

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

[Nom de l'étudiant:Laghzouli Amina]

[Date de la présentation:10/08/2023]

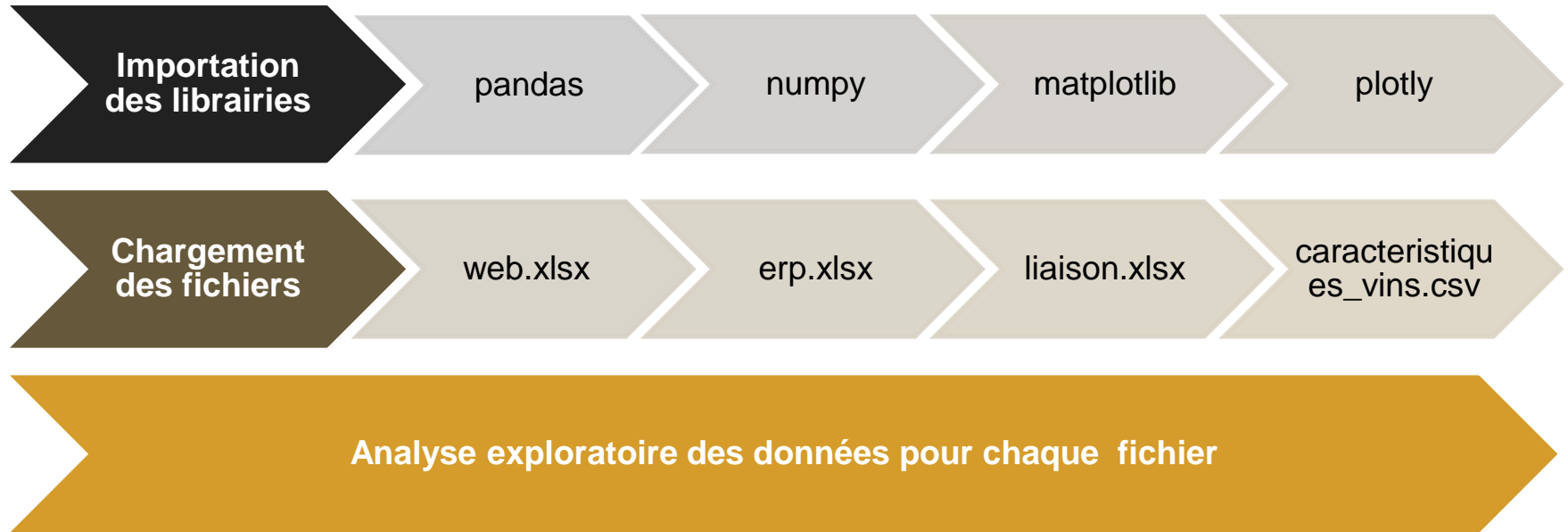
But de projet:

- Optimiser la gestion du stock de BottleNeck.
- Faire un rapprochement entre l'entreprise ressource planning « ERP » et le site de vente en ligne.
- Augmenter la visibilité d'analyse des ventes sur Internet.

Les objectifs de la présentation:

- Le rapprochement entre les 2 exports : un export de l'ERP contenant les références produit, leur prix de vente et leur état de stock, et un export d'une table de la boutique en ligne contenant les informations des produits commercialisés en ligne.
- Le calculer de chiffre d'affaires par produit, ainsi que le total du chiffre d'affaires réalisé en ligne.
- Les analyses effectuées sur la variable prix afin de détecter d'éventuelles valeurs aberrantes, de les lister et d'en faire une représentation graphique pour plus de lisibilité.

Analyses Exploratoires des Données



Analyses Exploratoires des Données

Df_erp

- Le tableau comporte 825 observations et 5 colonnes
- Corriger les données incohérentes
- Analyse les variables 'PRICE', 'STOCK', 'ONSALE_WEB'

Df_web

- Le tableau comporte 1513 observations et 28 colonnes
- Corriger les données manquantes, et supprimer les autres valeurs non renseignés
- Analyse la colonne 'SKU'

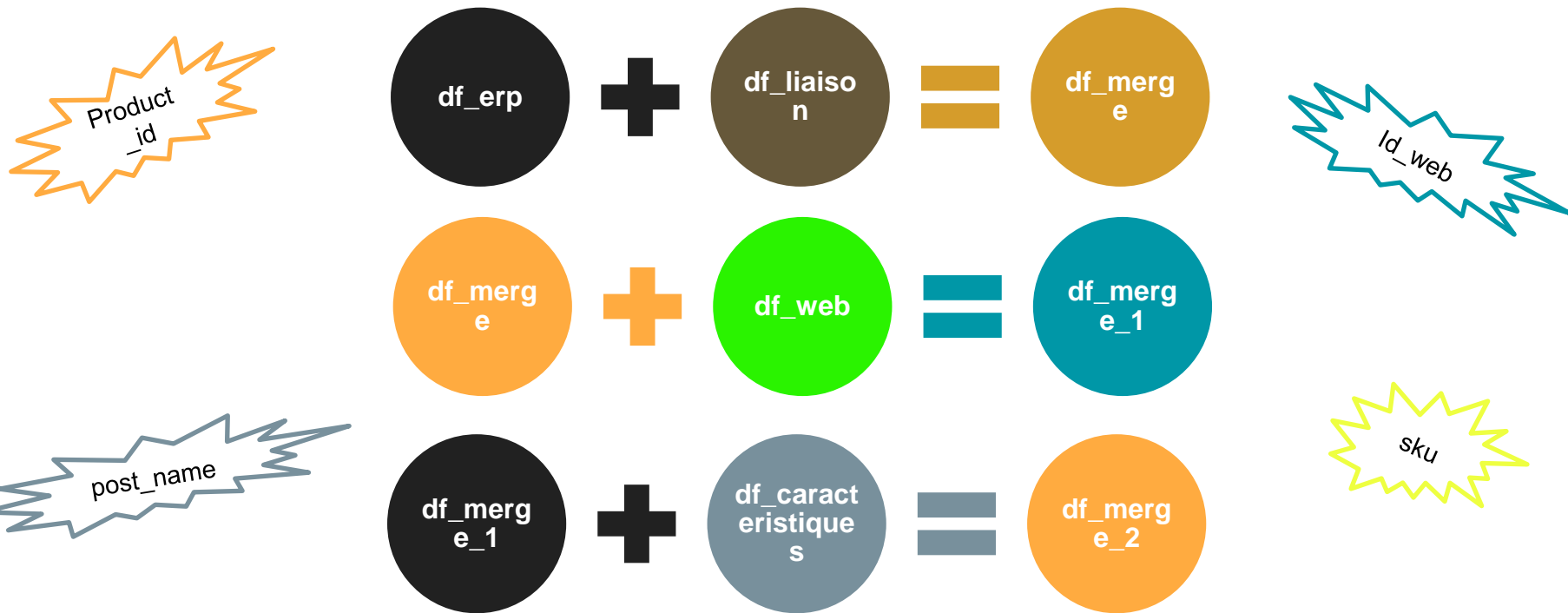
Df_liaison

- Le tableau comporte 825 observations et 2 colonnes
- Vérifier les valeurs de chaque colonne
- Trouver 90 articles sans correspondance

Df_caracteristiques_vins

- Le tableau comporte 611 observations et 13 colonnes
- Trouver les produits avec des informations manquantes, et modifier à partir des autres colonnes

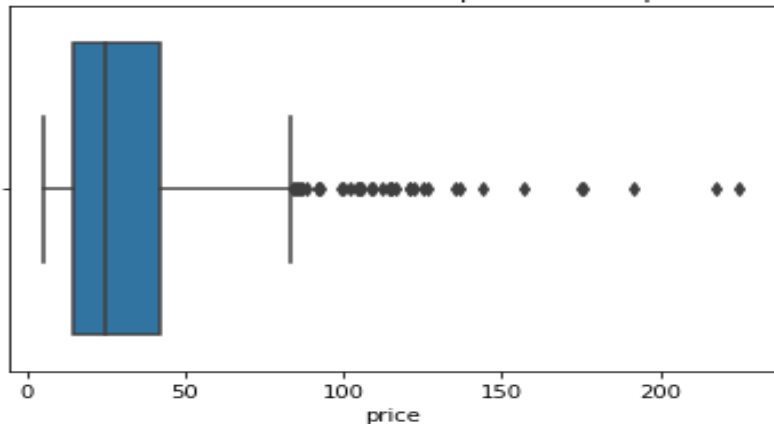
Fusion ou consolidations des données



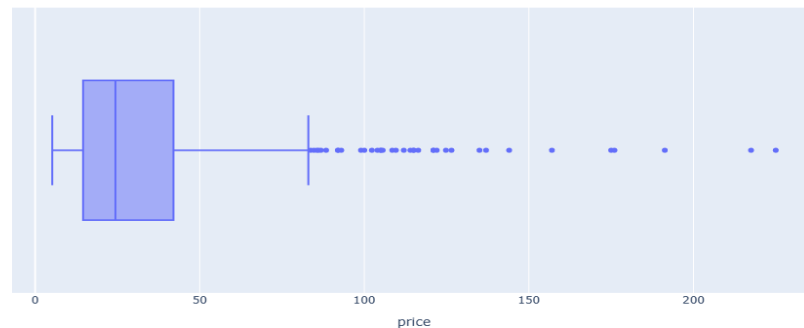
Analyses univariées du prix

1-Exploration par la visualisation de données :

Boite à moustache de la répartition des prix



la répartition des prix grâce à Pandas



Sur cette visualisation des données, les outliers sont représentés sous forme de points. Ils correspondent à des observations dont les valeurs sont supérieures à la valeur du 3ème quartile plus 1.5 fois l'intervalle interquartile.

Analyses univariées du prix

2-Exploration par l'utilisation de méthodes statistique :

Identification par le Z-index

Calculer la moyenne et l'écart-type du prix

Calculer z_score

Identifier le seuil prix dont z-score est supérieur à 3

Identification par l'intervalle interquartile

Etudier les mesures de dispersions

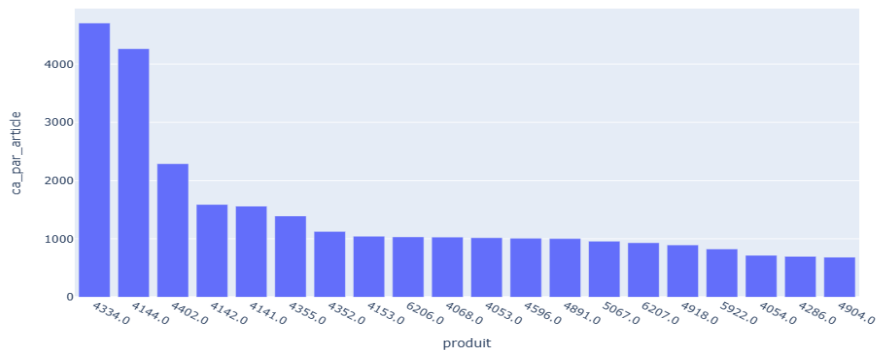
Définir un seuil pour les articles "outliers" en prix

Définir la proportion de l'ensemble du catalogue "outliers"

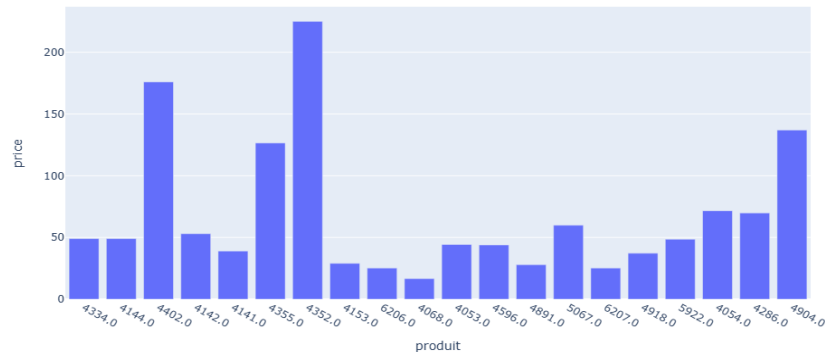
Analyses univariées du CA

1-Analyse des ventes en CA:

Graphique en barre des 20 premiers articles



Graphique en barre des 20 premiers articles



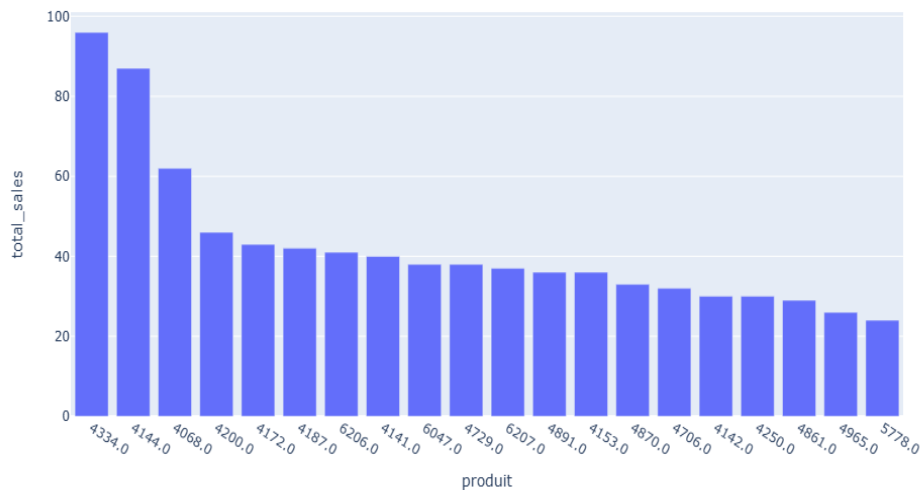
Le prix élevé par produit ne signifie pas dire que le chiffre d'affaires élevé et le contraire est vrai.

On a 130 articles représentant 80% du CA

Analyses univariées du CA

2-Analyse des ventes en Quantités:

Graphique en barre des 20 premiers articles



La quantité de ventes élevé par produit signifie que le chiffre d'affaires élevé.

On a 150 articles représentant 80% des ventes en quantité

Actions pour la suite

Créer une base de données relationnelle avec les différentes tables.

Fiabiliser les données.

Automatiser l'analyse de données.

Conclusion

- Faire le nettoyage de données pour optimiser la gestion du stock de BottleNeck.
- Relier l'entreprise ressource planning ERP avec le site de vente en ligne.
- Calculer le chiffre d'affaires par produit , le chiffre d'affaires total.
- Détecter les valeurs aberrantes de la variable prix