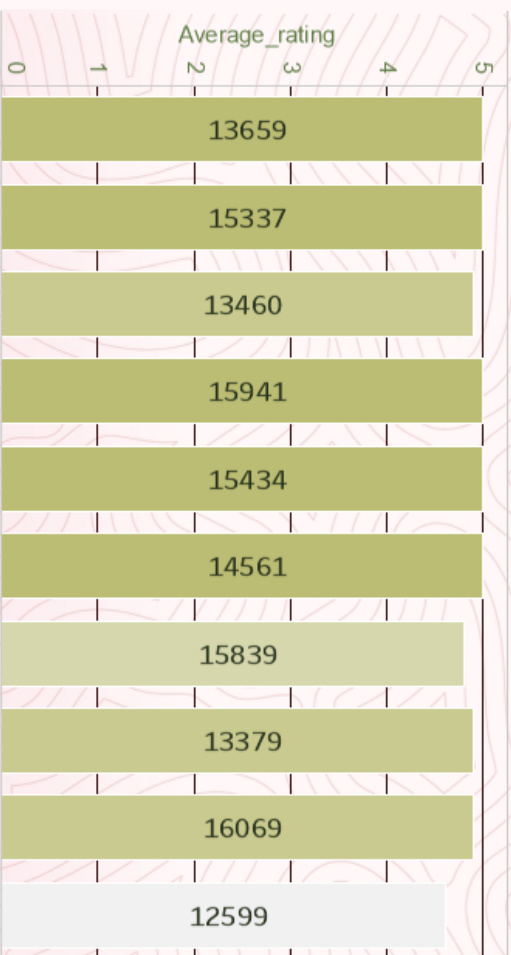
R..	Product_t.	Title	SKU	Avg..	Price	CA
1	Champagne	Champagne Egly-Ouriet Grand Cru Millésimé 2008	15940	3,8	225	15940
2	Champagne	Champagne Egly-Ouriet Grand Cru Brut Rosé	12587	3,5	79,5	12587
3	Champagne	Coteaux Champenois Egly-Ouriet Ambonnay Rouge 2022	14983	3,5	191,3	14983
4	Champagne	Champagne Egly-Ouriet Grand Cru Blanc de Noirs	12589	4,7	126,5	12589
5	Champagne	Champagne Agrapat & Fils Minéral Extra Brut Blanc de Blancs Grand Cru 2012	13913	4,2	86,8	13913
6	Champagne	Champagne Gosset Calébris Vintage 2007	14923	5	135	14923
7	Champagne	Champagne Lamandier-Bernier Grand Cru Vieilles Vignes du Levant 2008	13854	4	77	13854
8	Champagne	Champagne Egly-Ouriet Grand Cru Extra Brut V.P.	12588	3,4	79,5	12588
9	Champagne	Champagne Mailly Grand Cru Extra Brut Millésimé 2012	15648	4,5	33	15648
10	Vin	Agnès Levet Côte Rôte Améthyste 2021	15325	4,8	41,2	15325

Préférences des clients



Type Produit: (All) Date: (All) Segment: (All)

Top 10 Produits par Note Client



Distribution Note Client

Moyenne des notes attribuées par les clients.



Distribution de "Average_rating"



Moyenne bayésienne entre note moyenne et nombre d'avis

Pour établir le classement, on utilise la *moyenne bayésienne*, qui combine la note moyenne du produit (*Average_rating*) et son nombre d'avis (*Rating_count*). Si un produit n'a que quelques avis, sa note est ramenée vers la note moyenne globale: cela évite qu'un score élevé obtenu avec très peu d'avis le propulse injustement en tête. À l'inverse, plus un produit a d'avis, plus son score se rapproche de sa véritable note moyenne.

Les Produits préférés des clients

Produit	SKU	Avg. Price	Pur..	Notes		
1	Domaine Plageoles Vin de France Contre-Pied 2016	13659	12,8	6,94	16 notes	
2	Oratoire Saint Martin Cairanne Blanc Haut-Coustias 2021	15337	5	20,5	11,02	17 notes
3	Domaine Sérol Côte Roannaise Perdrière 2016	13460	4,9	16,6	8,32	17 notes
4	Domaine de la Jérôme Côtes du Rhône Village 201	15941	5	12,7	6,76	14 notes
5	La Préceporie Côtes du Roussillon Blanc Terres Nouvelles 2018	15434	5	19	10,31	13 notes
6	Argentine Mendoza Alamos Torrontes 2021	14561	5	11,1	5,68	11 notes
7	Domaine Sérol Côte Roannaise Les Millierands 2018	15839	4,8	15,5	8,25	14 notes
8	Domaine Rouge Garance Côtes du Rhône Villages Garances 20...	13379	4,9	12,5	6,46	11 notes
9	Plateau des Chênes Lirac 2015	16069	4,9	14,7	7,97	10 notes
10	Domaine de l'Écu Muscadet Orthogneiss 2014	12599	4,6	16,7	8,97	17 notes

Analyse par segment



Type Produit:

(All)

Date:

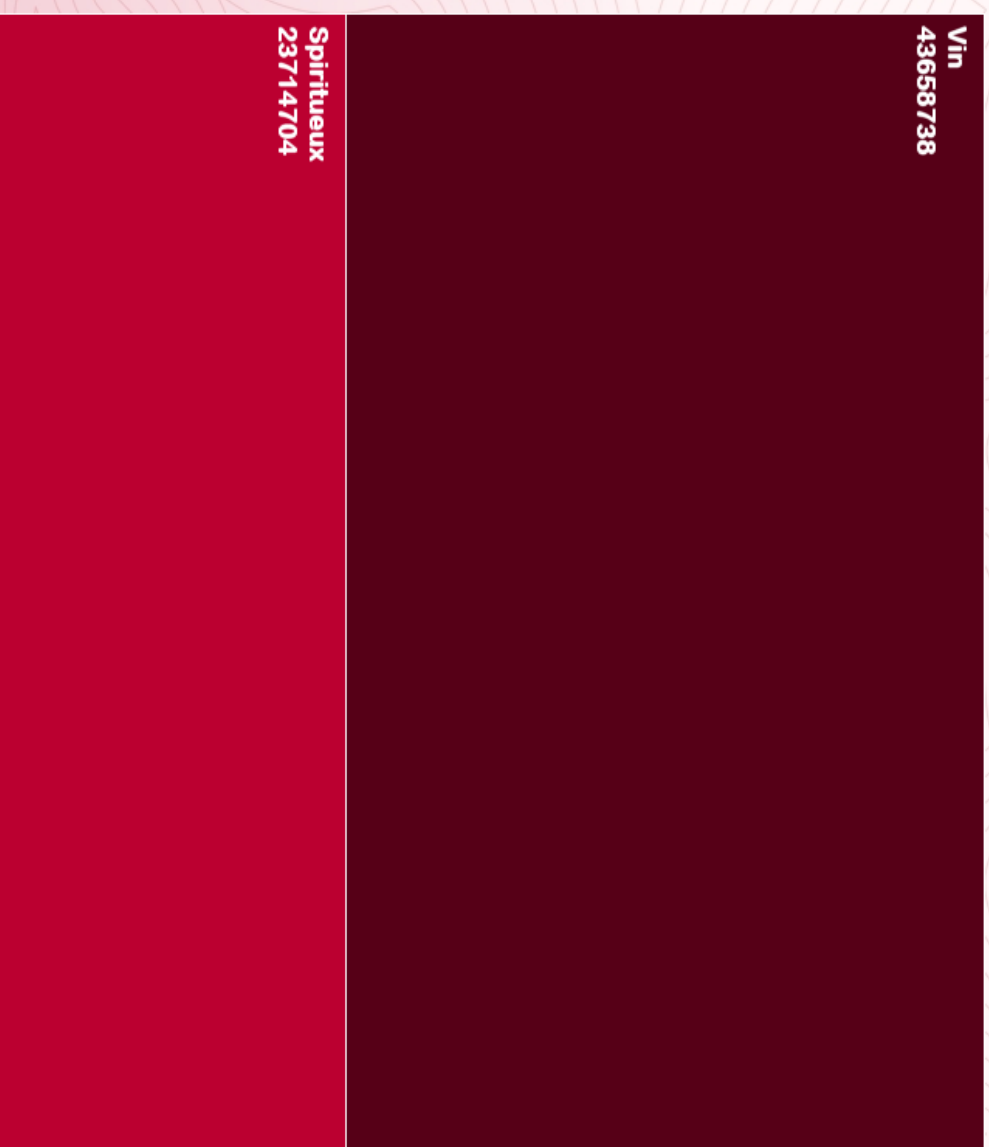
(All)

Segment:

(All)

Treemap Segment

Répartition du chiffre d'affaires par segment.



Engagement vs CA

Note Moyenne:

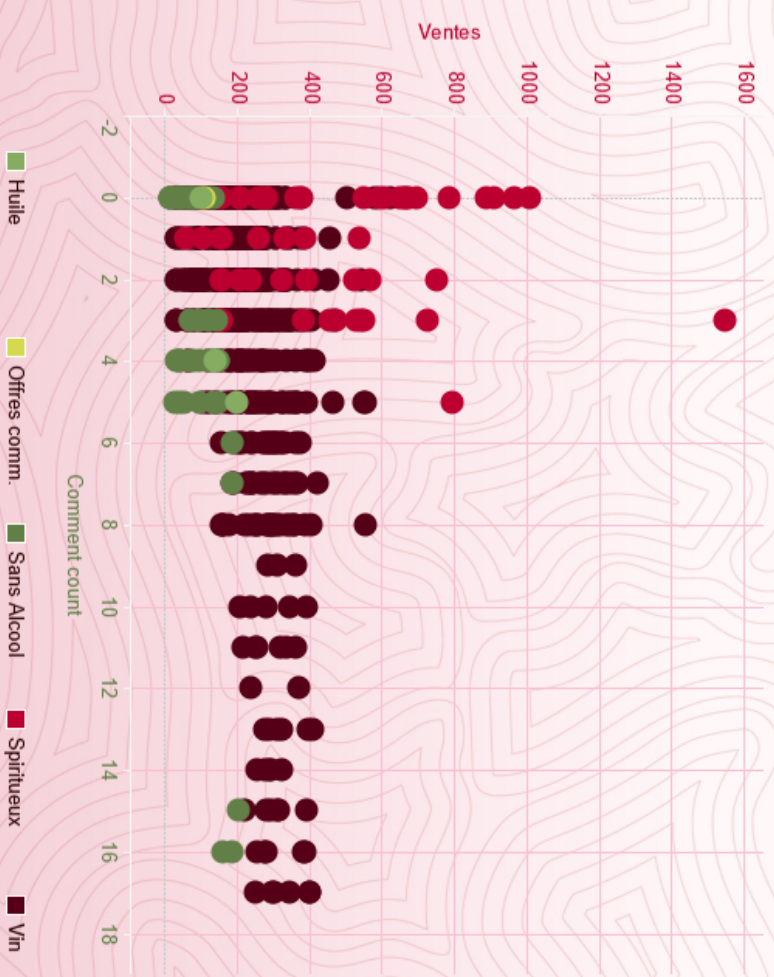
2,930



5,000

Ce graphique compare CA et engagement client par produit et segment, pour identifier rapidement opportunités et points faibles.

Haut droite: produits phares – forte demande et interaction ; maintenir stock et visibilité.
 Haut gauche: bons vendeurs sans feedback ; enrichir fiche produit (images, avis) pour fidéliser.
 Bas droite: fort intérêt mais peu de conversion ; revoir prix/présentation ou lancer une promo.
 Bas gauche: faible performance ; envisager promotion ciblée ou suppression de la référence.



Niveau de stock des produits

Etat actuel des stocks par produit

Rotation Stock

* Le "Marge_Prix" est la différence entre le prix d'achat et le prix de vente.

Rank..	Product_t.	Title	SKU	Rotation..	Rotation..	Stock	Prix de ..	Marge_..	AVG_Ra.	Rating_..	Comme..
1	Vin	Château de Meursault Puligny-Montrachet 1er Cru Champ Canet 2014	12857	0,07	0,04	0	421	114	56	5	3
2	Vin	Domaine Weinbach Gewurztraminer Grand Cru Furstentum SGN 201..	14915	0,08	0,04	0	379	126	58	3	2
3	Vin	Zind-Humbrecht Riesling Grand Cru Rangen De Thann Clos Saint-Ur..	15382	0,09	0,05	0	310	100	47	3	0
4	Vin	David Dubard Chammes-Chamberlin Grand Cru 2014	14581	0,12	0,08	0	371	224	100	3	0
5	Vin	Clos du Mont-Olivet Côtes-du-Rhône Blanc 2019	16028	0,12	0,07	0	226	14	7	4	3
6	Huile d'olive	Huiles d'Olive Extra Vierge Planeta 3x 10cl	13849	0,13	0,13	0	1.523	20	7	4	5
7	Vin	Domaine des Comtes Laton Volnay 1er Cru Champans 2016	15070	0,13	0,09	0	335	129	61	3	1
8	Vin	Château Palsance Fronton Rouge 2017	15575	0,14	0,14	0	1.273	9	4	5	1
9	Huile d'olive	Huile d'Olive Extra Vierge Planeta 50cl	11258	0,14	0,14	0	1.013	13	5	3	4

Range du Vin:

Rotation Stock	Interprétation	Action conseillée
> 3	Produit très demandé, risque de rupture de stock	Augmenter le stock, pas de promo
1 - 3	Équilibre	Aucune action urgente
≤ 1	Ventes lentes, risque de surstock	Promotion, révision de la stratégie

Range des Spiritueux:

Rotation Stock	Interprétation	Action conseillée
> 2,5	Produit très demandé, risque de rupture de stock	Augmenter le stock, pas de promo
0,9 - 2,5	Équilibre	Aucune action urgente
≤ 0,8	Ventes lentes, risque de surstock	Promotion, révision de la stratégie

Range du Sans Alcool:

Rotation Stock	Interprétation	Action conseillée
> 5,5	Produit très demandé, risque de rupture de stock	Augmenter le stock, pas de promo
2,1 - 5,5	Équilibre	Aucune action urgente
≤ 2	Ventes lentes, risque de surstock	Promotion, révision de la stratégie

Range de l'Huile:

Rotation Stock	Interprétation	Action conseillée
> 3	Produit très demandé, risque de rupture de stock	Augmenter le stock, pas de promo
1,3 - 3	Équilibre	Aucune action urgente
≤ 1,2	Ventes lentes, risque de surstock	Promotion, révision de la stratégie

Type Produit: (All)
 Date: (All)
 Segment: (Multiple val...)

Legend:
■ Overstock

Etat des stocks:

- 50 Propos. Promo
- Overstock
- Equilibre
- Rupture

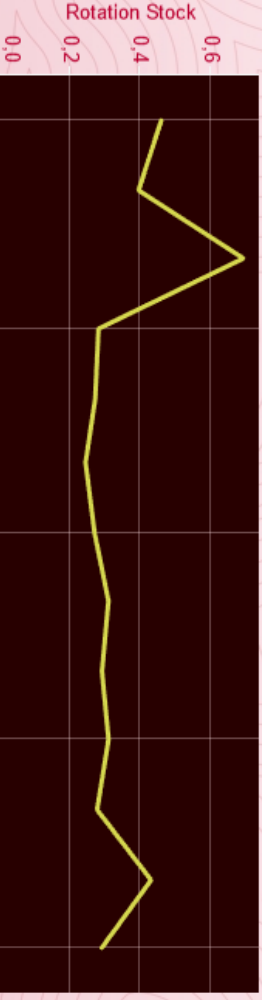
Etat des produits

Equilibre	36
Overstock	676
Rupture	1

Rotation Stock (avg)

0,2931

Rotation Stock dans le temps



Taux de Conversion

Le Taux_conversion mesure donc dans quelle mesure les commandes fournisseurs correspondent aux ventes. Un taux > 1 signifie qu'on a commandé plus que ce qui a été vendu (potentiel surstock). Un taux < 1 peut indiquer un risque de rupture de stock ou une sous-estimation de la demande.



