

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

Thomas Hébert
Data Analyst
13/09/2024



Contexte

Phase 1 : Agréger les fichiers pour exploiter les données



ERP



Site Web



Table de
liaison

Phase 2 : Analyser les données

Analyse exploratoire des données

ERP

Colonne	Données non_nulles	Type
product_id	825	int
onsale_web	825	int
price	825	float
stock_quantity	825	int
stock_status	825	object
purchase_price	825	float

Table de liaison

Colonne	Données non_nulles	Type
id_web	734	object
product_id	825	int

Site Web

Colonne	Données non_nulles	Type
sku	1428	object
virtual	1513	int
downloadable	1513	int
rating_count	1513	int
average_rating	1430	float
total_sales	1430	float
tax_status	716	object
tax_class	0	float
post_author	1430	float
post_date	1430	date
post_date_gmt	1430	date
post_content	0	float
product_type	1429	object
post_title	1430	object
post_excerpt	716	object
post_status	1430	object
comment_status	1430	object
ping_status	1430	object
post_password	0	float
post_name	1430	object
post_modified	1430	date
post_modified_gmt	1430	date
post_content_filtered	0	float
post_parent	1430	float
guid	1430	object
menu_order	1430	float
post_type	1430	object
post_mime_type	714	object
comment_count	1430	float

Analyse exploratoire des données

ERP

Colonne	Données non_nulles	Type
product_id	825	int
onsale_web	825	int
price	825	float
stock_quantity	825	int
stock_status	825	object
purchase_price	825	float

Erreur de type

Valeurs aberrantes

Valeurs incohérentes

Données non pertinentes

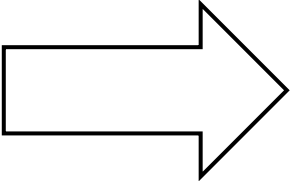
Colonne	Données non_nulles	Type
product_id	821	object
onsale_web	821	bool
price	821	float
stock_quantity	821	int
purchase_price	821	float

Analyse exploratoire des données

Site Web

Colonne	Données non_nulles	Type
sku	1428	object
virtual	1513	int
downloadable	1513	int
rating_count	1513	int
average_rating	1430	float
total_sales	1430	float
tax_status	716	object
tax_class	0	float
post_author	1430	float
post_date	1430	date
post_date_gmt	1430	date
post_content	0	float
product_type	1429	object
post_title	1430	object
post_excerpt	716	object
post_status	1430	object
comment_status	1430	object
ping_status	1430	object
post_password	0	float
post_name	1430	object
post_modified	1430	date
post_modified_gmt	1430	date
post_content_filtered	0	float
post_parent	1430	float
guid	1430	object
menu_order	1430	float
post_type	1430	object
post_mime_type	714	object
comment_count	1430	float

Valeurs manquantes
Données non pertinentes
Erreur de type
Erreur de formatage
Erreur de contradiction
Valeurs aberrantes



Colonne	Données non_nulles	Type
sku	714	object
total_sales	714	int
product_type	714	object
post_title	714	object

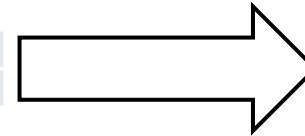
Analyse exploratoire des données

Table de liaison

Colonne	Données non_nulles	Type
id_web	734	object
product_id	825	int



Erreur de type
Valeurs manquantes

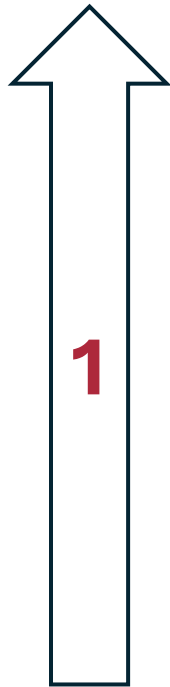


Colonne	Données non_nulles	Type
sku	734	object
product_id	825	object

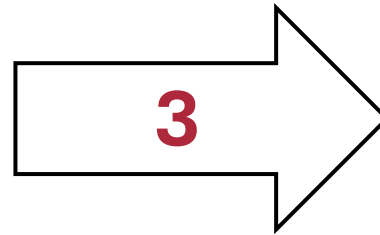
Fusion des données

ERP

Colonne	Données non_nulles	Type
PRODUCT_ID	821	object
onsale_web	821	bool
price	821	float
stock_quantity	821	int
purchase_price	821	float



Colonne	Données non_nulles	Type
PRODUCT_ID	821	object
onsale_web	821	bool
price	821	float
stock_quantity	821	int
purchase_price	821	float
sku	732	object



Merge

Colonne	Données non_nulles	Type
product_id	732	object
onsale_web	732	bool
price	732	float
stock_quantity	732	int
purchase_price	732	float
sku	732	object

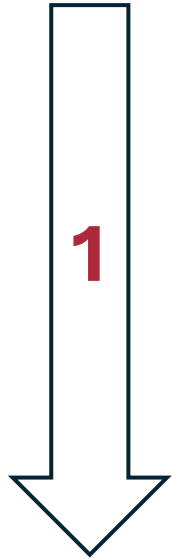
Table de liaison

Colonne	Données non_nulles	Type
sku	734	object
PRODUCT_ID	825	object

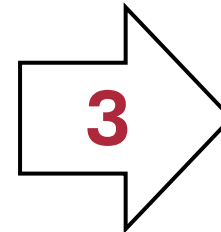
Fusion des données

Merge

Colonne	Données non_nulles	Type
product_id	732	object
onsale_web	732	bool
price	732	float
stock_quantity	732	int
purchase_price	732	float
SKU	732	object



Colonne	Données non_nulles	Type
SKU	714	object
total_sales	714	int
product_type	714	object
post_title	714	object
product_id	713	object
onsale_web	713	object
price	713	float
stock_quantity	713	float
purchase_price	713	float

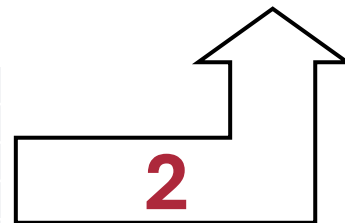


Final

Colonne	Données non_nulles	Type
product_id	713	object
sku	713	object
post_title	713	object
product_type	713	object
price	713	float
purchase_price	713	float
stock_quantity	713	int
total_sales	713	int

Site Web

Colonne	Données non_nulles	Type
SKU	714	object
total_sales	714	int
product_type	714	object
post_title	714	object

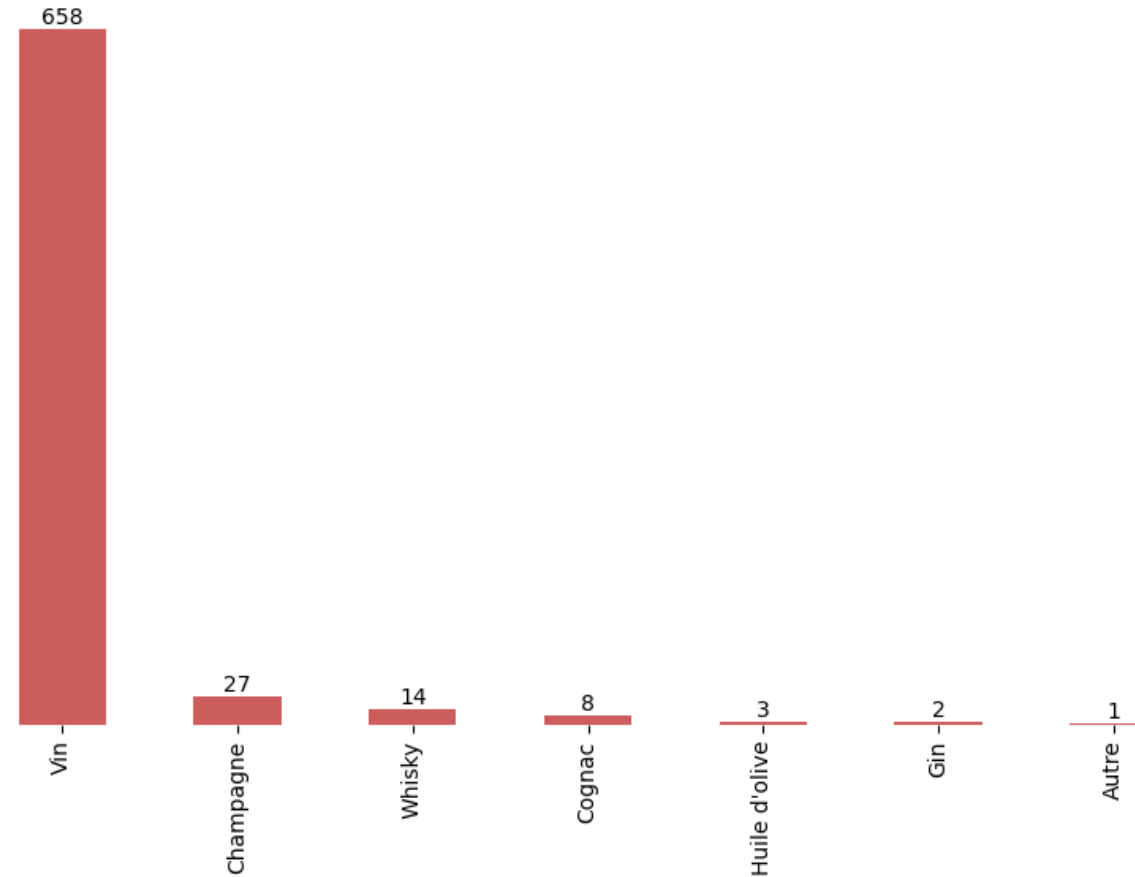


Recommandations pour l'ERP

Erreurs	Suggestions
Erreur de type/formatage/irrégularité	Imposer un format de données
Valeur aberrante	Définir une règle (pas de valeurs négatives ou confirmation au-delà d'une plage de valeurs)
Valeur incohérente/erreur de contradiction	Définir une règle (selon la valeur d'une autre variable)
Valeur manquante	Imposer le remplissage
erreur lexicale	Créer un dictionnaire
doublon	Définir une règle (clé unique)

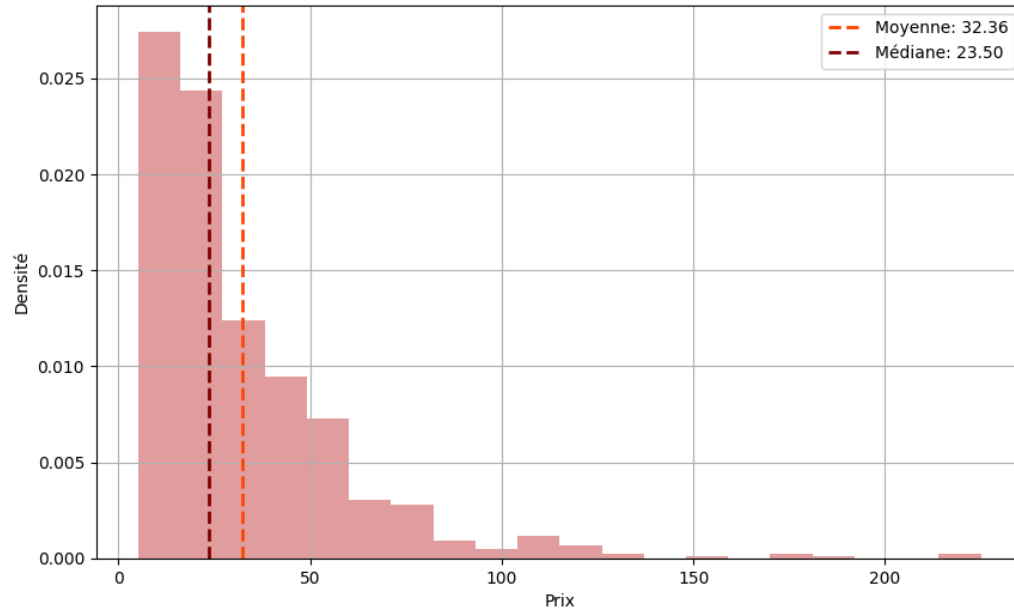
Analyses des types de produits

Nombre de Références par Type de Produit

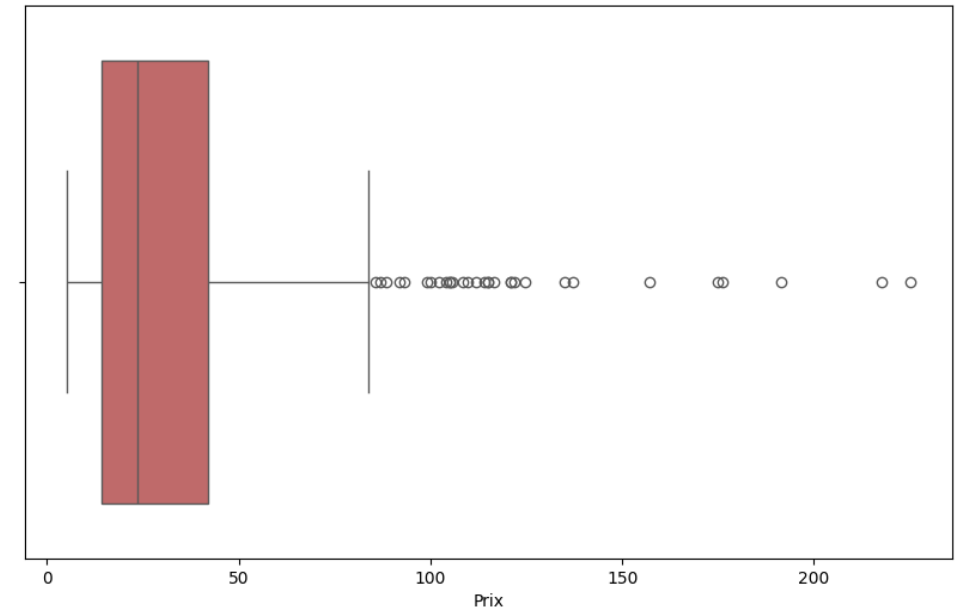


Analyses des prix

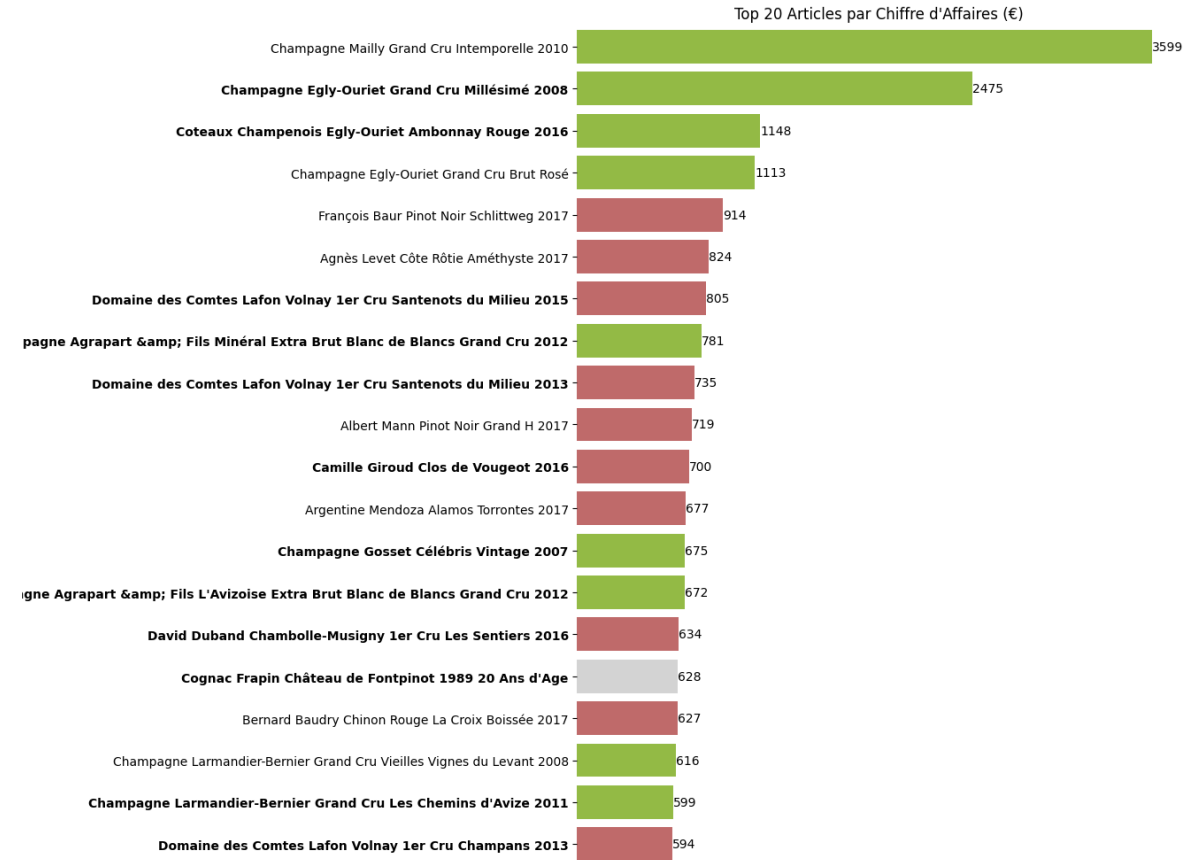
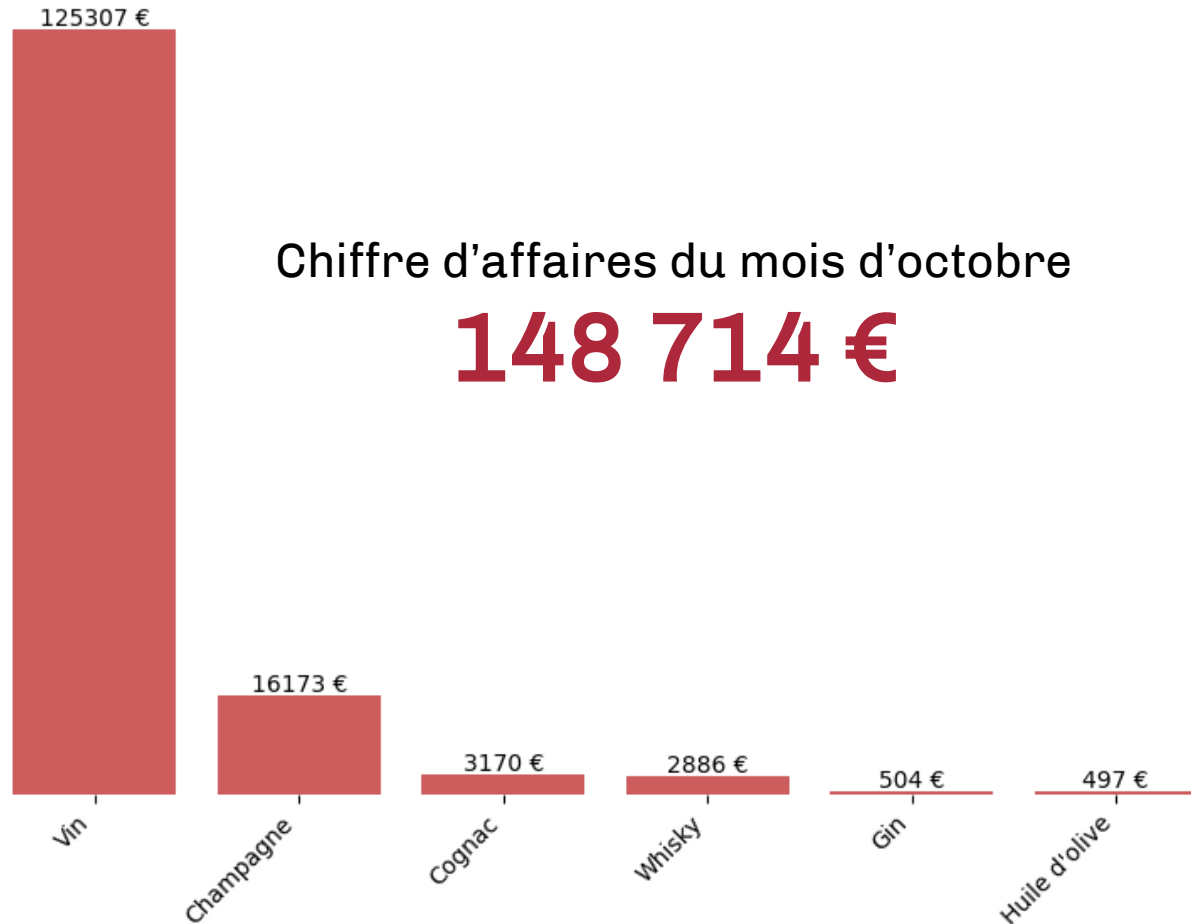
Histogramme des prix avec lignes de moyenne et médiane



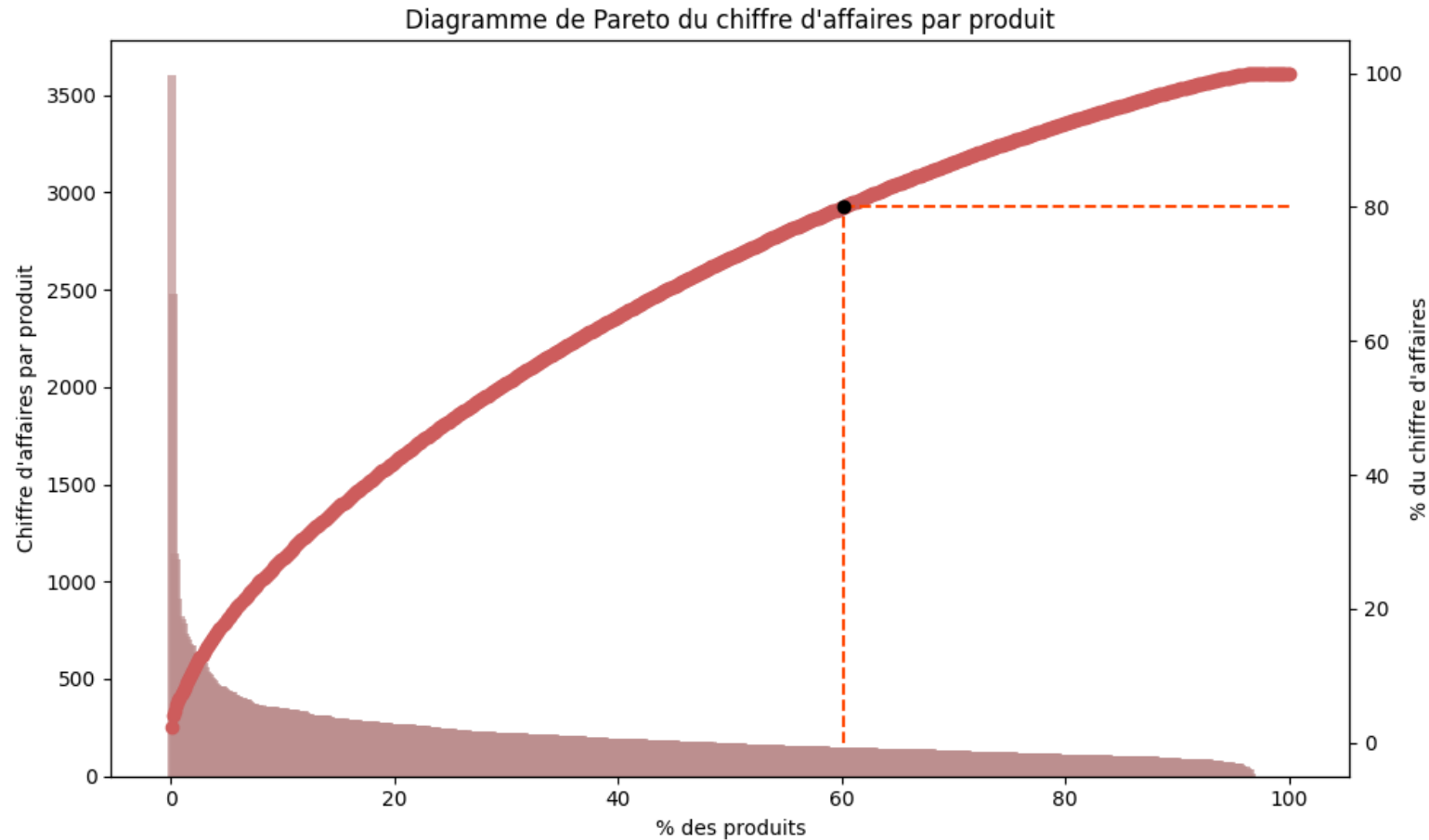
Répartition des Prix des Produits



Analyses du chiffre d'affaires

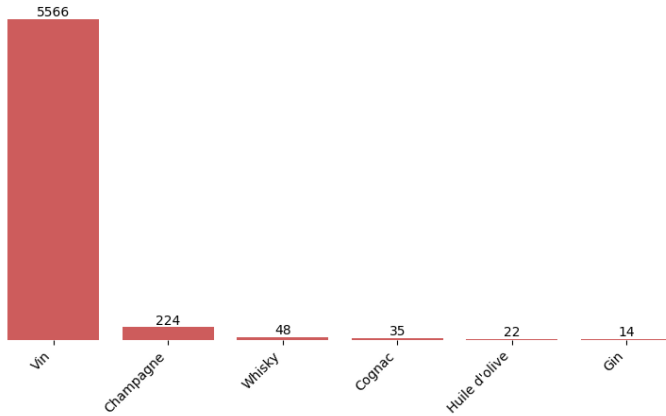


Analyses du chiffre d'affaires

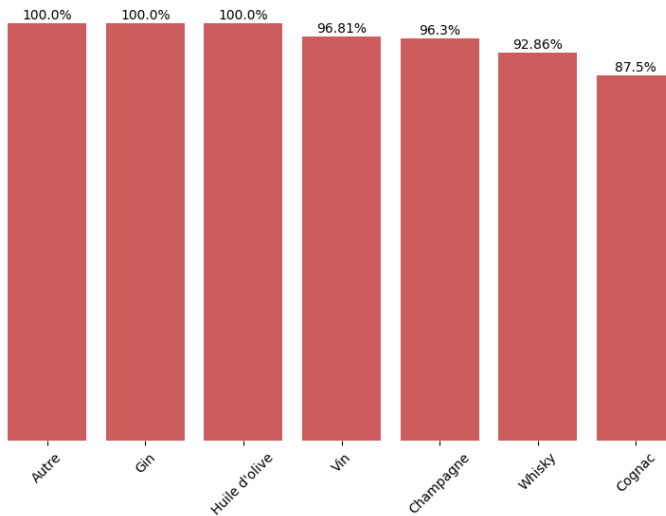


Analyses des ventes

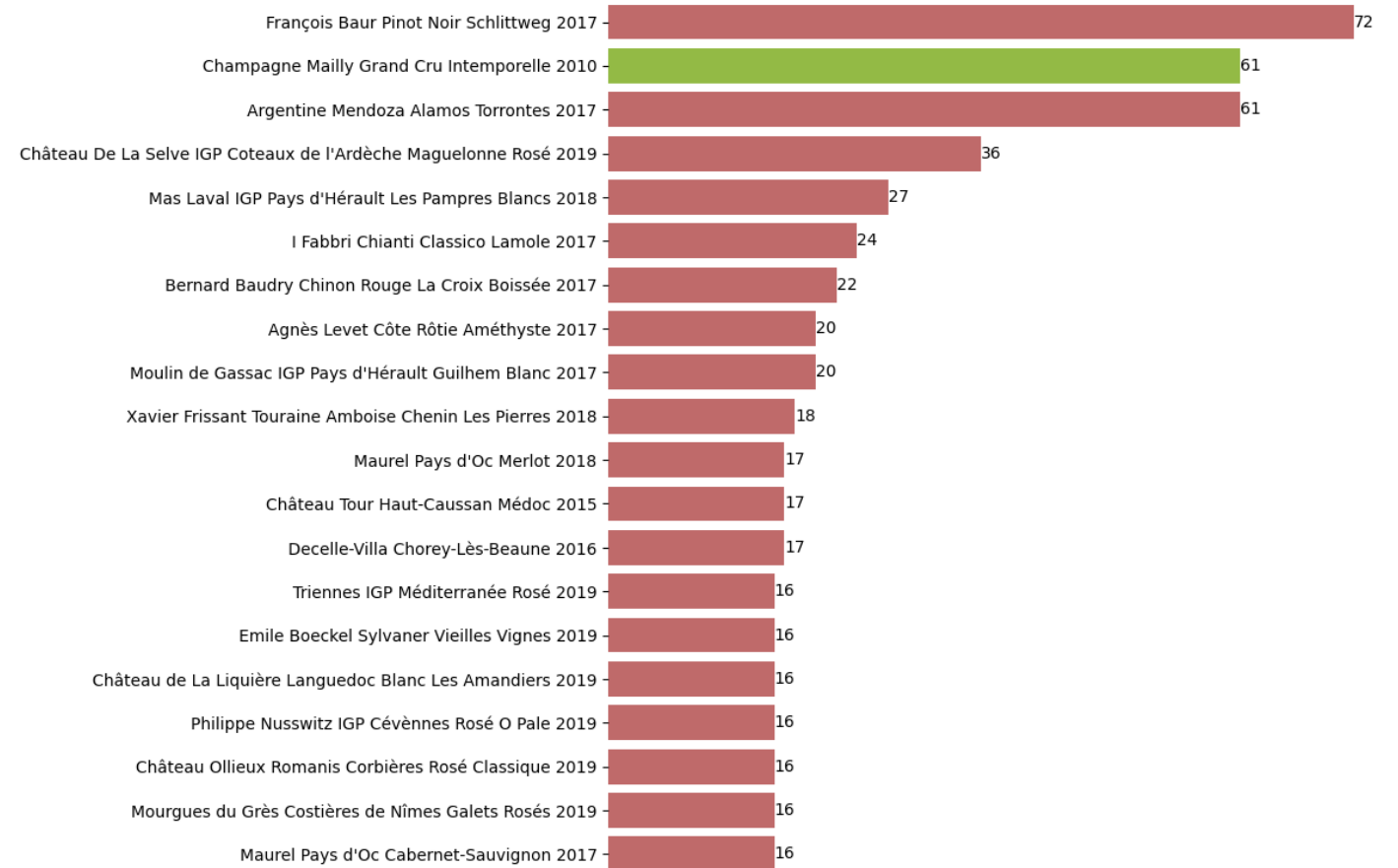
Quantité vendue par type de produit



Pourcentage de références ayant généré des ventes par catégorie

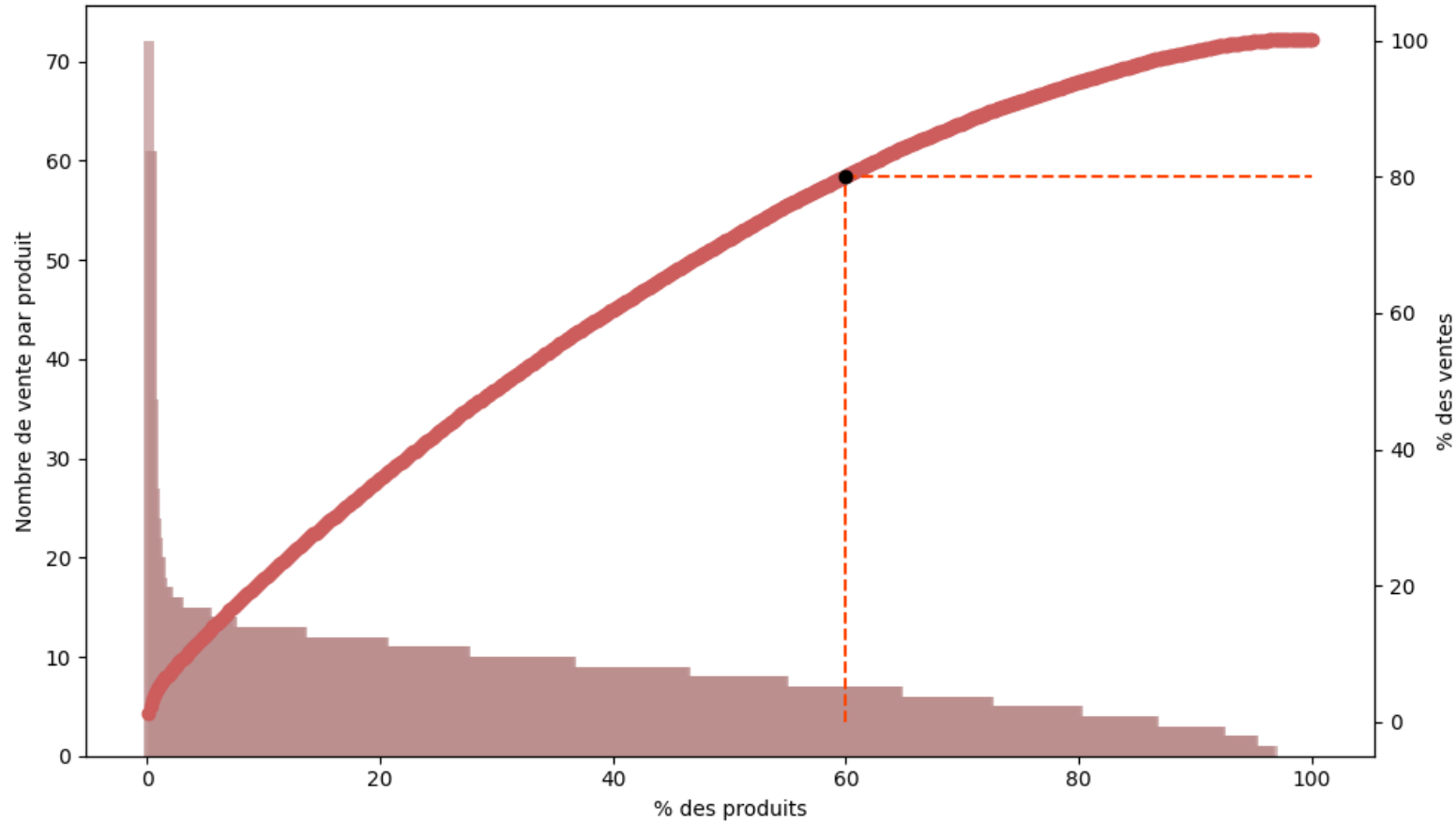


Top 20 Articles par quantités vendues



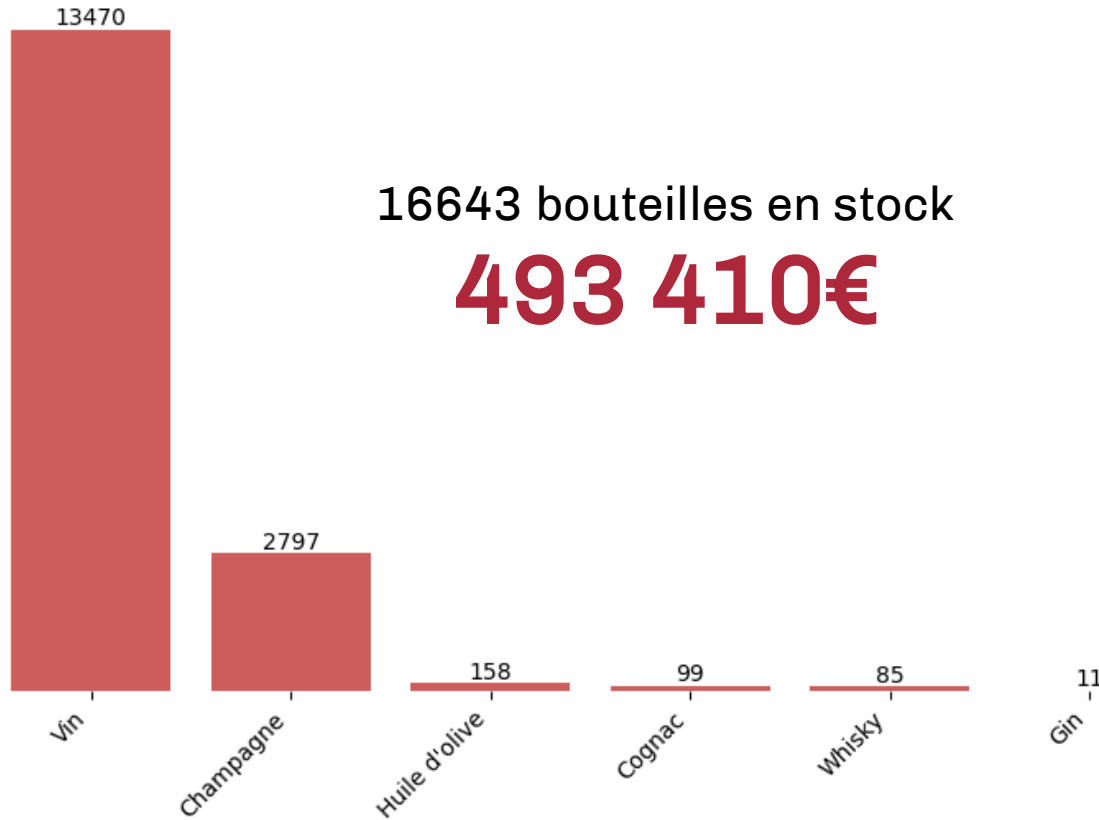
Analyses des ventes

Diagramme de Pareto des ventes par produit



Analyses des stocks

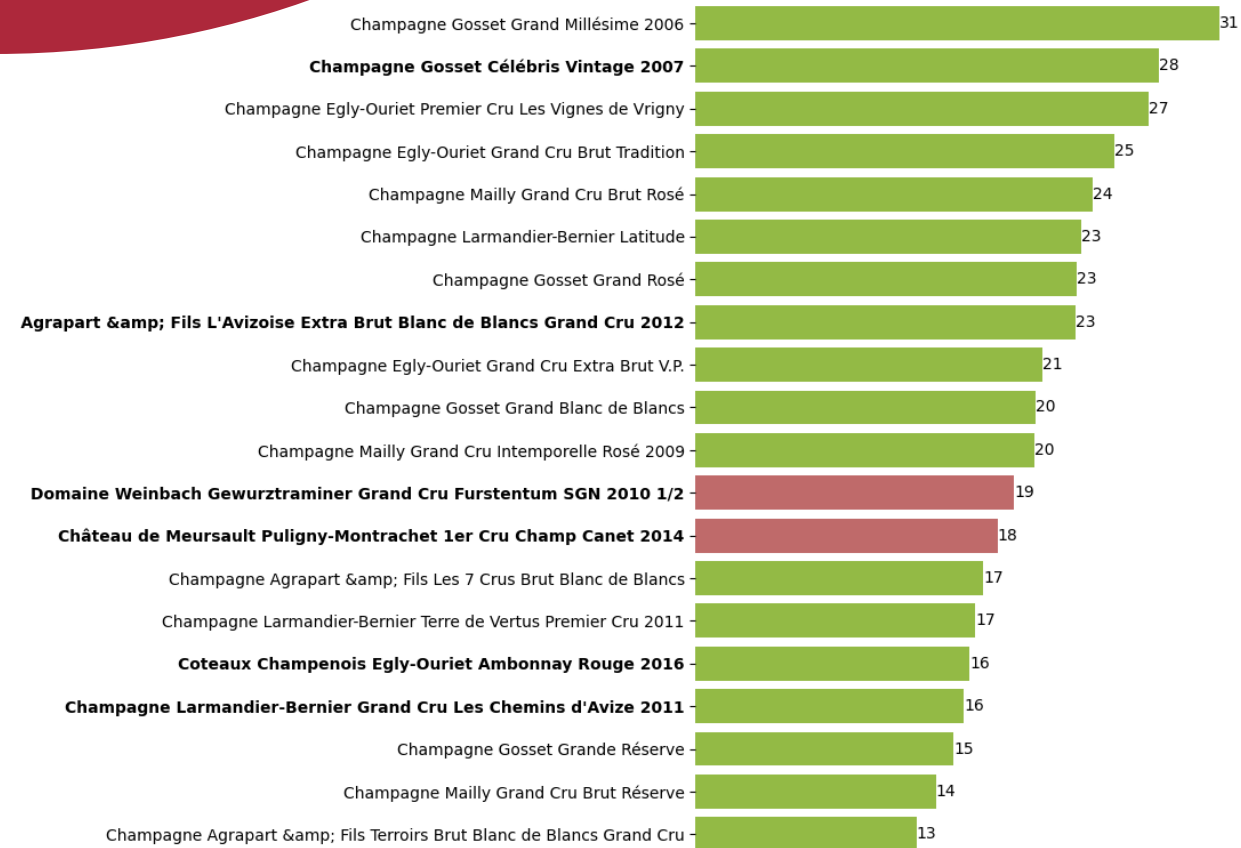
Stock par type de produit



16643 bouteilles en stock

493 410€

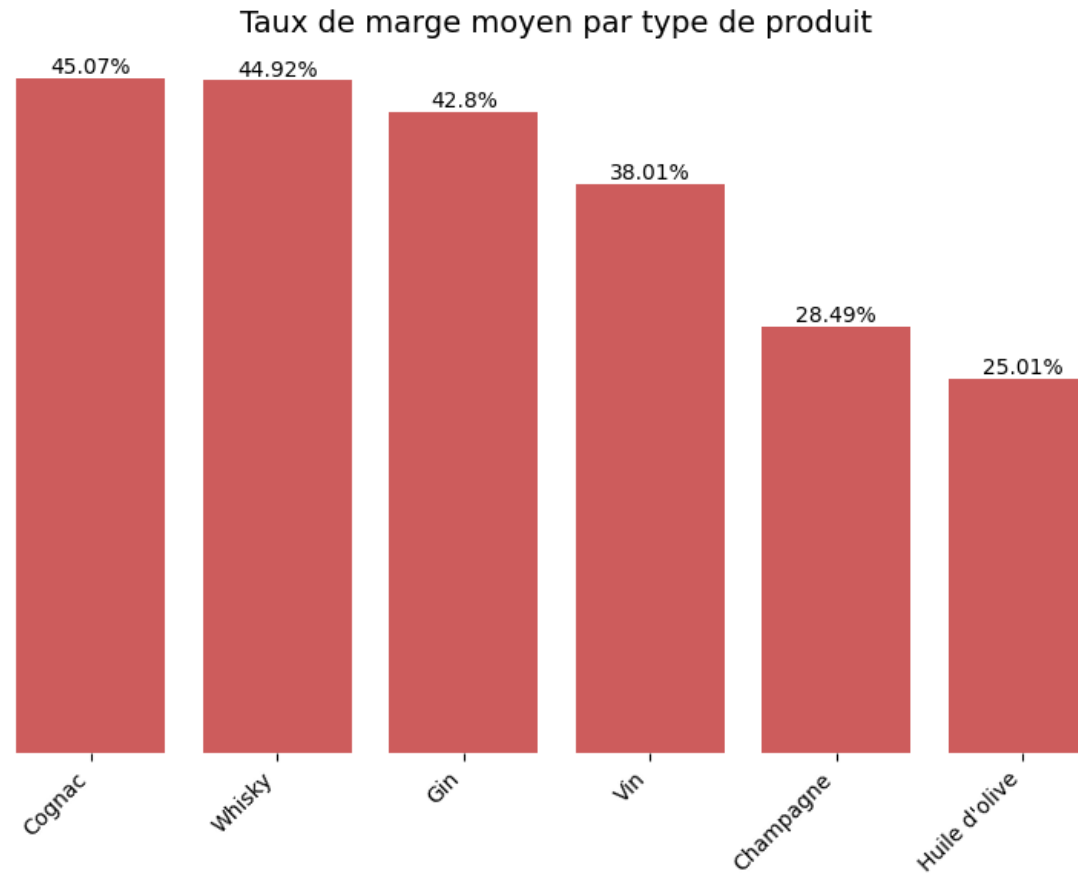
Flop 20 Articles qui ont le plus de mois de stock



Taux de rotation global du site

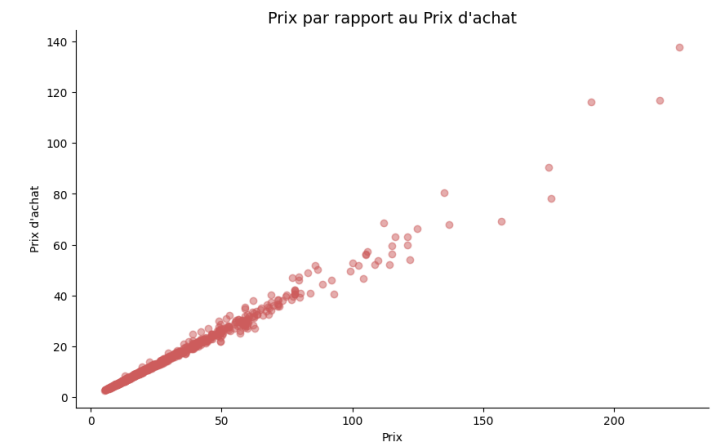
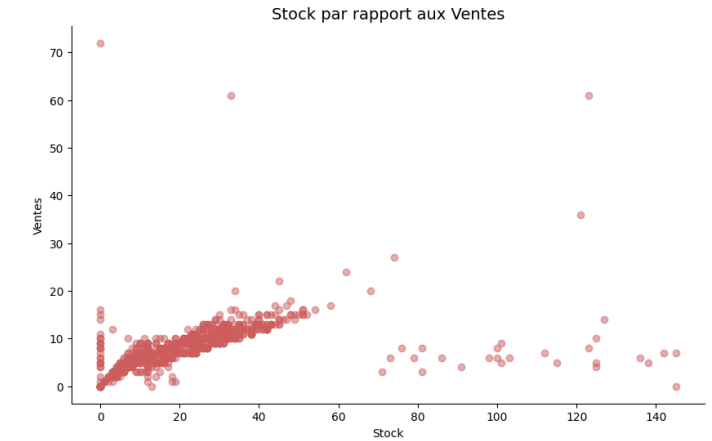
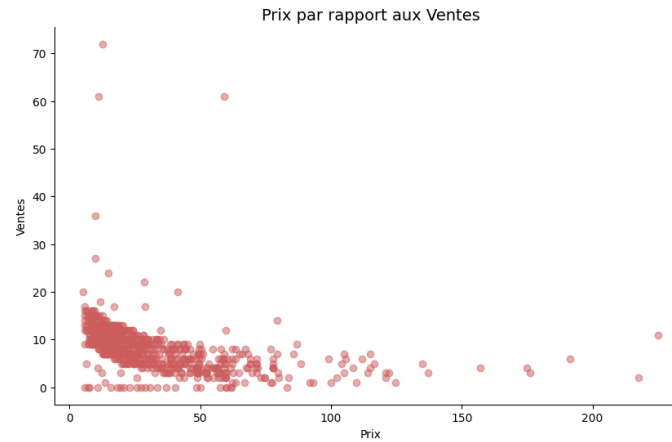
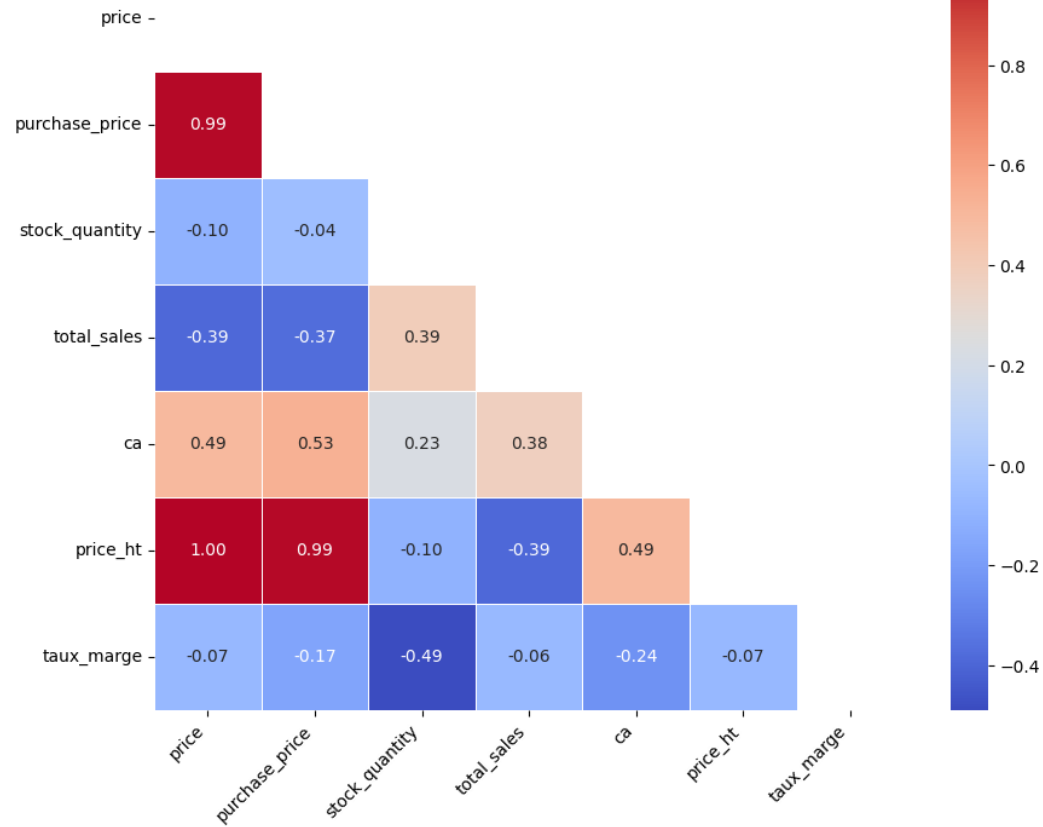
3,32

Analyses du taux de marge



Corrélations entre les variables

Heatmap des corrélations



Conclusion d'analyses

Les analyses sont limitées par des données provenant uniquement du mois d'octobre. Sur cette base, nous pouvons conclure que :

- Le vin est un marché globalement performant.
- Le champagne, bien que figurant parmi les meilleurs contributeurs au chiffre d'affaires, a une faible rotation et occupe une part importante du stock.
- Les spiritueux peinent à performer, ce qui remet en question leur pertinence dans l'offre du catalogue.
- L'huile d'olive ne représente que 0,3 % du chiffre d'affaires, avec un impact très limité sur les performances globales, même suggestion que les spiritueux.
- Moins le prix est élevé plus nous faisons de ventes
- Les prix sont cohérents
- Nous avons trop de stock et/ou peu de ventes sur certains produits

Pour approfondir, nous pourrions analyser chaque marché individuellement, avoir accès à des données anonymisées sur les habitudes d'achat des clients. Des données trimestrielles ou annuelles permettraient d'affiner les analyses et de mettre en évidence des schémas plus pertinents.