

# Optimisez la gestion des données d'une boutique avec R ou Python

Pauchet Kilian  
Data Analyste  
20/02/2026

# Environnement de travail initial

---



**ERP**

825 références (*stocks, prix d'achat*)



**Liaison**

*Table de correspondance*



**Web**

714 références (*Ventes, prix de vente*)



Respect RGDP, absence de données personnelles, données centrées sur les produits, données traiter au sein de l'U.E

*Suppression des doublons.*



*Harmonisation des colonnes (Id\_web -> Sku)*

*Conservation des colonnes utiles à nos analyses (product\_id, sku, price, stock\_quantity, total\_sales, product\_date)*

# Audits et qualité des données

---

*Erreur entre les différents fichiers.*



*Nom de colonnes clé différents.*

*91 champs orphelins entre les tables web et erp.*

*Présence de doublons 'sku' et 'id\_web'.*

*Gestion des cellules vides ou NaN.*

*Formatage des données au format attendu.*

*Incohérences des données.*



*Produits avec des stocks négatifs.*

*Produits avec des prix négatifs.*

*Problème d'affichage des produits qui sont InStock et OutOfStock.*

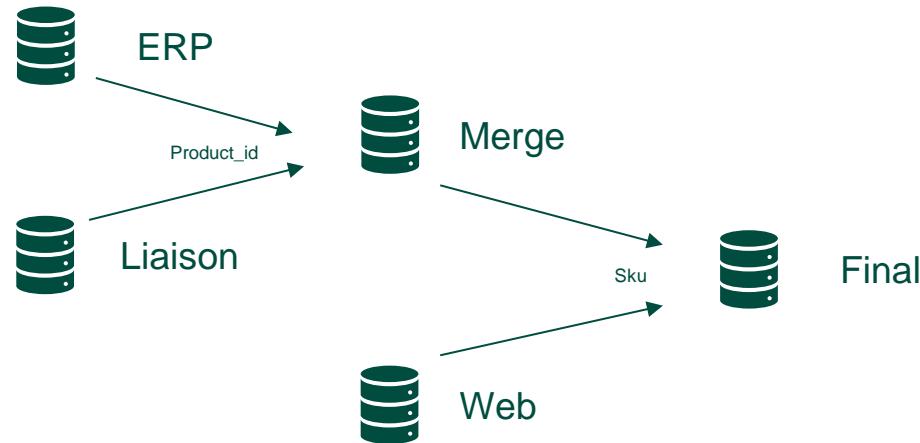
*Produits avec des prix d'achat supérieurs au prix de revente.*

# Méthodologie des jointures

---

Unicité des clés de jointure sku et product\_id

Inclusion des produits vendu absents du web grâce à un outer Join.



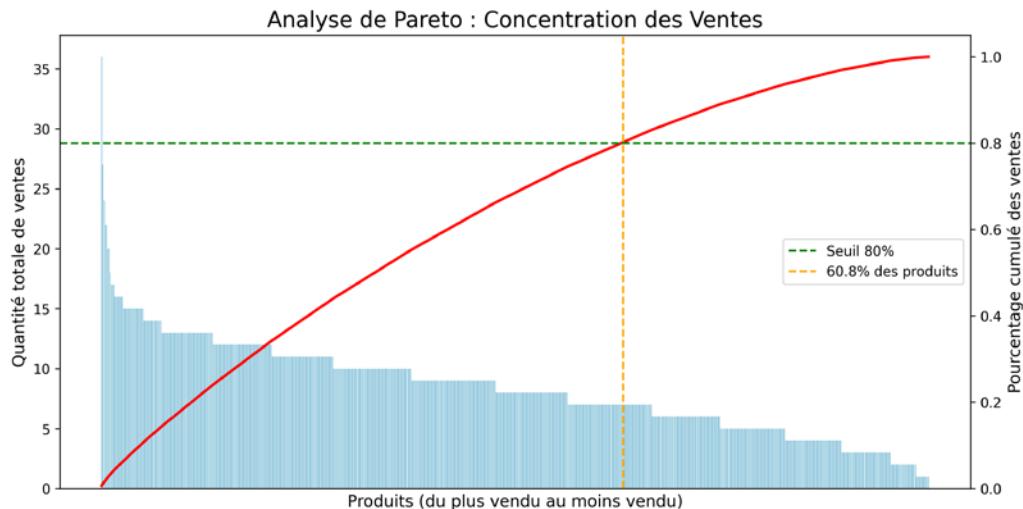
*Gestion des NaN pour les produits sans correspondances.*

# Analyse du CA

---

C.A calculé: 143 680,1 Euros

Plus de 60 % (434 références) du catalogue représente 80 % des ventes.



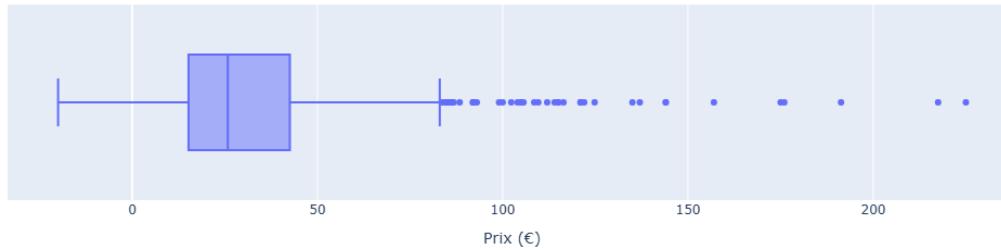
# Analyses des prix, des marges et des Outliers.

---

*Cœur de gamme de prix est au alentour des 25-30 euros.*

*Les outliers sont les prix supérieurs a 84 euros selon la méthode interquartile. Les prix 'Aberrent' sont en réalité nos produits les plus prestigieux.*

Répartition des prix (Plotly Express)

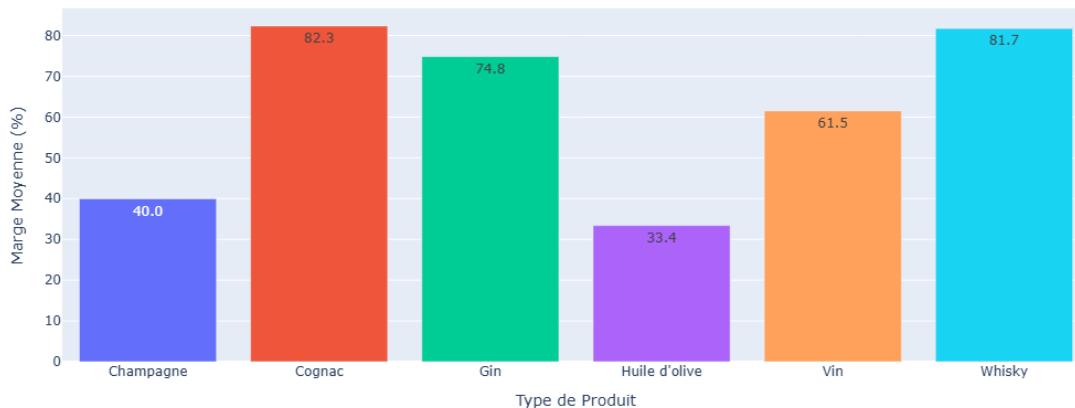


# Analyses des prix, des marges et des Outliers.

---

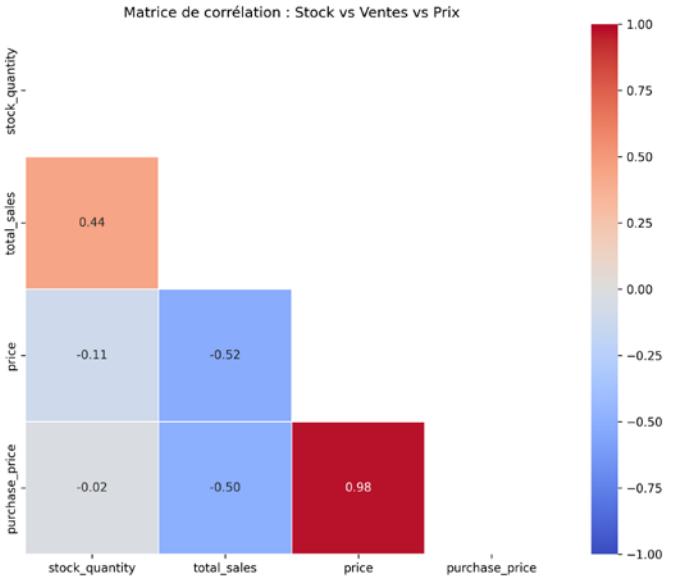
Taux de marges sont en général plus élevé pour les spiritueux et alcool fort.

Rentabilité moyenne par type de produit



# Analyses multivariées.

---



*On a conclu à 3 corrélations entre le total des ventes, le prix des produits et les stocks de produits.*

*Corrélation forte entre le prix et le prix d'achat, les prix sont cohérents et permettent une optimisation des marges de ventes.*

*Corrélation faible entre le prix et le total des ventes, les prix n'ont pas d'influence sur le nombre de vente, les clients cherchent des produits de qualités.*

*Corrélation faible et positive entre le stock de produit et le nombre de ventes, le stock de certains produits n'est peut-être pas en adéquation avec la demande client.*

# Gestion et rotation des stocks

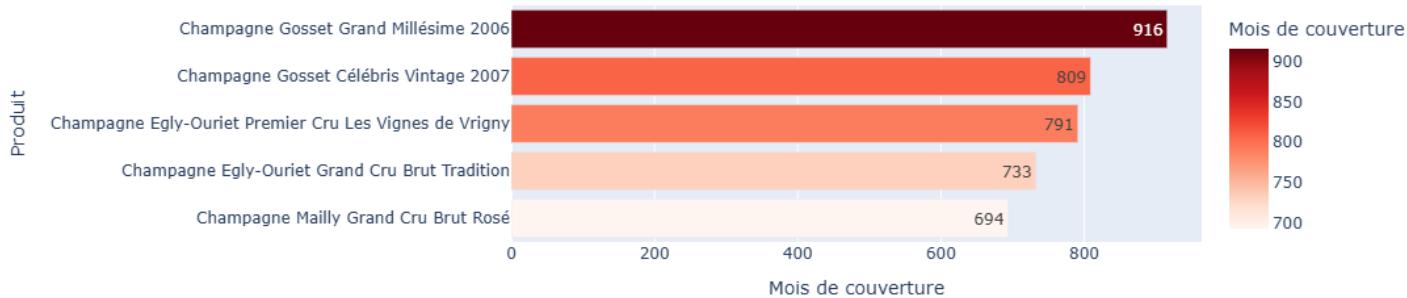


Final

Calcul du nombre de mois de la période de vente via la fonction days()

*(Date la plus ancienne - date la plus récente).days / 30,44 (nombre de jour par mois en moyenne)*

Top 5 des produits avec le plus de stock dormant (en mois)



*On exclut délibérément les ventes = 0*

# Le livrable Final

Création d'un fichier Excel grâce à la fonction to\_excel

Tableau synthétique des données pour les produits de l'ERP et du Web

Création de colonnes d'indicateur calculé.

Fichier standardisé et utilisable par tous les collaborateurs de l'entreprise.

stock_status	purchase_price	sku	total_sales	post_date	post_title	product_type	post_type	_merge	CA	z_score	ca_par_article	ventes_mensuelles	mois_de_stock	valeur_stock_vente	valeur_stock_achat	price_ht	taux_marge
instock	4,22	38	10	2018-04-18 12:25:58	Emile Boeckel Crémant Brut Blanc de Blancs	Vin	product	both	86	-0,860030352	86	0,341255605	76,1892247	223,6	109,72	7,166666667	69,82622433
instock	20,12	41	6	2018-04-14 12:01:43	Marcel Windholtz Eau de Vie de Marc de Gewurztraminer	Vin	product	both	246	0,314038706	246	0,204753363	53,72317127	451	221,32	34,166666667	69,81444665
instock	24,86	304	8	2018-02-13 12:57:44	Champagne Gosset Grande Réserve	Champagne	product	both	312	0,241565307	312	0,273004484	450,5420499	4797	3057,78	32,5	30,73209976
instock	27,18	523	0	2019-04-06 15:25:58	Cognac Normandin Mercier VFA	Cognac	product	both	0	0,998912323	0	0	0	778,7	353,34	49,916666667	83,65219524
instock	13,78	531	8	2018-07-18 15:58:02	Champagne Petit Lebrun & Fils Blanc de Blancs Grand Cru	Champagne	product	both	180	-0,356340231	180	0,273004484	278,3837057	1710	1047,28	18,75	36,0676343
instock	38,49	791	1	2018-04-14 11:54:47	Marcel Windholtz Eau de Vie de Baie de Houx	Vin	product	both	76,8	1,611312542	76,8	0,034125561	29,30354796	76,8	38,49	64	66,27695505
instock	33,21	793	1	2018-04-14 11:52:47	Marcel Windholtz Eau de Vie de Coligny	Vin	product	both	62,4	1,089504071	62,4	0,034125561	29,30354796	62,4	33,21	52	56,57934357
instock	31,92	798	8	2018-04-14 12:05:04	Marcel Windholtz Eau de Vie de Kirsch d'Alsace Réserve Particulière	Vin	product	both	499,2	1,089504071	499,2	0,273004484	43,95532194	748,8	383,04	52	62,90726817
instock	21,07	802	5	2018-04-14 11:48:19	Marcel Windholtz Eau de Vie de Marc de Muscat	Vin	product	both	198	0,263307327	198	0,170627803	46,88567674	316,8	168,56	33	56,62078785
instock	27,61	804	3	2018-04-14 11:59:32	Marcel Windholtz Eau de Vie de Mirabelle d'Alsace Réserve Particulière	Vin	product	both	157,2	0,727137078	157,2	0,102376682	48,83924661	262	138,05	43,666666667	58,1525776
instock	28,62	805	2	2018-04-14 11:45:24	Marcel Windholtz Eau de Vie de Poire Williams Réserve Particulière	Vin	product	both	116,6	0,940933604	116,6	0,068251121	29,30354796	116,6	57,24	48,583333333	69,75308642
instock	27,88	807	2	2018-04-14 11:43:35	Marcel Windholtz Eau de Vie de Pruneille Sauvage	Vin	product	both	105,8	0,745255428	105,8	0,068251121	58,60709593	211,6	111,52	44,083333333	58,1181253

# Conclusion.

---

Nettoyage des données entre les fichiers

Réconciliation des données ERP et Web

Création d'un support de pilotage des produits web et ERP.

CA Web : 143 680,1 euros

60% du catalogue génère 80% du CA

Suggestion stratégique :

Lancer une opération de déstockage pour les produits dont la rotation des stocks est faible.

Pérenniser ce fichier centralisé pour en faire un tableau de bord mensuel.