



# Optimisation de notre gestion de données & Analyse stratégique

par Aurélie RIVIERE



# Objectifs et **contexte**

Amélioration de nos outils d'analyse de donnée pour l'aide à l'analyse des ventes et à la gestion des stocks.

- ➡ Mise en place d'une méthode de collecte et de nettoyage des données
- ➡ Observations des corrections à apporter pour la mise en place d'un processus de travail
  - \*Travail sur un échantillon test de nos ventes sur le mois d'octobre*
- ➡ Premières analyses statistiques

Point de départ du futur projet  
de data visualisation



Octobre 2024



# Source des Données

Sources des données : Export au format excel

## ERP

- Identifiant produit
- Vendu en ligne
- Prix et prix d'achat
- Quantité et état du stock

+

## SITE WEB (Wordpress)

- Identifiant produit catalogue
- Nom, description et infos produit
- Informations complémentaires propre à Wordpress.

Identifiant ERP  $\neq$  Identifiant Wordpress

Création d'un fichier de liaison avec correspondance des identifiants produit du site et de la base ERP.



Intégration des données dans  
un **notebook Jupyter**.



Analyse des données avec  
**Python et ses bibliothèques**  
(Pandas, Plotly Express)

# Analyse Exploratoires des Données ERP

Exploration du fichier ERP afin de vérifier la qualité des données

Fichier ERP : 6 colonnes et 825 lignes

➔ Vérification d'identifiants en double

➔ Vérifications des stocks (statut et quantité)

La colonne stock statut indique si du stock est présent ou non. Cette information est liée à la quantité en stock.

- Correction des erreurs de stocks.
- Informations à valider avec l'équipe en charge du stock et remontée pour correction dans la base ERP

Articles avec des prix négatifs :

	product_id	price
151	4233	-20
469	5017	-8
739	6594	-9.1

➔ Vérification des données prix

Recherche des prix non renseignés, nul ou négatif

- Correction des prix négatifs en prix positif

➔ Vérification des Prix d'achat

Nos produits ont une fourchette de prix d'achat situé entre 2,74€ et 137,81€.

- Si tous les prix d'achat sont bien renseignés, nous avons relevé certaines incohérences.

	stock_status	stock_status_2	stock_quantity
4	outofstock	instock	3
398	instock	outofstock	0
449	outofstock	instock	-10
573	outofstock	instock	-1

	product_id	price	purchase_price
210	4355	12.65	77.48
391	4864	8.3	9.99
724	6324	92	99
817	7196	31	31.2

# Analyse Exploratoires des Données Web

Exploration du fichier WEB afin de vérifier la qualité des données

Fichier WEB : 29 colonnes et 1513 lignes

## ➔ Analyse des colonnes

- Suppression des colonnes vides ou non pertinentes

## ➔ Vérification des identifiants

Recherche des identifiants non renseignés, des doublons et des erreurs de structure

- Suppression lignes vides
- Suppression des identifiants en doublon, du à l'export des lignes de type attachement (images du produit)
- Identification de données non structurées

	sku	total_sales	tax_status	product_type	post_title	post_excerpt	post_type
272	13127-1	4	taxable	Vin	Clos du Mo...	Nez gracie...	product
1387	bon-cadeau-25-euros	7	taxable	nan	Bon cadeau...	<span styl...	product

Colonnes pertinentes conservées :

- sku
- total\_sales
- tax\_status
- product\_type
- post\_title
- post\_excerpt
- post\_type

# Fusion et consolidation des données

Le fichier liaison permet de faire correspondre les identifiants du fichier ERP avec les identifiants du fichier WEB

→ Travail réalisé manuellement

→ Fichier liaison : 825 lignes et 2 colonnes

➡ Vérification d'identifiants en double

➡ Vérification des correspondances

- Correspondances manquantes (dû au produit non vendu en ligne)

## Regroupement des informations des différentes sources

➡ Fusion du fichier ERP + LIAISON  
sur la base des identifiants produits afin de lier l'identifiant web

➡ Fusion du fichier ERP + WEB  
Utilisation d'une jointure externe et vérification des correspondance entre les lignes

Fichier fusionné : 825 lignes.



# Nettoyage de données et améliorations à apporter

## Types d'erreurs trouvés

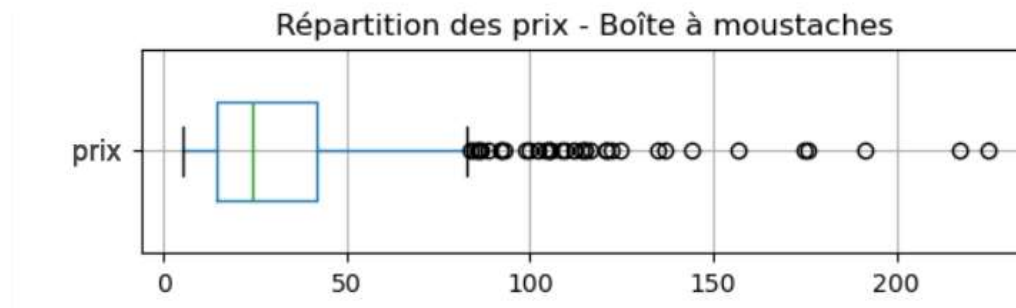
- Incohérence entre stock\_statut et stock\_quantity
- Saisie de stock négatif
- Saisie du prix en négatif
- Incohérence entre prix d'achat et prix de vente
- Ligne sans identifiants ni informations
- Export de doublons du aux attachements
- Erreur de codification des identifiants

## Améliorations à envisager

- Alignement des identifiants ERP et WEB
- Optimisation de l'export Wordpress
- Utilisation d'extensions d'exports personnalisables
- Amélioration des outils de saisie de données
- Amélioration de la saisie des données manuelle
- Remontée des erreurs trouvées aux services concernés

# Analyse univarié du prix

➔ Représentation en boîte à moustache de la dispersion des prix de vente



Moyenne du prix : 32,28€  
Ecart type du prix : 26,6€

➔ Calcul du Z-Score et des écarts interquartile

➔ Définition d'une  
Stratégie d'outliers  
(valeurs aberrantes ou extrêmes)  
Seuil supérieur à : 83.25€

	product_id	price	stock_quantity	purchase_price	total_sales	product_type	post_title
476	5026	86.8	101	50.13	9.0	Champagne	Champagne Agrapart & Fils Minéral Extra Br...
475	5025	112.0	136	68.60	6.0	Champagne	Champagne Agrapart & Fils L'Avizoise Extra...
208	4352	225.0	0	137.81	11.0	Champagne	Champagne Egly-Ouriet Grand Cru Millésimé 2008
214	4359	85.6	112	51.93	7.0	Champagne	Champagne Larmandier-Bernier Grand Cru Les Che...
635	5892	191.3	98	116.06	6.0	Champagne	Coteaux Champenois Egly-Ouriet Ambonnay Rouge ...
697	6126	135.0	138	80.33	5.0	Champagne	Champagne Gosset Célébris Vintage 2007



# Analyse des ventes en CA

## Méthode statistique employée

- ➔ Calcul du Chiffre d'affaire par produit -  $\text{prix} \times \text{total des ventes}$
- ➔ Calcul du Chiffre d'affaire total - Somme des CA

## Analyse du 20/80

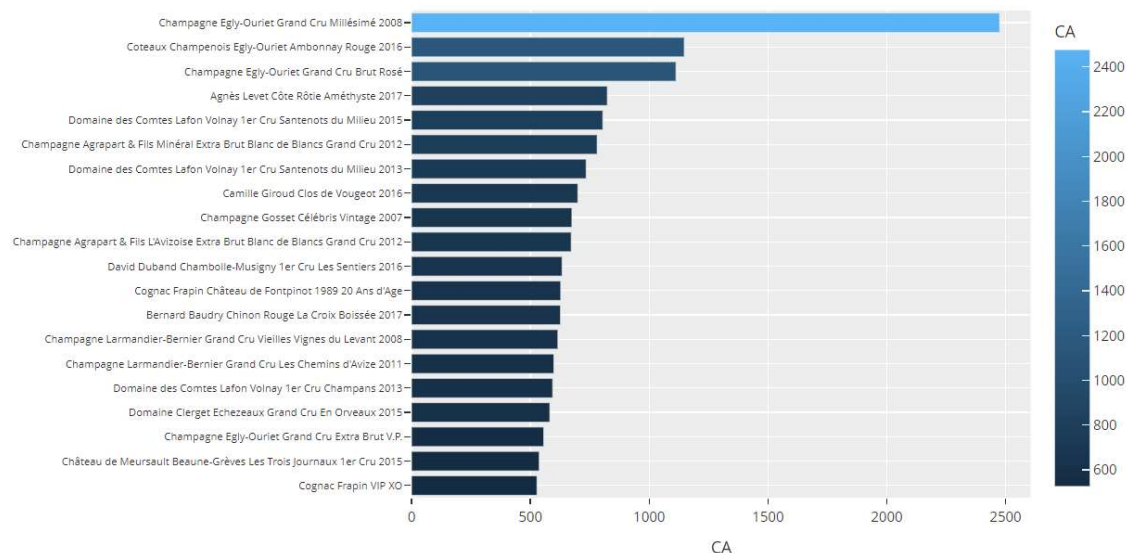
- ➔ Calcul de la part du CA -  $\text{CA Produit} / \text{CA Total}$   
Déterminée pour chaque produit par rapport au chiffre d'affaires total
- ➔ Somme Cumulative du CA

Chiffre d'affaire Total :  
**143 680 €**

Nbre d'articles représentant 80% du CA :  
**434 produits / 815 références**

Proportion des articles représentant 80% du CA :  
**52.61%**

Top 20 des produits  
avec le meilleur  
chiffre d'affaire

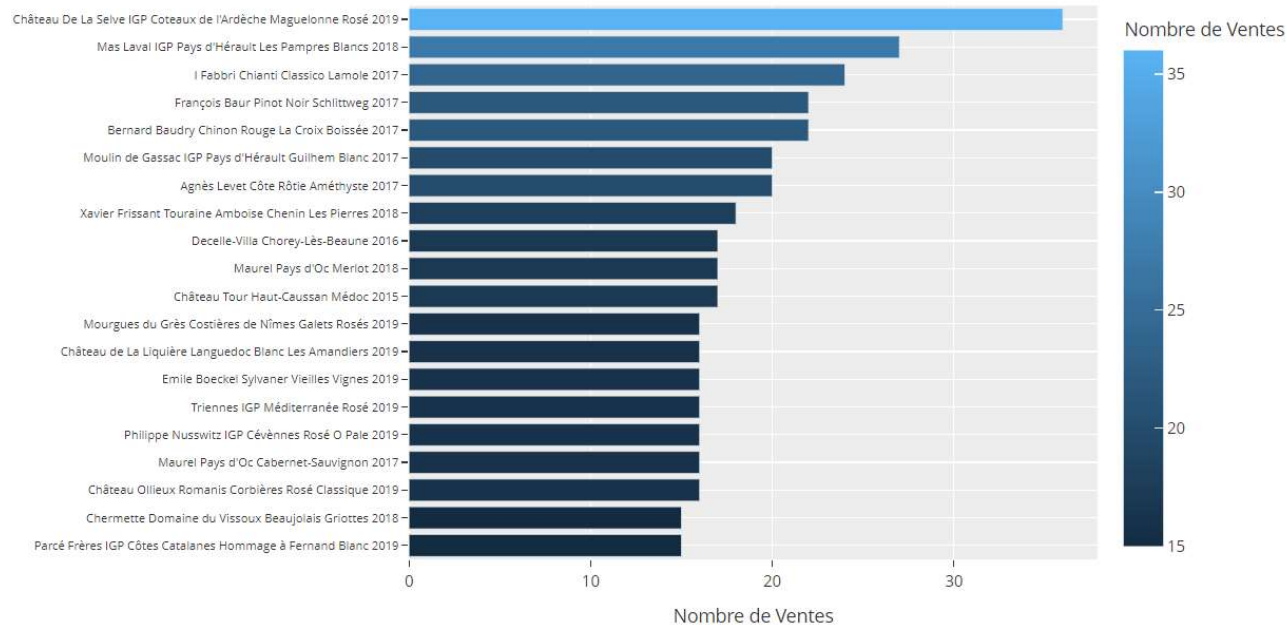


# Analyse des ventes en quantité

Nombre d'articles représentant  
80% des ventes en quantité :  
**433 produits**

Proportion des articles représentant  
80% des ventes en quantité :  
**52.48%**

Top 20 des produits  
Avec le meilleur nombre  
de ventes



# Analyses des Stocks

## ➔ Calcul de la rotation des stocks

Rotation des stocks =  $CA / \text{Quantité en stock}$

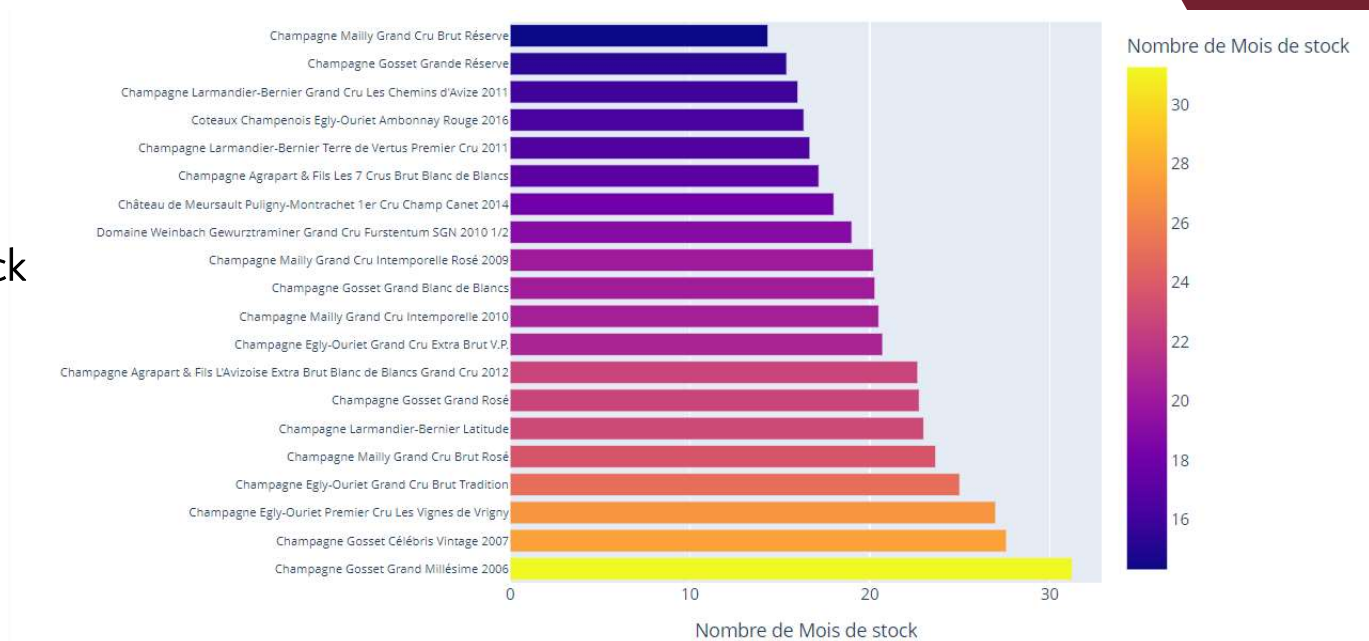
## ➔ Calcul des mois de stock

Mois de stock =  $\text{Quantité en stock} / \text{Ventes moyennes par mois}$

Valorisation totale des stocks :  
532 119.10 €

Nombre de produit en stock :  
17 822 bouteilles

## Produits ayant le plus de stock (valorisé en mois)

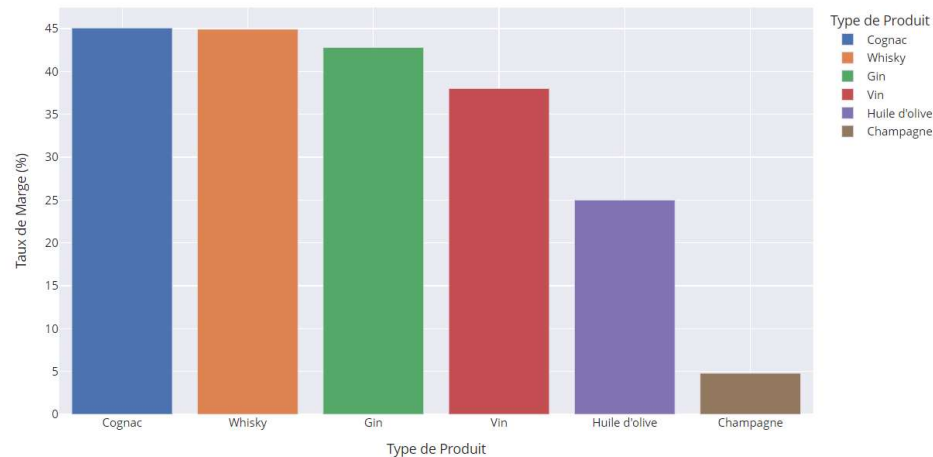


# Analyses du Taux de marge

➔ Calcul du Prix de vente HT

➔ Ajout d'une colonne Taux de marge  
Pourcentage du chiffre d'affaires qui reste après déduction des coûts d'achat ou de production

$$\text{Taux de marge} = (\text{PrixDeVente} - \text{CoûtAchat}) / \text{PrixDeVente} * 100$$



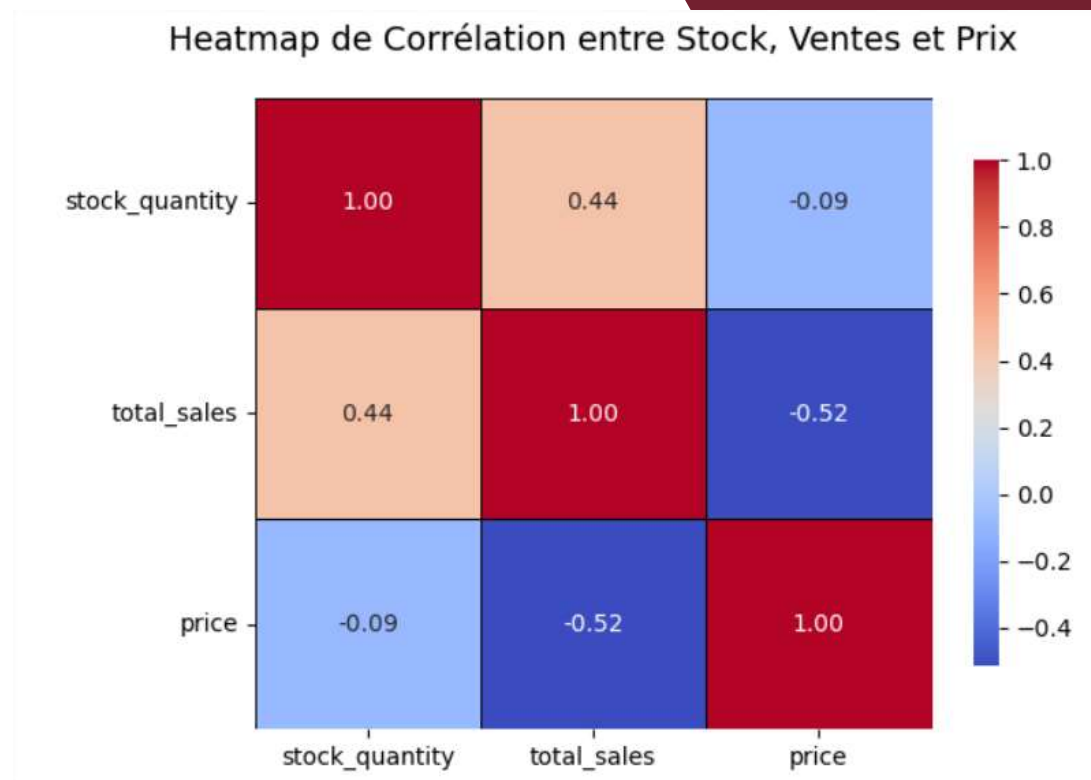
Taux de marge le plus faible :  
22.78%

Taux de marge le plus élevé :  
47.76%

Répartition des taux de marge  
par catégorie de produits

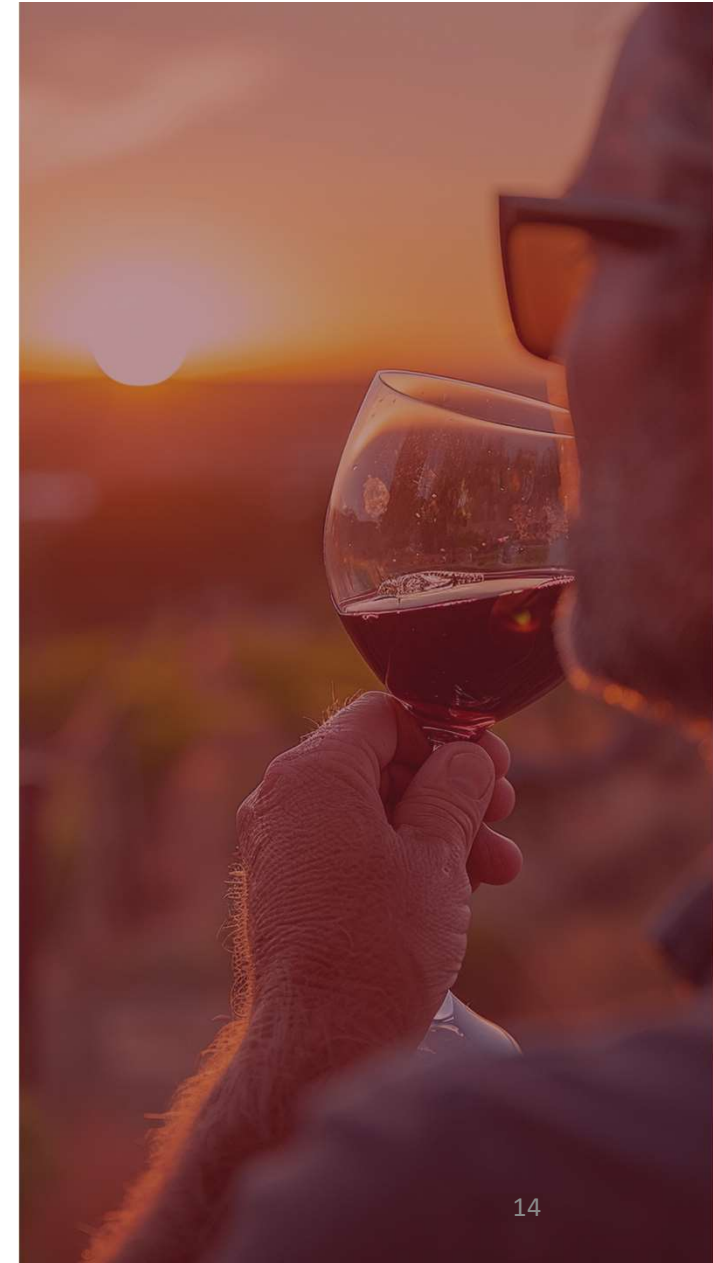
# Analyses des corrélations entre stock, ventes et prix

- ➔ Stock & Ventes : Corrélation : 0.44  
Assurer une stratégie de gestion de stock basée sur la disponibilité des produits à forte demande.
- ➔ Ventes & Prix : Corrélation : -0,52  
Sensibilité des Consommateurs au Prix
- ➔ Stock & Prix : Corrélation : -0,09  
Assurer une stratégie de gestion de stock basée sur la disponibilité des produits à forte demande.



## Actions pour la suite

- ➡ **Amélioration de la qualité des données :** Standardisation des données, élimination des doublons, procédure d'export.
- ➡ **Amélioration de la saisie des données :** Mise en place d'automatisation et formation des équipes.
- ➡ **Mise en place de tableaux de bord :** Suivi des indicateurs clés, détection des anomalies et des opportunités d'amélioration.
- ➡ **Maintenance :** Evaluation et amélioration des outils



**Merci !**  
**pour votre attention**