

P6 - Optimisez la gestion du stock d'une boutique en nettoyant ses données

Leonardo MONTILLA

Business Intelligence Analyst

22/12/2023

Analyses Exploratoires des Données

ERP.xlsx

825 observations et 5 caractéristiques

1 valeur incohérente :
4954 – bon cadeau
25€

Prix : de 5,2€ à 225€

WEB.xlsx

714 observations (dont
deux doublons) et 28
caractéristiques

Plusieurs colonnes
avec des valeurs vides
(NaN)

Liaison.xlsx

825 observations (dont
90 avec des valeurs
vides)

2 caractéristiques
(product_id, id_web)

Caractéristiques_vin.xlsx

611 observations et 13
caractéristiques

Toutes les valeurs ne
sont pas renseignées
(plus grand niveau
de détail : Millésime,
Appellation...)

Fusion ou consolidations des données

df_erp



df_liaison

Clé : 'product_id'



df_merge

```
_merge  
both          825  
left_only     0  
right_only    0  
Name: count, dtype: int64
```

Fusion ou consolidations des données

df_merge



df_web

Clés
'id_web' & 'sku'



df_final

```
_merge  
both          710  
left_only     115  
right_only    2  
Name: count, dtype: int64
```

Fusion ou consolidations des données

df_final



df_caracteristiques

Clés
'post_name' & 'post_title' → *Normalisation des colonnes + left join*



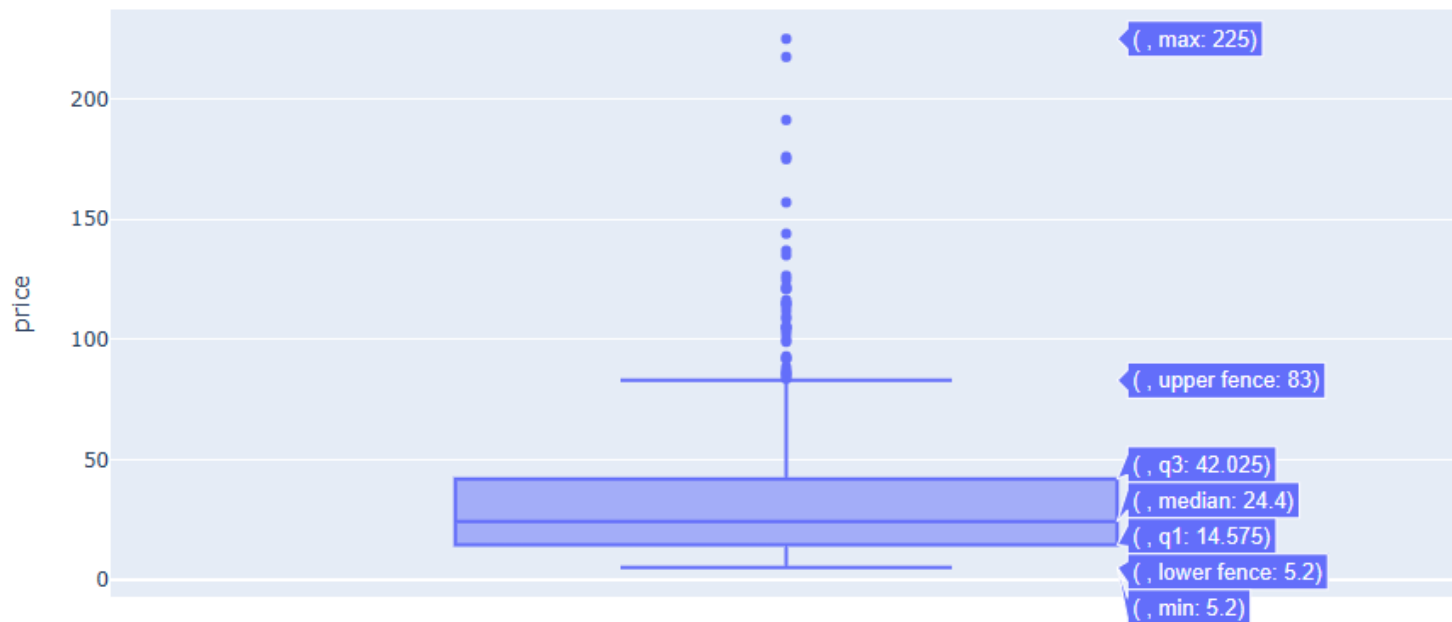
df_final_1

	product_id	onsale_web	price	stock_quantity	stock_status_2	id_web	sku	total_sales	post_title	post_type	_merge
0	3847	0	1.0	24.2	0.0	outofstock	15298	15298	6.0	Pierre Jean Villa Saint-Joseph Préface 2018	product both
1	3849	0	1.0	34.3	0.0	outofstock	15296	15296	0.0	Pierre Jean Villa Saint-Joseph Rouge Titidé 2017	product both
2	3850	0	1.0	20.8	0.0	outofstock	15300	15300	0.0	Pierre Jean Villa Crozes-Hermitage Accroche Co...	product both
3	4032	0	1.0	14.1	0.0	outofstock	19814	19814	3.0	Pierre Jean Villa IGP Collines Rhodaniennes Ga...	product both
4	4099	0	1.0	46.0	0.0	outofstock	19815	19815	0.0	Pierre Jean Villa Côte Rôtie Carmina 2017	product both
...
822	7247	0	1.0	54.8	23.0	instock	13127-1	NaN	NaN	NaN	NaN left_only
823	7329	0	0.0	26.5	14.0	instock	14680-1	NaN	NaN	NaN	NaN left_only
824	7338	0	1.0	16.3	45.0	instock	16230	16230	0.0	Domaine Saint-Nicolas Vin de France Blanc Les ...	product both
825	NaN	NaN	NaN	NaN	NaN	NaN	NaN	19823	0.0	Bon cadeau de 25€	product right_only
826	NaN	NaN	NaN	NaN	NaN	NaN	NaN	19824	0.0	Pierre Jean Villa Condrieu Jardin Suspendu 2018	product right_only

Analyses univariées du prix

Boîte à moustache

avec répartition des prix



Analyses univariées du prix

Z-Score

Et prix moyen

	price	z_score_price
0	24.2	-0.306601
1	34.3	0.070323
2	20.8	-0.433486
3	14.1	-0.683525
4	46.0	0.506958
...

822	54.8	0.835367
823	26.5	-0.220767
824	16.3	-0.601423
825	NaN	NaN
826	NaN	NaN

827 rows x 2 columns

Z-Score : **1,21**
Prix moyen : **32,42€**

Analyses univariées du prix

Intervalle interquartile

Mesures de dispersion pour la colonne 'price':

```
count    825.000000
mean      32.415636
std       26.795849
min        5.200000
25%       14.600000
50%       24.400000
75%       42.000000
max      225.000000
Name: price, dtype: float64
```

Intervalle interquartile (IQR): 27.4

Outliers

Seuil : **83,1€**

Nombre d'outliers : **37**

Justifiés ou pas ?

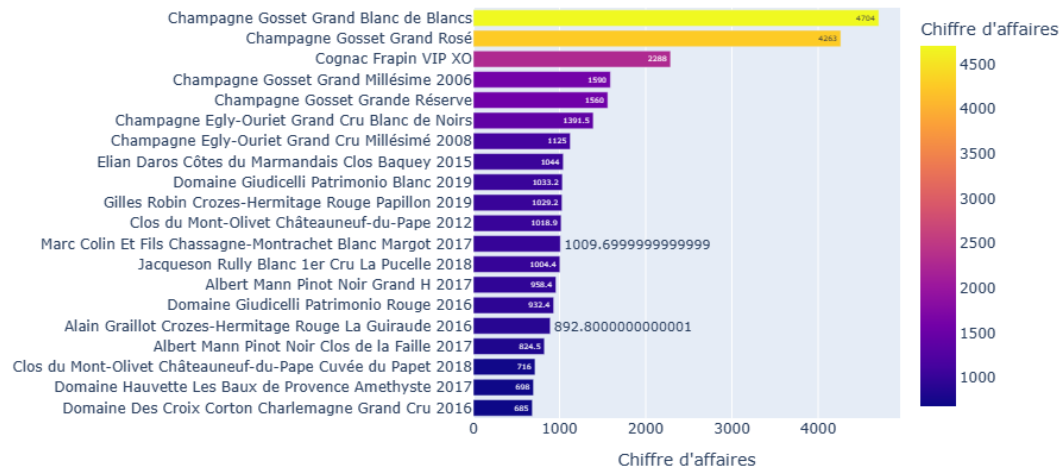


Analyses univariées du CA

CA total du site web

70 290,9€

Palmarès des 20 premiers articles en CA



80/20 (en nombre d'articles)

80 => 470

20 => 79

Actions pour la suite



Analyse du fichier Excel

- *Traitement des données vides*
- *Actions commerciales*
- *Ajout de données temporelles*
- *Ajout de produits sur le site web + création de sku*

	A	B	C	D	E	F	G	H	I
	product_id	onsale_web	price	stock_quantity	stock_status_2	id_web	sku	total_sales	post_title
1									
2	4041	1	32,7	15	instock	14975	14975	0	Pierre Gaillard Condrieu 2018
3	4042	1	31,2	34	instock	16042	16042	7	Pierre Gaillard Cornas 2017
4	4043	1	60	12	instock	14980	14980	3	Pierre Gaillard Côte Rôtie Esprit de Blonde 2017
5	4045	1	42,6	66	instock	16041	16041	14	Pierre Gaillard Côte Rôtie 2018
6	4047	1	18,3	36	instock	14977	14977	0	Pierre Gaillard Côtes-du-Rhône Blanc Les Gendrines 2018
7	4048	1	22,8	22	instock	16044	16044	7	Pierre Gaillard Saint-Joseph Rouge Clos de Cuminaille
8	4049	1	19,3	46	instock	16043	16043	1	Pierre Gaillard Saint-Joseph Rouge 2018
9	4051	1	7,7	41	instock	16045	16045	6	Jeanne Gaillard IGP Collines Rhodaniennes Syrah Terre
10	4052	1	33,7	35	instock	16030	16030	6	Clos du Mont-Olivet Châteauneuf-du-Pape Blanc 2019
11	4053	1	44,3	16	instock	13127	13127	23	Clos du Mont-Olivet Châteauneuf-du-Pape 2012
12	4057	1	8,7	112	instock	16039	16039	18	Mourgues du Grès Costières de Nîmes Galets Dorés 2018
13	4058	1	8,7	267	instock	16318	16318	11	Mourgues du Grès Costières de Nîmes Galets Rouges 2018
14	4059	1	8,7	36	instock	16275	16275	7	Mourgues du Grès Costières de Nîmes Galets Rosés 2018
15	4060	1	11,9	73	instock	16498	16498	9	Mourgues du Grès Costières de Nîmes Terre d'Argence 2018
16	4062	1	11,9	60	instock	16320	16320	3	Mourgues du Grès IGP Pont du Gard Terre d'Argence 2018
17	4063	1	14,5	10	instock	16319	16319	3	Mourgues du Grès Costières de Nîmes Terre de Feu 2018
18	4064	1	14,4	27	instock	15966	15966	3	Oratoire Saint Martin Cairanne Rouge Réserve des Seigneurs 2018
19	4065	1	19,5	35	instock	15022	15022	0	Oratoire Saint Martin Cairanne Rouge Les Douyes 2018

Point sur les compétences apprises

Qu'est-ce qui s'est bien passé pour vous dans ce travail de nettoyage ?

- *Les fusions de fichiers*
- *La compréhension des données dans les fichiers*

Qu'est-ce que vous avez trouvé le plus difficile ?

- *L'import de certains fichiers*
- *L'identification de valeurs qui ne respectent pas une règle (sku)*
- *La création de graphiques*
- *Les calculs basés sur le CA et les ventes*
- *Recommencer le travail après avoir éteint l'ordinateur (réimporter les packages)*

Sur quelles tâches est-ce que vous pensez avoir besoin de plus d'entraînement ?

- *L'identification de valeurs manquantes et les actions à entreprendre*
- *La création de graphiques*
- *Comprendre les enjeux de l'espace de stockage et des différentes notions sur Python*



Merci pour votre attention !

**P6 - Optimisez la gestion du stock d'une boutique en
nettoyant ses données**

Leonardo MONTILLA
Business Intelligence Analyst
22/12/2023