

[Optimisez la gestion du stock d'une boutique en nettoyant ses données]

[COEFE Robet Christian]
[Business Intelligence Analyst]

[26/11/2023]

Analyses Exploratoires des Données

- **Le dataset est composé de 4 fichiers**
- **Importation des Librairies**
- ❖ **Pandas pour la manipulation et l'analyse des données**
- ❖ **Seaborn permet la Création de statistiques pour exploration et la compréhension des données**
- ❖ **NumPy permet de manipuler des tableaux multidimensionnels pour réaliser des calculs mathématiques**
- ❖ **Chardet permet de déterminer l'encodage du fichier**
- ❖ **Plotly permet de réaliser des graphiques pour la visualisation des données**

- **Caractéristiques des fichiers**

- ❖ **Le fichier Erp sous format Excel**
- ❖ **Le fichier Web sous format Excel**
- ❖ **Le fichier Liaison sous format Excel**
- ❖ **Le fichier Caractéristique sous format CSV**

- **Traitement réalisés**

- ❖ **Recherche des colonnes vides ou comportant des doublons**
- ❖ **Recherche des lignes vides ou comportant des doublons**
- ❖ **Recherche des colonnes incohérents ou des données manques**
- ❖ **Suppressions des lignes, des colonnes vides et des doublons**
- ❖ **Réalisation de calculs et de graphiques**

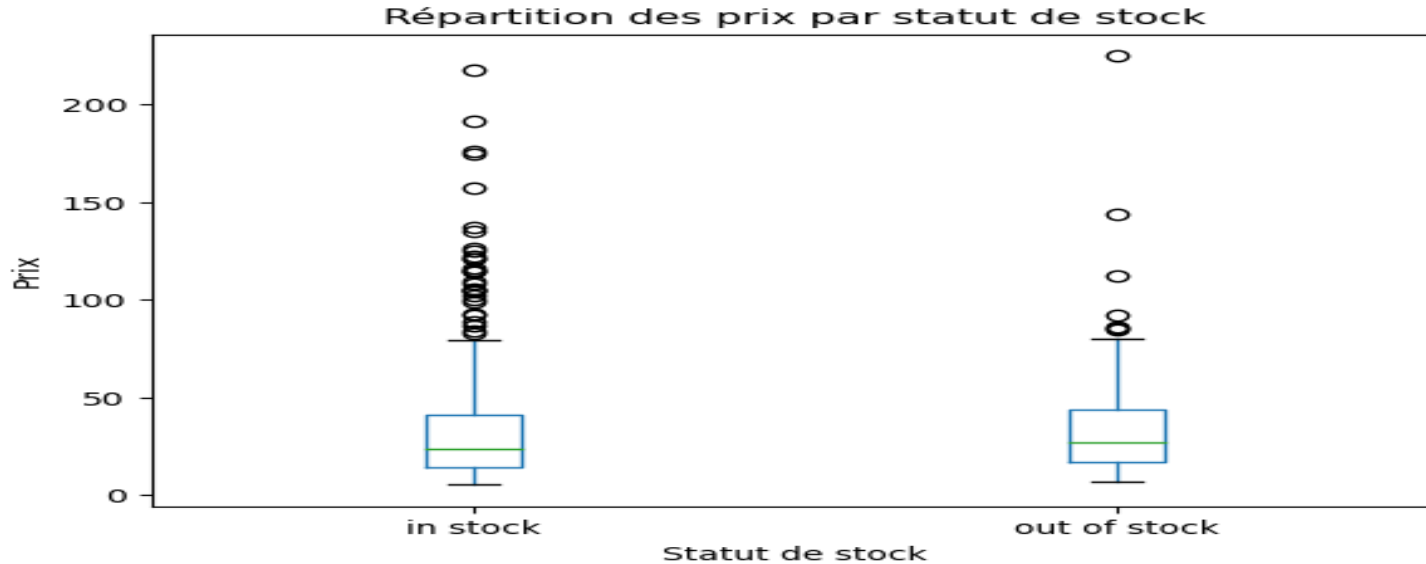
Fusion ou consolidations des données

- **Choix des attributs**
- **Clés utilisés**
 - ❖ **Product_id** permet de lier le fichier Erp et Liaison
 - ❖ **Id_web et SKU** permet de lier le fichier Df_final et Web
 - ❖ **Post_name** permet de lier le fichier Df_final et Df_cara
- **Vigilances particulières au cours du traitements**
 - ❖ **Le choix de la jointure**
 - ❖ **La suppression des lignes en doubles**
 - ❖ **Les lignes sans correspondance**

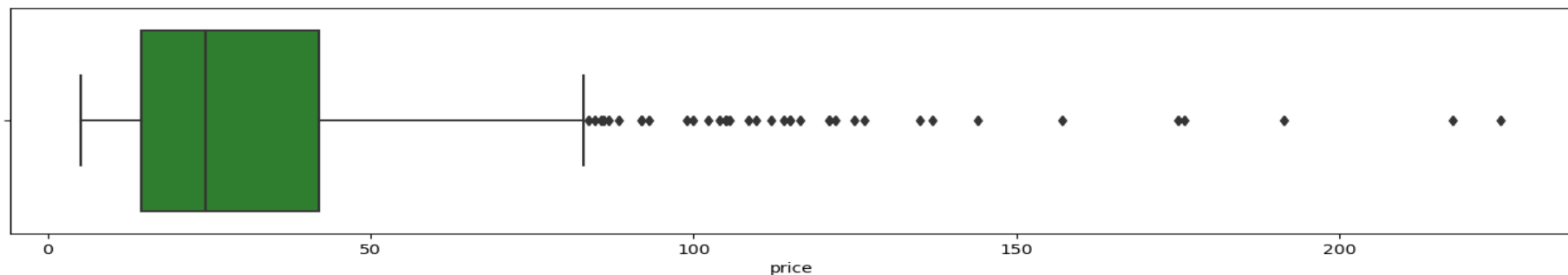
- Difficultés ou pièges rencontrés
- ❖ L'encodage du fichier Caractéristique
- ❖ Renommer la colonne SKU
- ❖ Choix des colonnes à conserver

	product_id	sku	onsale_web	price	stock_quantity	total_sales	post_name
0	3847	15298	1	24.2	0	6.0	pierre-jean-villa-saint-joseph-preface-2018
1	3849	15296	1	34.3	0	0.0	pierre-jean-villa-saint-joseph-tilde-2017
2	3850	15300	1	20.8	0	0.0	pierre-jean-villa-croze-hermitage-accroche-coe...
3	4032	19814	1	14.1	0	3.0	pierre-jean-villa-igp-gamine-2018
4	4039	19815	1	46.0	0	0.0	pierre-jean-villa-cote-rotie-carmina-2017

Analyses univariées du prix



- La majorité des prix des produits sont inférieur à 150€
- La majorité des produits commercialisés sont en stock



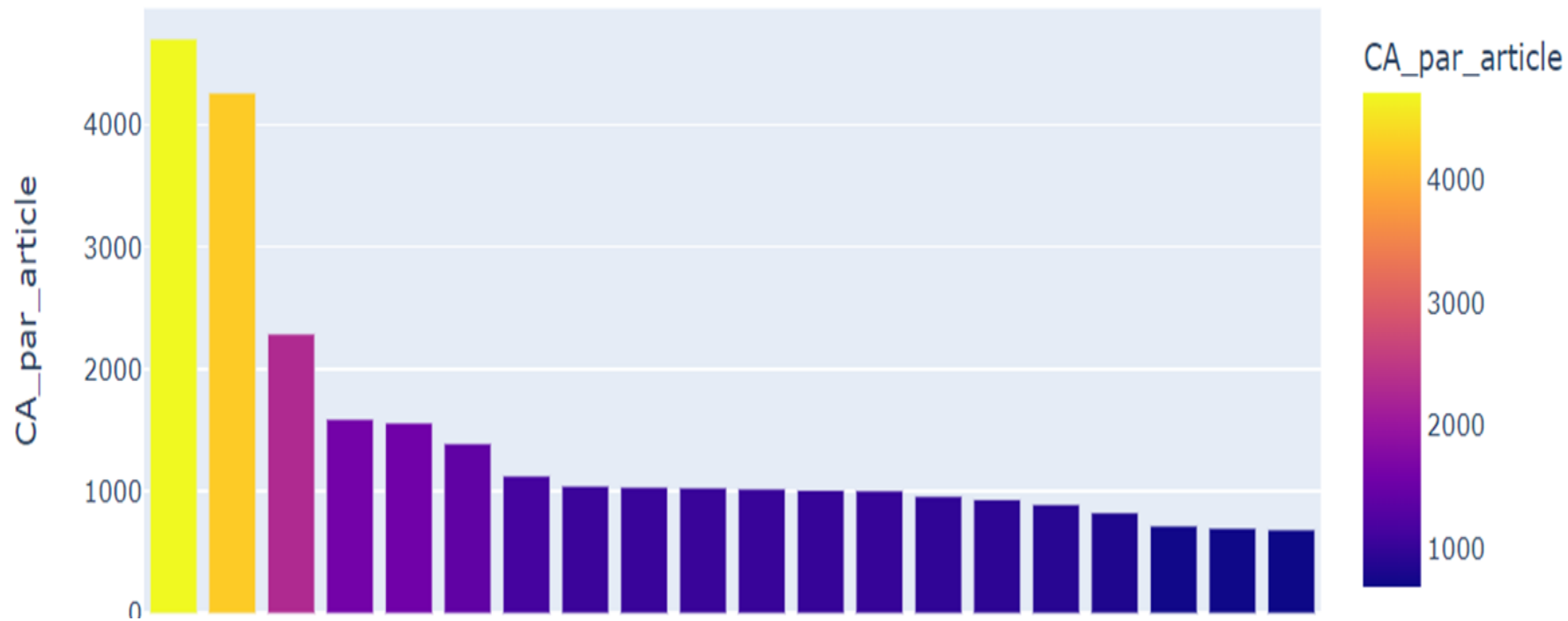
Etude du prix:

count	825.000000
mean	32.415636
std	26.795849
min	5.200000
25%	14.600000
50%	24.400000
75%	42.000000
max	225.000000

Etude du Z-score:

count	8.250000e+02
mean	-8.612639e-17
std	1.000000e+00
min	-1.015666e+00
25%	-6.648655e-01
50%	-2.991372e-01
75%	3.576809e-01
max	7.187097e+00

Analyses univariées du CA

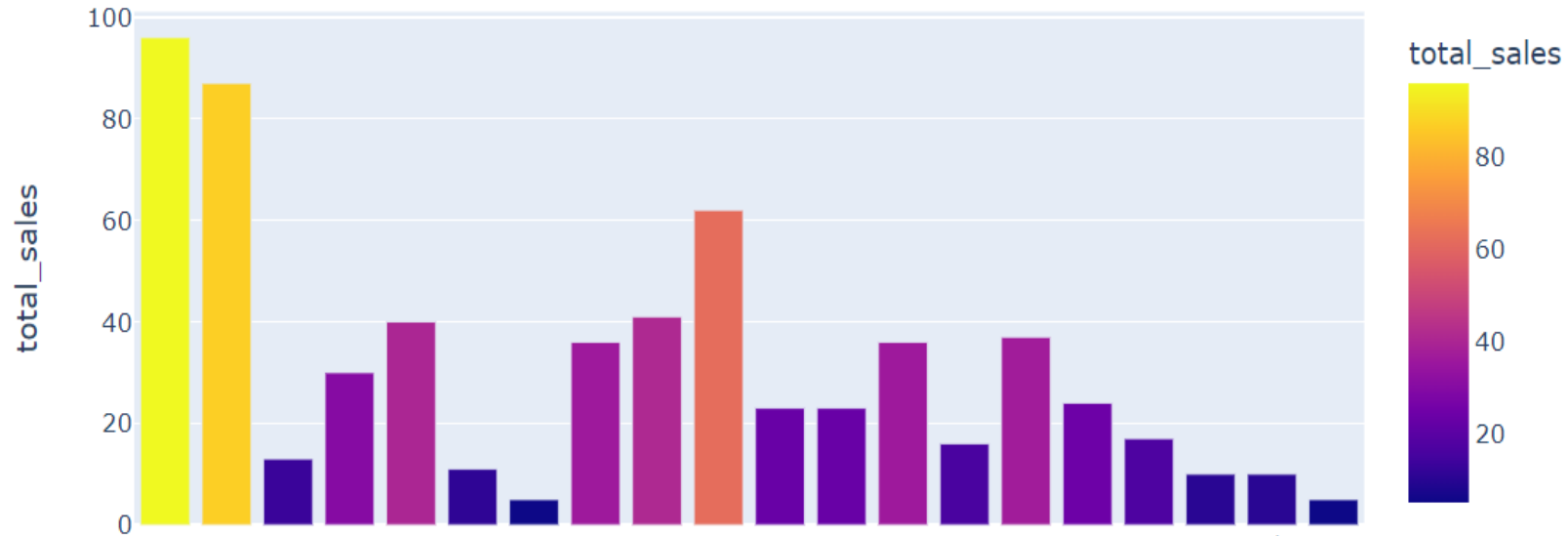


- ❖ **Le chiffre d'affaires est de 70 318,6€.**
- ❖ **La proportion de la participation du top 20 des ventes du site web au chiffre d'affaires est de 56% soit 462 articles sur 825 articles.**
- ❖ **Le top 20 se compose 6 champagnes, 1 cognac, et 13 vins.**

Etude de la quantite totale vendue:

count	712.000000
mean	4.009831
std	8.525871
min	0.000000
25%	0.000000
50%	1.000000
75%	4.000000
max	96.000000

- ❖ **328 articles vendus représentent 40% du total des articles vendus**
- ❖ **Le top 20 des articles vendus se compose , en quantité du vin, du champagne et du cognac.**



Actions pour la suite

- ❖ **Mettre en place une veille au niveau des stocks concernant les meilleurs afin de pas avoir de rupture de stock pour ces produits.**
- ❖ **Eliminer de nos stocks les produits non-vendus et garder uniquement les produits qui se vendent.**
- ❖ **Comme il s'agit d'un site internet on pourrait faire les publicités sur différents réseaux sociaux.**

- ❖ **Etudier les sites des concurrents pour améliorer le notre**
- ❖ **Mettre en place un tunnel de ventes pour augmenter la fréquentation et améliorer les ventes**
- ❖ **Payer des référencementements sur google pour augmenter la visibilité sur internet**

Point sur les compétences apprises

- **Qu'est-ce qui s'est bien passé pour vous dans ce travail de nettoyage ?**
- ❖ **Le rapprochement des fichiers**
- ❖ **Detecter les caracteristiques des produits**
- ❖ **Supressions des lignes en double**

- **Qu'est-ce que vous avez trouvé le plus difficile ?**
- ❖ **Création du dataset final**
- ❖ **Création des graphes**
- **Sur quelles tâches est-ce que vous pensez avoir besoin de plus d'entraînement ?**
- ❖ **L'utilisation des différentes bibliothèques python**
- ❖ **Les formules de statistique**
- ❖ **La maitrise du logiciel python**