

The background of the slide is a photograph of a large, modern wine cellar. The room is filled with rows of wooden wine racks, each containing numerous bottles of wine. The racks are arranged in long aisles that recede into the distance. The lighting is soft and focused on the wine bottles, creating a professional and sophisticated atmosphere. The ceiling is dark with exposed wooden beams, and the floor is a light-colored, polished surface.

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

[Loïc Stéphane BAMENOU]
[Business Intelligence Analyst]
[23/04/2024]

Plan de présentation

- I. Contexte du projet ;*
- II. Analyses exploratoires des données ;*
- III. Fusion et consolidations des données ;*
- IV. Analyse univariée du prix*
- V. Analyse univarié du Chiffre d'affaires (CA)*
- VI. Actions pour la suite*
- VII. Point des compétences acquises*

Contexte du projet

*Dans un souci d'avoir une vision globale du chiffre d'affaires lié à ces produits de prestige, l'entreprise « **Bottleneck** » fait appelle à notre aide. Il faut le rappel que pour elle utilise des outils un peu artisanaux et obsolètes en attendant l'adoption d'un outil centralisé pour la gestion des stocks et la vente .*

***Enjeu du projet :** la mise en place d'un outil temporaire permet analyse les ventes en utilisant les données du back-office à mettre à niveau*

Objectifs de la présentation :

- *Analyser et consolider les données du stock et de la vente ;*
- *Analyser le chiffre d'affaires réalisés en ligne de l'entreprise ;*
- *Produire un fichier synthèse*
- *Etablir un plan d'action pour continuer d'avoir une vision globale*



Analyses Exploratoires des Données

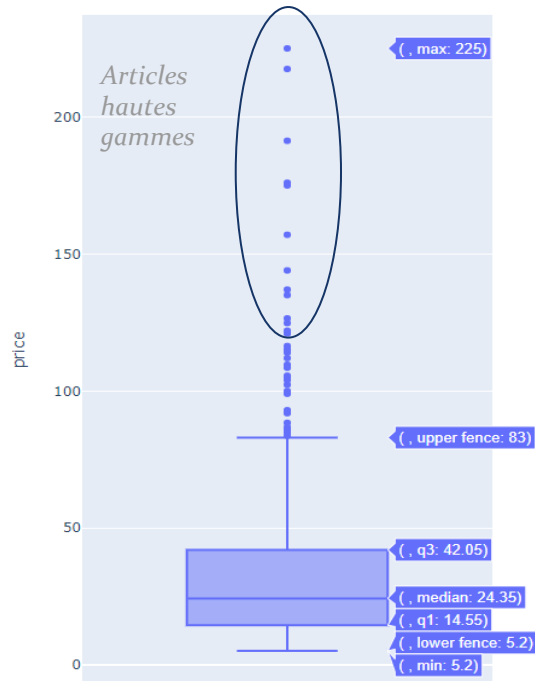
- *Outils d'analyse : Python via google Colab ;*
- *Fichiers (datasets) à synthétisés*
 - *"ERP"*
 - *"LIAISON"*
 - *"WEB"*
 - *"CARACTERISTIQUES "*
- *Premières reflexes à avoir*
 - *Attention à l'encodage des fichiers à exploiter ;*
 - *Cartographie des données ;*
 - *Contrôle de la qualité des données (valeurs manquantes, doublons, outliers et autres)*
- *Traitements réalisés*
 - *Nettoyages des données*
 - *Sélection des variables souhaitées ;*
 - *Repérage des données incohérentes ;*
 - *Features engineering*
 - *Correction du statut des stock ;*
 - *Contrôle du contenu des variables ;*
- *Remarques éventuelles, pièges ou difficultés rencontrées*
 - *Ligne incohérente dans le fichier "ERP" ;*
 - *Absence de données à jour sur le catalogue de produits ("CARACTERISTIQUES") ;*
 - *Non respect de la typologie de certaines variable « sku » :*

Fusion ou consolidations des données

- *Choix des attributs*
- *Clés utilisés pour donner lieu à notre fichier "df_merge"*
 - *Entre les fichiers "ERP" et "LIAISON" la clé de jointure est « product_id » ;*
 - *Entre les fichiers "LIAISON" et "WEB" les clés de jointure sont « id_web » et « sku » ;*
 - *Entre les fichiers "WEB" et "CARACTERISTIQUES" la clé de jointure est « post_name » ;*
- *Vigilances particulières au cours du traitements*
 - *Vérification du nombre de observation dans le fichier issu de la consolidation ;*
 - *Contrôle de la nature de la jointure souhaitée ;*
- *Difficultés ou pièges rencontrés*
 - *Création de doublons lors de la fusion de datasets à dimensions inégales*
 - *Formulation des attentes*

Analyses univariées du prix

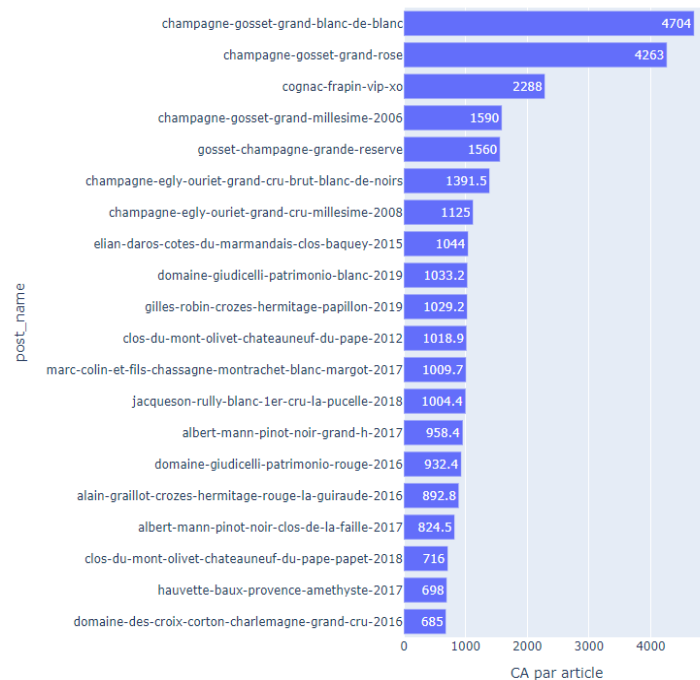
- *Boîte à moustache de la répartition des prix*



- *Le coût moyen d'une bouteille est de 32,43 €;*
- *Nos bouteilles de vins ont un prix variant entre 5,2 € et 225 € avec la moitié d'entre elles à moins de (prix médian) 24,35 € ;*
- *On note une certaine homogénéité des coûts avec un écart considérable de 28 € sur la moitié de nos bouteilles (soit l'intervalle interquartile) ;*
- *On note également la présence de valeurs atypiques dans le quantile supérieur représentant nos bouteilles les plus prestigieuses à partir de 114 € (valeur obtenue à partir de la Z-score)*

Analyses univariées du CA

- *Classements du top 20 des produits mieux vendus*



- *Sur les 20 000 produits proposés à notre clientèle, nous avons réussi à écouler 13% de nos articles (soit 713 articles distincts articles en ligne, 86,89 %);*
- *Le chiffre d'affaires généré s'élève à un montant de 70 318,6 €*
- *Le succès de notre gamme de champagnes « gassets grands » sur l'extrade et des produits de l'année « 2019 »;*

Actions pour la suite

- *Réaliser audits de hebdomadaire ou mensuel pour actualiser les données ;*
 - *Collecter et préparer les données (ventes , stock, fournisseurs) ,*
 - *Nettoyer les données*
- *Former et sensibiliser le personnel à l'usage des outils de la Science de la Données ;*
 - *Analyse exploratoire des données*
 - *Analyses multivariées*
- *Automatiser le processus de calcul des Kpi d'entreprise (interne ou externe) ;*
- *Mettre en place un modèle prédictif sur le niveau des stocks ;*
- *Intégrer un système centralisé de suivi des stocks et des ventes à temps réel (Dashboard) ;*

Point sur les compétences apprises

- *Qu'est-ce qui s'est bien passé pour vous dans ce travail de nettoyage ?*
 - *L'étape d'identification des erreurs contenus dans les fichiers (manquants, erreurs, doublons et outliers)*
 - *La réalisation des statistiques de bases (descriptive, valorisation d'opérations et)*
- *Qu'est-ce que vous avez trouvé le plus difficile ?*
 - *L'apprentissage à départager les termes méthodes et fonctions dans la mise en pratique*
 - *La mise en forme d'un graphique sans les librairies graphiques (Matplotlib, Seaborn ou plotly express) ;*
 - *La découverte des paramètres graphiques et des filtres avancés sur les datasets*
 - *Le choix des variables à sélectionner dans le cadre de l'étude ;*
 - *La mise en place des tableaux empiriques (concept un peu flou sur la pratique sur python)*
 - *La préparation du fichier Excel à exporter sur google colab*
- *Sur quelles tâches est-ce que vous pensez avoir besoin de plus d'entraînement ?*
 - *La prise de décision dans l'étape de l'analyse exploratoire ;*
 - *La prise en main des librairies graphiques et de Numpy ;*
 - *L'analyse des valeurs outliers ;*