

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

Jérémy GUINAULT  
Bottleneck

# Analyses Exploratoires des Données

---

- *Trois datasets :*
  - *Fichier ERP (référence produits, stocks, prix),*
  - *Fichier Web (SKU, quantités vendues, description),*
  - *Table de liaison.*
- *Traitement réalisés :*
  - Analyse des datasets (dimension dataset, nature des données),*
  - Correction du status des stocks,*
  - Recherche d'erreurs (prix ou stocks négatifs),*
  - Suppression de certaines colonnes du fichier Web,*
  - Suppression des doublons, ou articles sans référence.*

# Fusion ou consolidations des données

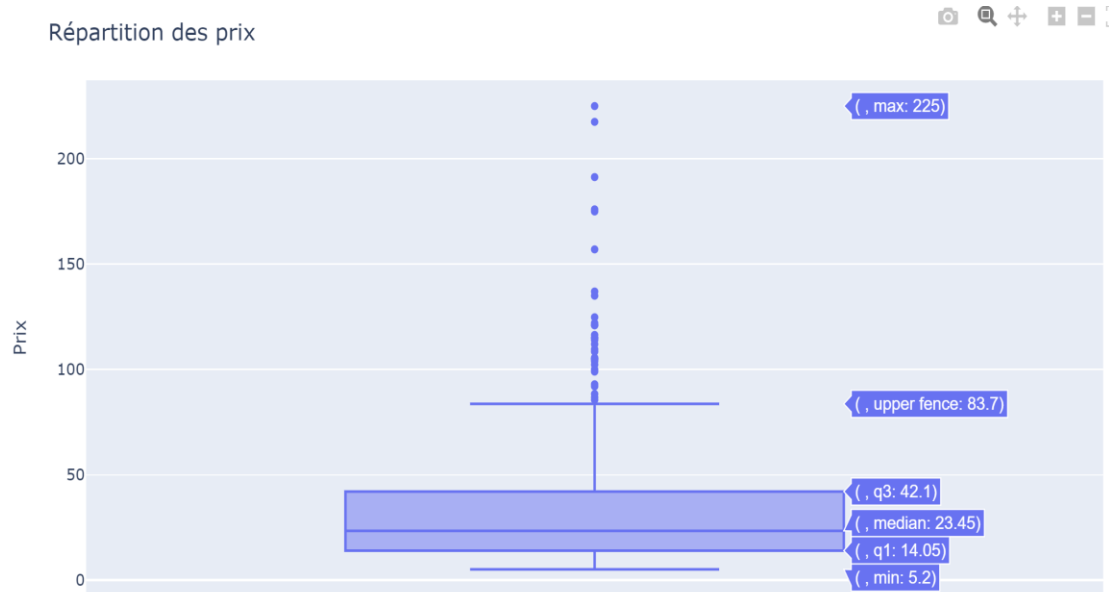
---

- *Utilisation de la table de liaison pour fusionner les fichiers.  
Table de liaison composée de **product\_id** et **id\_web**.*
- *Fusion du dataset ERP et de la table de liaison via product\_id.*
- *Fusion de ce nouveau dataset avec le dataset Web où  
id\_web = SKU.*
- *Obtention d'un fichier (df\_tot) unique contenant les informations  
des datasets ERP et Web.*

# Analyses univariées du prix

- Création d'un graphique « boîte à moustache » de la répartition des prix.

Répartition des prix



- Prix étalés de 5,2€ à 225€,

- Prix médian de 23,45€,

- La moitié des articles ont un prix compris entre 14,05€ et 42,10€.

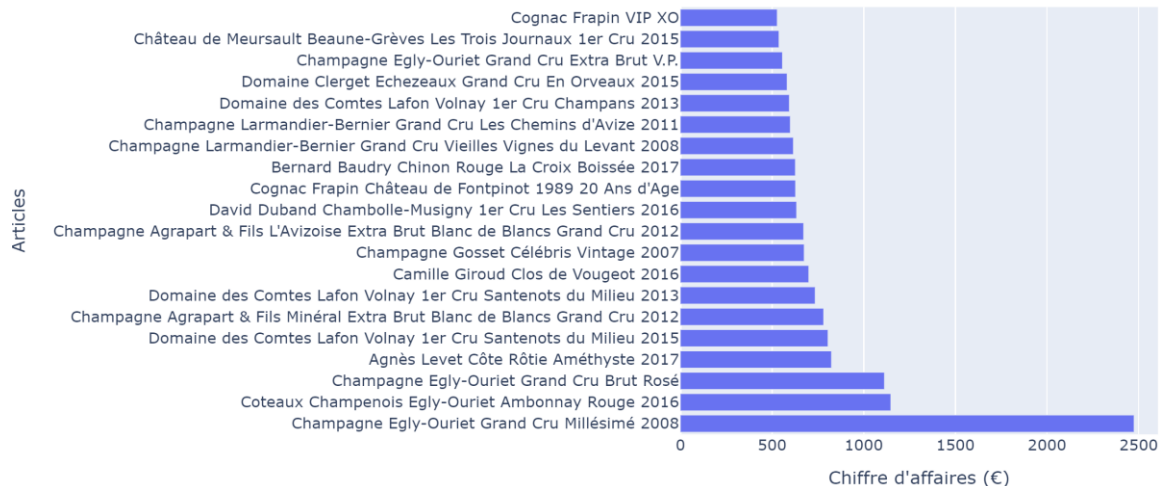
# Analyses univariées du prix

---

- *Utilisation de « describe » pour les mesures de dispersions.*
- *Le prix moyen est de 32,33€.*
- *L'écart-type du prix est de 27,60€.*
- *Le seuil maximum en prix est de 84,09€, 31 produits du catalogue ont un prix supérieur, soit 4,34% du catalogue.*
- *Ces prix pourraient être des valeurs aberrantes, mais ici ils sont justifiés car ils concernent des produits haut de gamme.*

- Le chiffre d'affaires du site internet est de **143680,10€** sur le mois.
- Graphique des 20 produits générant le plus de chiffre d'affaires :

Chiffre d'affaires des 20 premiers articles



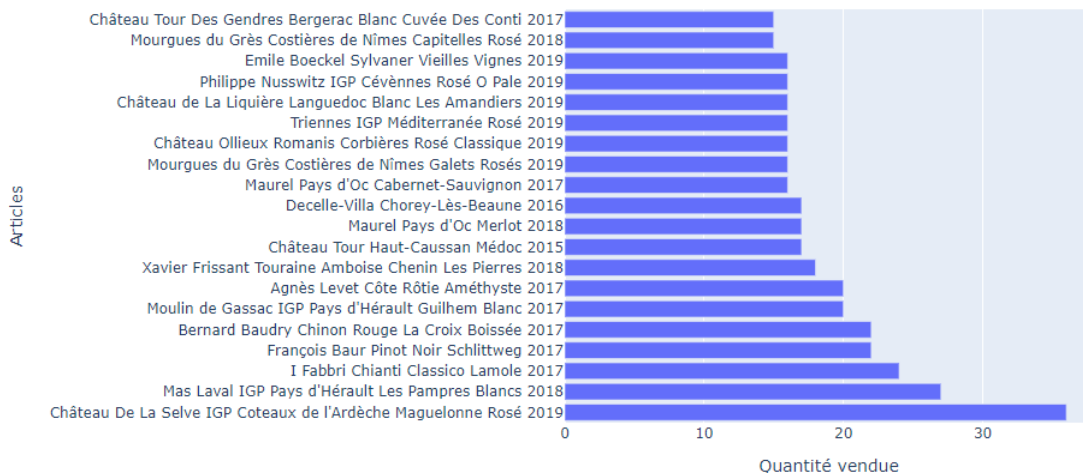
- *Ces 20 produits représentent 2,8% du catalogue et génèrent 11% du chiffre d'affaires.*
- *433 produits, soit 60,6% du catalogue, génèrent 80% de notre chiffre d'affaires.*
- *Sur ce mois certains produits ne se sont pas vendus et n'ont donc généré aucun chiffre d'affaires.*

## Analyses complémentaires

### Quantités vendues

- Un total de **5751 produits** vendus sur le mois.
- *Graphique des 20 produits les plus vendus pour octobre :*

Les 20 articles les plus vendus en quantité

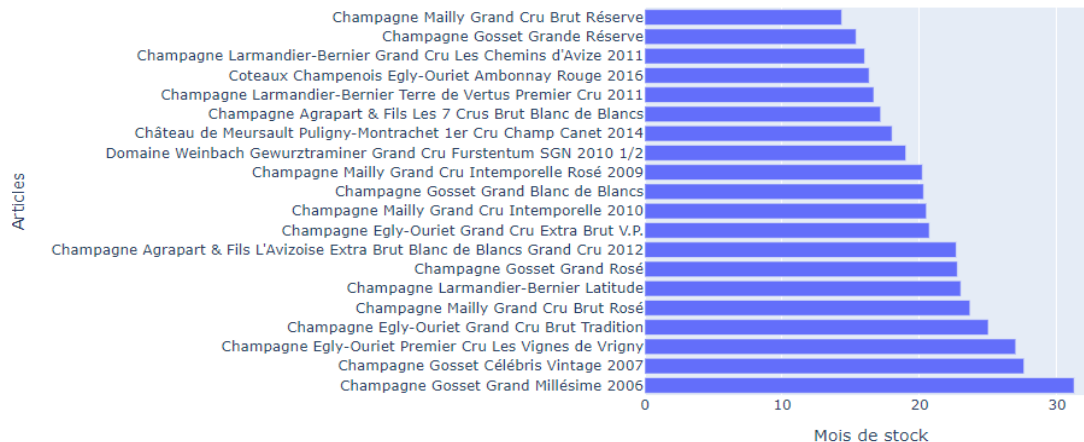




- *Ces 20 produits représentent 382 ventes sur le mois.*
- *432 produits, soit 60,6% du catalogue, génèrent 80% des produits vendus.*
- *Sur ce mois-ci 25 produits n'ont pas été vendu une seule fois, soit 3,5% du catalogue.*

- Un stock total de **16739 produits**, soit une valorisation de **494593.4€** de celui-ci.
- Ci-dessous un graphique présentant les 20 articles ayant le plus de mois de stock, par rapport aux ventes de ce mois-ci :

Les 20 articles avec le plus de mois de stock

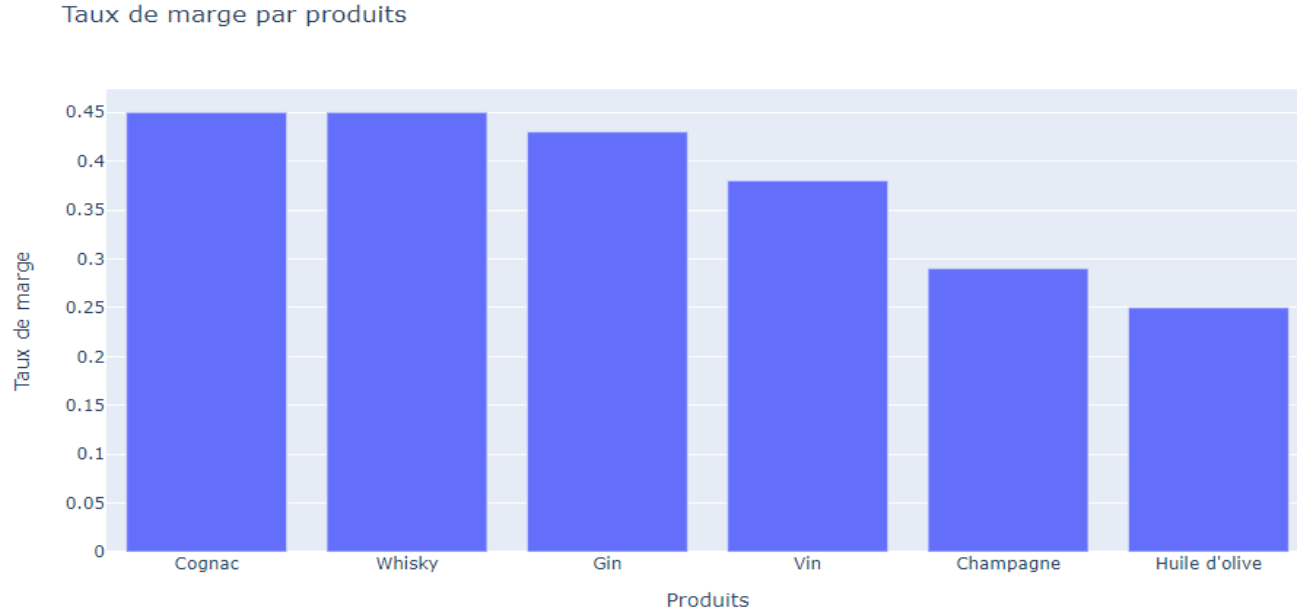


- *Taux de marge variable :*
  - *Le taux maximum est de 48%,*
  - *Le taux minimum est de -635%,*
  - *Le taux moyen est de 36%.*
- Il est anormal d'avoir un taux de marge négatif, car cela signifie qu'on le vend à perte.
- Le produit concerné est le 'Champagne Egly-Ouriet Grand Cru Blanc de Noirs' acheté 77,48€ et revendu 10,54€ HT.

## Analyses complémentaires

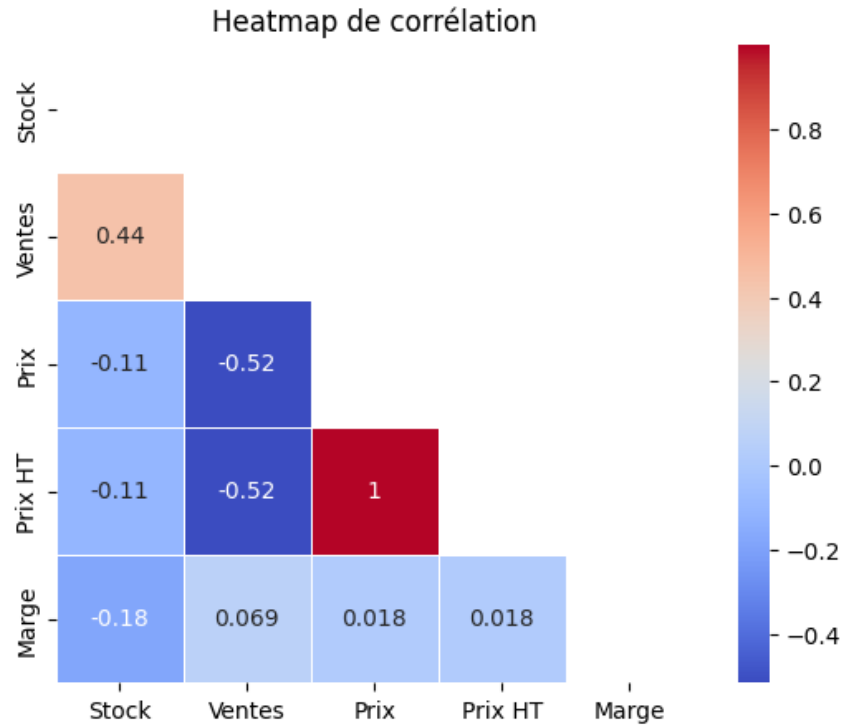
### Taux de marge

- *Graphique présentant le taux de marge par produits :*



## Analyses complémentaires

### Heatmap de corrélation



- On constate une corrélation modérée entre le nombre de ventes et le prix des produits.  
Plus le produit a un prix « attractif » plus il sera vendu.
- Une corrélation modérée également entre le nombre de ventes et le stock.  
En effet les produits en rupture de stock ne génèrent pas de ventes et inversement.

# Actions pour la suite

---

*Afin de développer au mieux l'ERP :*

- *Limiter au maximum les erreurs (prix, stock...) afin d'avoir des données cohérentes,*
- *Conserver uniquement des informations pertinentes,*
- *Lié si possible l'ERP au site web, pour avoir les données à jour continuellement.*