

Présentation CODIR Analyse des données Bottleneck

RAMAT Romain

PROJET 6

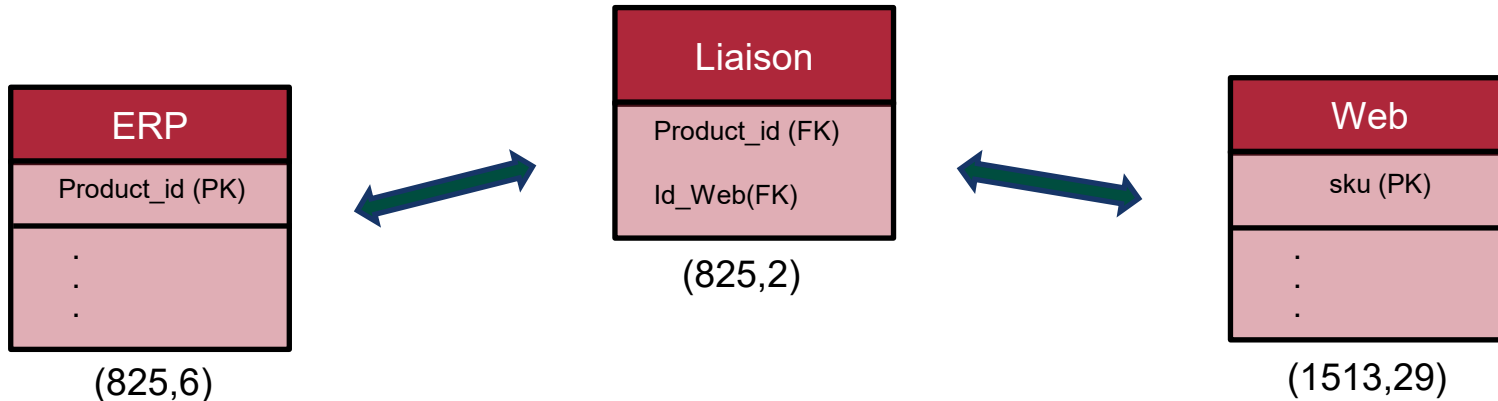
01 Mars 2025

Contexte

- L'utilisation d'outils artisanaux au sein de Bottleneck rendent complexe l'analyse des données et la gestion des stocks.
 - Présentation de la structure de données
 - Analyse des points de faiblesse
 - Présentation des résultats
 - Préconisation d'amélioration
- Pour la RGPD : collecte d'information sur les « posts » utilisateurs, il convient de s'assurer de la mise en place:
 - Du consentement de l'utilisateur
 - De la mise en place d'un droit à l'oubli

Analyses Exploratoires des Données

Composition des données de la base Bottle Neck



Les commandes et fonctions utilisées

- Le nettoyage des données

- `.isnull()` / `.isna()`
- `.notnull()` / `.notna()`
- `.is_unique`
- `.duplicated()`
- `.value_counts()` / `value_counts(normalize = True)`
- `.dropna`
- `.apply(lambda x: isinstance(x, (int)))`
- `.info()`

- Les analyses des données

- `.mean()`
- `.std()`
- `.var()`
- `.mode()`
- `.median()`
- `.skew()`
- `.kubertosis`
- `.describe()`
- `.value_counts()`
- `.hist()`

Fusion ou consolidations des données

Pré-Analyse des données de la table ERP:

ERP

Product_id (PK)

(825,6)

Colonne	Commentaire
Product_id	<ul style="list-style-type: none">Données uniques (clé primaire table ERP)
Stock Status / Stock Quantity	<ul style="list-style-type: none">Incohérences relevées entre le nombre de stock et l'état de stock associé -> CorrigéPrésence de stocks négatifs (2) -> Remis à 0
Price / Purchase Price	<ul style="list-style-type: none">7 produits vendus sous le prix d'achat (price < purchase price)3 prix de ventes négatifs (price < 0)Grande variance des prix
On Sale web	<ul style="list-style-type: none">716 articles en vente sur internet sur les 825 soit 86,7 %

Fusion ou consolidations des données

Pré-Analyse des données de la table Liaison:

Liaison

Product_id (FK)

Id_Web(FK)

(825,2)

Colonne	Commentaire
Product_id	<ul style="list-style-type: none">Données uniques
Id_web	<ul style="list-style-type: none">91 valeurs nulles (ventes non-web)Nom différent de la base web (sku)3 valeurs de type string « bon-cadeau-25-euros, 13127-1, 14680-1 »1 valeur id_web non-référencée en vente web « 14680-1 »

Fusion ou consolidations des données

Pré-Analyse des données de la table Web:

Web
sku (PK)
.
.
.

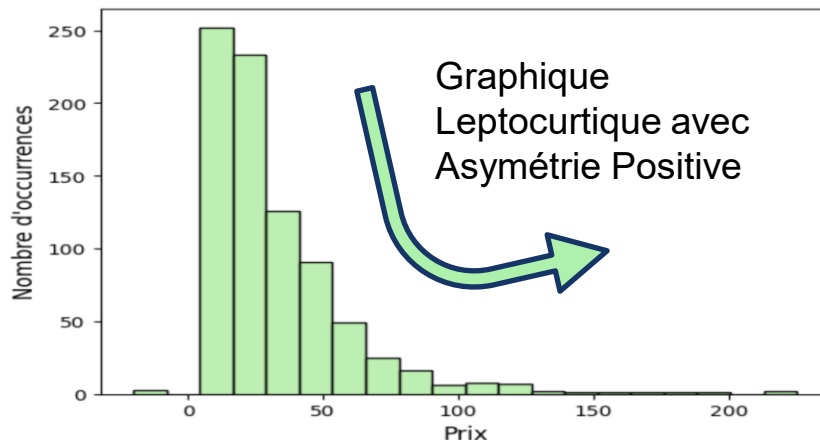
(1513,29)

Colonne	Commentaire
sku	<ul style="list-style-type: none">• Nom de colonne n'est pas aligné avec la table liaison (id_web) -> Corrigé• 1513 lignes dont :• 83 lignes nulles --> Supprimées• 2 valeurs inconnues -> Supprimées• 4 valeurs non codifiées (« 13127-1 » et « bon-cadeau-25-euros » X2)• Lignes en Doublons avec les « valeurs product vs attachement » → Supprimées
tax_class / post_content / post_password / post_content_filtered	<ul style="list-style-type: none">• Colonnes systématiquement nulles (NaN)
Virtual / downloadable / rating_count / average_rating / post_parent / menu_order / comment_count	<ul style="list-style-type: none">• Colonne avec des valeurs à « zero » ou NaN
post_type	<ul style="list-style-type: none">• Product ou attachement (si image), double les lignes

Analyse univariée du prix de vente

Analyse des données « Prix »

Variance	713,53
Ecart-Type	26,71
Moyenne	32,18
Médian	24,3
Mode	45
Skewness	2,69
kurtosis	10,71



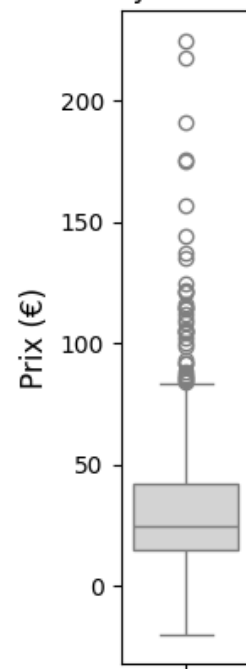
Analyse :

Les prix sont asymétriques, avec asymétrie positive.

Un prix fréquent de 45 €, une moyenne plus élevée que la médiane indiquant de nombreux prix élevés.

Un kurtosis à 10, indique beaucoup de données concentrées autour de la moyenne, mais quelques valeurs extrêmes.

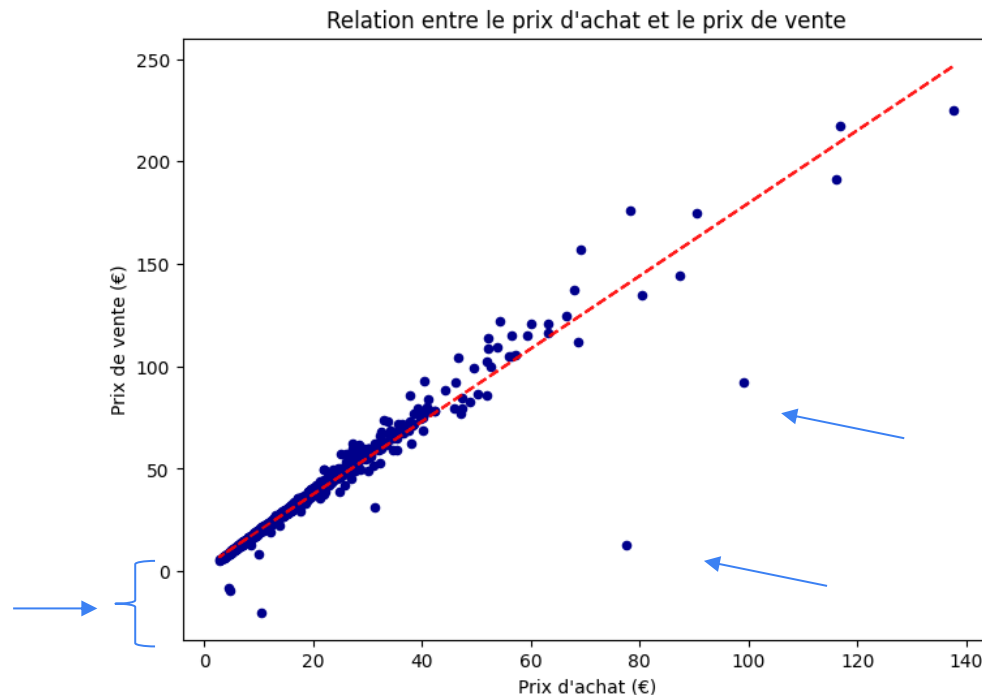
Analyse des prix



Analyse bivarié prix d'achat et prix de vente

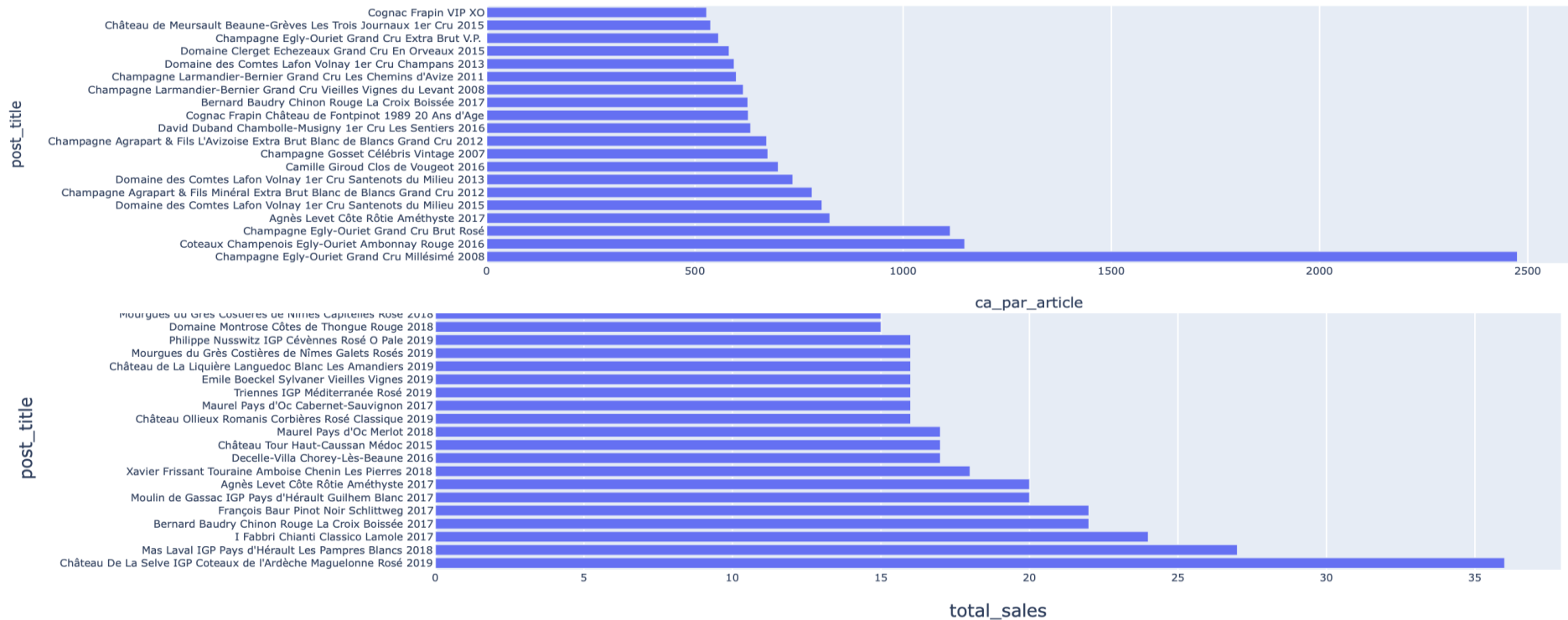
Analyse :

- Cohérence globale entre les prix d'achat et prix de vente.
- Quelques cas atypiques : prix de achat bien inférieur au prix de vente



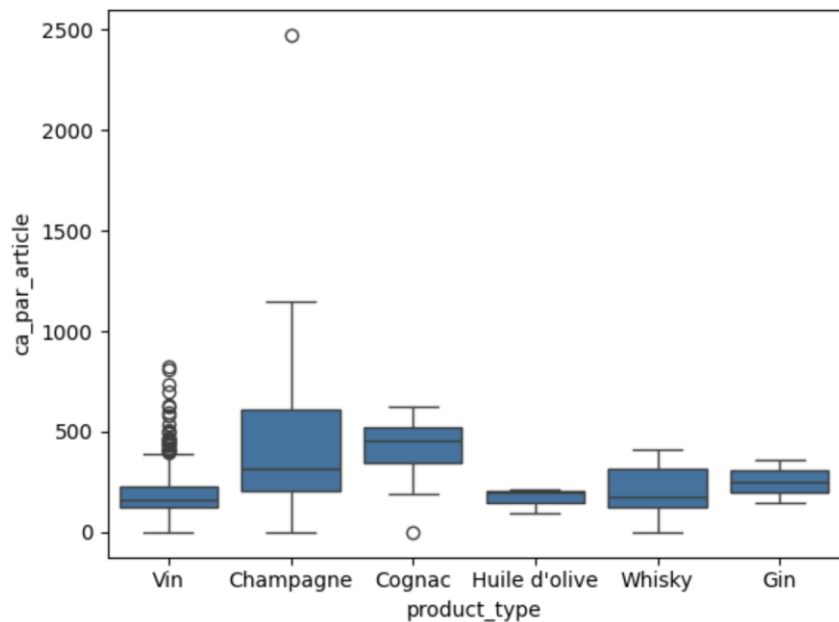
Analyses complémentaires

CA, quantités, stocks, taux de marge et correlations



Analyses complémentaires

CA, quantités, stocks, taux de marge et correlations



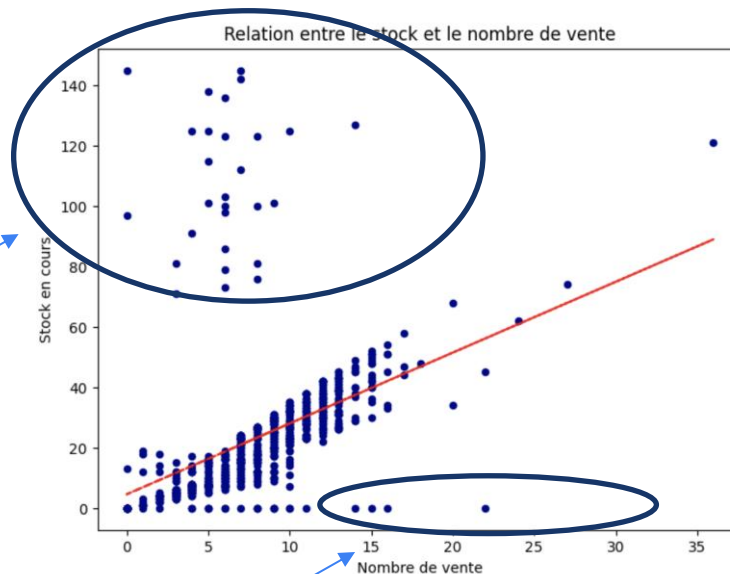
Analyse :

- Chiffre d'affaires Total de 143 598,9 €
- Meilleur chiffre d'affaires médian sur le Cognac

Analyses complémentaires CA, quantités, stocks, taux de marge et correlations

- 16707 articles en stock
- 711 Produits
- Valorisation du stock à 277 225,11€

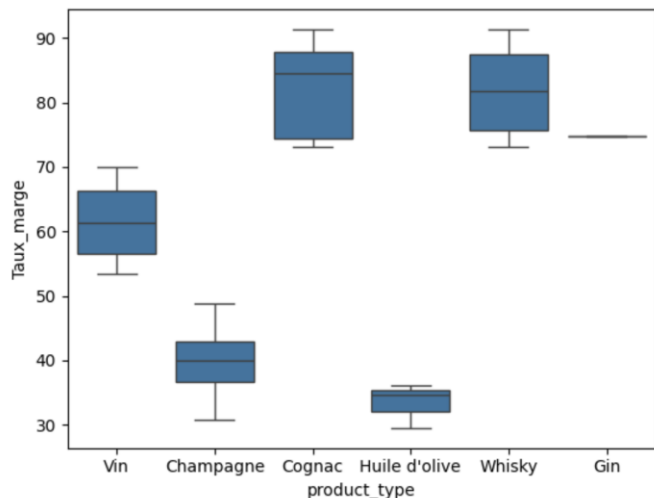
Disposition de stock sur
des produit qui se
vendent peu



Pas de stock sur des
produits qui se vendent
bien

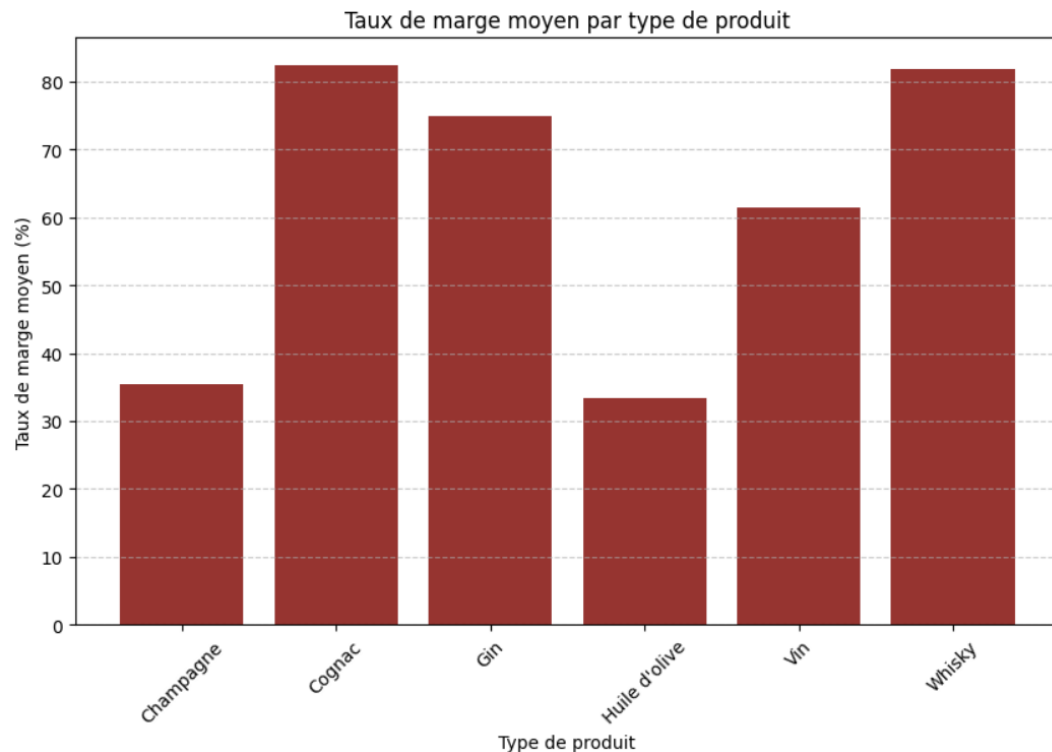
Analyses complémentaires

CA, quantités, stocks, taux de marge et correlations



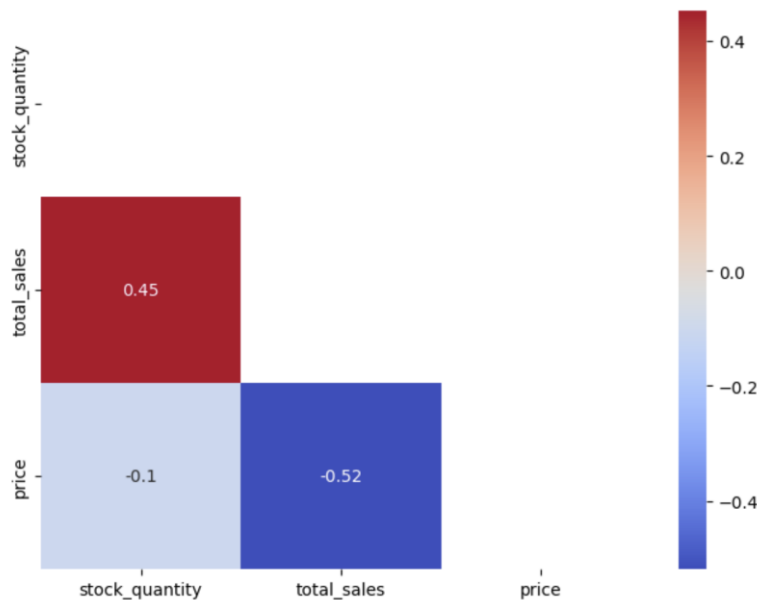
Analyse :

- Taux de marge important sur le Cognac et Whisky
- 2 Gins vendus au même prix



Analyses complémentaires

CA, quantités, stocks, taux de marge et correlations



Analyse des corrélations :

- 1 – Mesure négative entre prix et total des ventes
→ Lorsque le prix augmente, plus les ventes baissent
- 2 – Mesure positive entre total des ventes et stock
→ Les ventes augmentent lorsqu'il y a du stock
- 3 – légère mesure négative entre le prix et le stock
→ lorsque le stock augmente, le prix à légèrement tendance à diminuer

Actions pour la suite

- *Fiabiliser la base de données :*
 - *Appliquer des contraintes « Price / Stock / Product_id »*
 - *Supprimer les colonnes inutiles*
 - *Point d'attention sur RGPD*
- *Intégrer les valeurs :*
 - *Prix_HT, Marge, CA, rotation de stock*
- *Mettre en place un dashboard dynamique:*
 - *Identification immédiate des écarts*

Point sur les compétences apprises

- *Mise en place des graphiques selon typologie des données : Qualitative / Quantitative*
- *Meilleure appréhension des notions de « Statistiques »*
- *Manipulation globale des DataFrames avec Pandas*