



Optimisation & Analyse des données

Manon DARGET

Business Intelligence Analyst

MARS 2024

Analyses Exploratoires des Données



df_erp

825 lignes
6 colonnes

	product_id	onsale_web	price	stock_quantity	stock_status	purchase_price
0	3847	1	24.2	16	instock	12.88
1	3849	1	34.3	10	instock	17.54
2	3850	1	20.8	0	outofstock	10.64
3	4032	1	14.1	26	instock	6.92
4	4039	1	46.0	3	outofstock	23.77

Traitements réalisés : Vérification des doublons 'product_id'

Dataset df_erp

Vérification des statuts de stock

	product_id	onsale_web	price	stock_quantity	stock_status	purchase_price	stock_status_2
4	4039	1	46.0	3	outofstock	23.77	instock
398	4885	1	18.7	0	instock	9.66	outofstock

	product_id	onsale_web	price	stock_quantity	stock_status	purchase_price	stock_status_2
151	4233	0	-20.0	0	outofstock	10.33	outofstock
469	5017	0	-8.0	0	outofstock	4.34	outofstock
739	6594	0	-9.1	19	instock	4.61	instock

Vérification des prix

Vérification des quantités en stock

	product_id	onsale_web	price	stock_quantity	stock_status	purchase_price	stock_status_2
449	4973	0	10.0	-10	outofstock	4.96	outofstock
573	5700	1	44.5	-1	outofstock	22.30	outofstock

Analyses Exploratoires des Données



df_web

1513 lignes

29 colonnes

Traitements réalisés : Suppression 11 colonnes inutiles

	sku	total_sales	tax_status	post_author	post_date	post_date_gmt	product_type	post_title	post_excerpt	post_status
8	NaN	NaN	NaN	NaN	NaT	NaT	NaN	NaN	NaN	NaN
20	NaN	NaN	NaN	NaN	NaT	NaT	NaN	NaN	NaN	NaN
30	NaN	NaN	NaN	NaN	NaT	NaT	NaN	NaN	NaN	NaN
37	NaN	NaN	NaN	NaN	NaT	NaT	NaN	NaN	NaN	NaN
41	NaN	NaN	NaN	NaN	NaT	NaT	NaN	NaN	NaN	NaN
...

**Suppression
valeurs nulles**

Dataset df_web

Vérification et suppression des doublons

	sku	total_sales	tax_status	post_author	post_date	post_date_gmt	product_type	post_title	post_excerpt
0	11862	3.0	NaN	2.0	2018-02-12 13:46:23	2018-02-12 12:46:23	Vin	Gilles Robin Hermitage Rouge 2012	NaN
1	16057	5.0	NaN	2.0	2018-04-17 15:29:17	2018-04-17 13:29:17	Vin	Domaine Pellé Sancerre Rouge La Croix Au Garde...	NaN
2	14692	5.0	taxable	2.0	2019-03-19 10:06:47	2019-03-19 09:06:47	Vin	Château Fonréaud Bordeaux Blanc Le Cygne 2016	<div>Grâce à la complémentarité des 3 cépages ...

Extrait de la liste des doublons df_web

Après traitement



df_web

714 lignes
18 colonnes

Analyses Exploratoires des Données



df_liaison

825 lignes
6 colonnes

Traitements réalisés : Vérification des doublons

	id_web	product_id
19	NaN	4055
49	NaN	4090
50	NaN	4092
119	NaN	4195
131	NaN	4209
...
817	NaN	7196

**Recherche articles
sans correspondances**

Fusion des données



Fusion en 2 temps

	product_id	onsale_web	price	stock_quantity	stock_status	purchase_price	id_web
0	3847	1	24.2	16	instock	12.88	15298
1	3849	1	34.3	10	instock	17.54	15296
2	3850	1	20.8	0	outofstock	10.64	15300
3	4032	1	14.1	26	instock	6.92	19814
4	4039	1	46.0	3	instock	23.77	19815
...

Extrait de df_merge, fusion entre df_erp et df_liaison

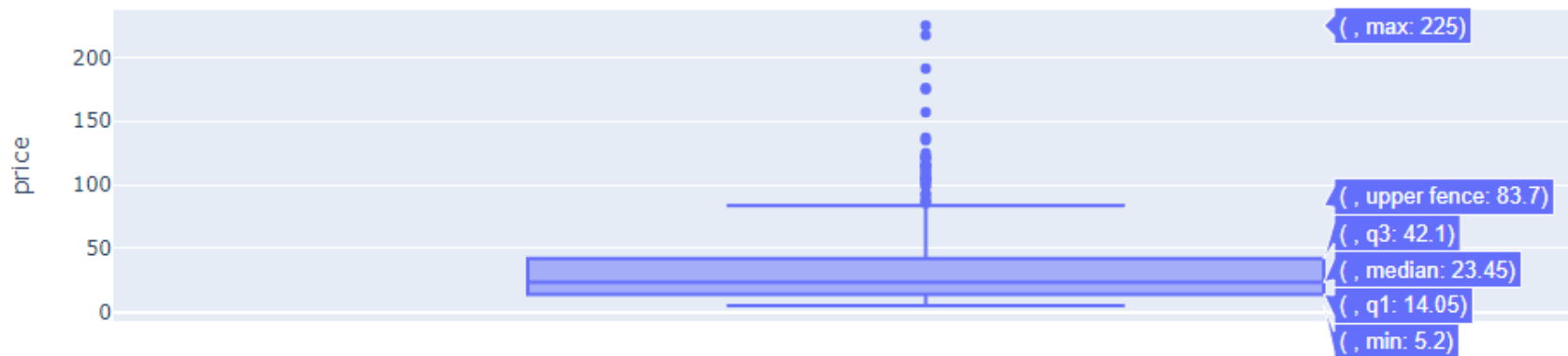


Fusion df_merge et df_web avec les clés id_web et sku

Analyses univariées du prix

Détection des anomalies de prix, 13 articles outliers/aberrants

Répartition des articles par prix



Analyses univariées du prix



Comparaison des outliers avec les prix du marché



CLOS DE VOUGEOT GRAND CRU ROUGE 2016 CAMILLE GIROUD 75CL

179,99 €

Prix extrait de www.comptoirnourisson.com

Analyse du CA

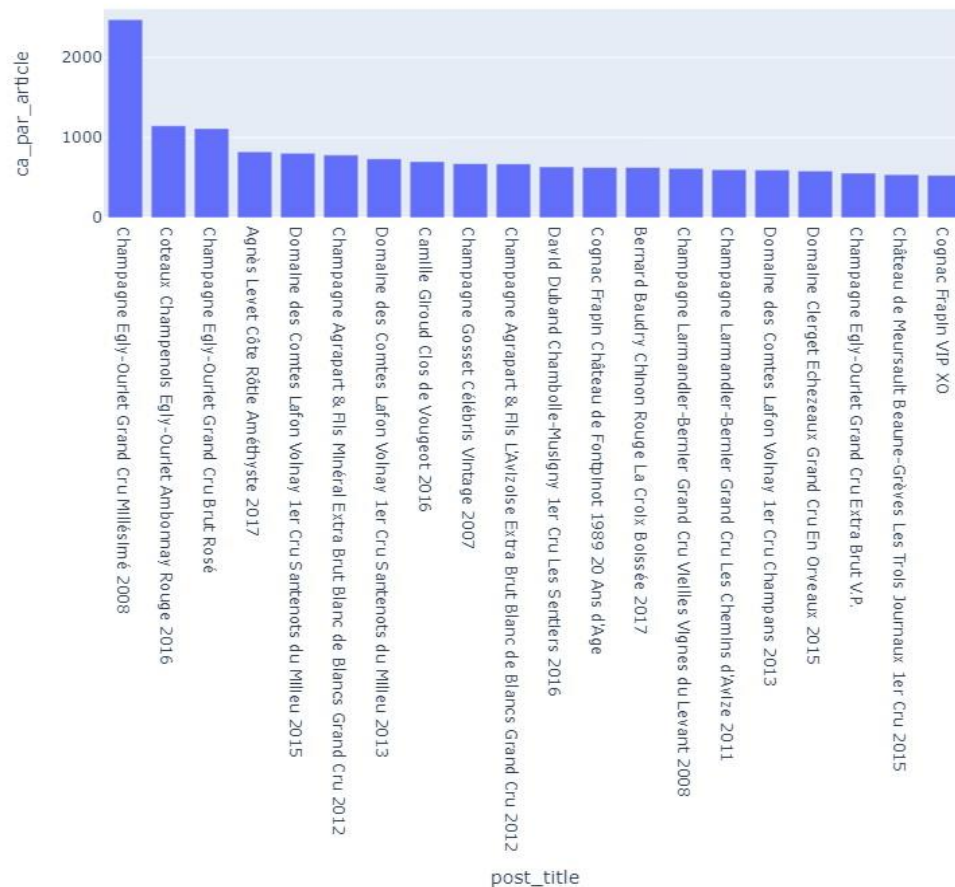


CA du site web
143 680 €



124 articles génèrent
80 % du CA

Top 20 des articles par CA

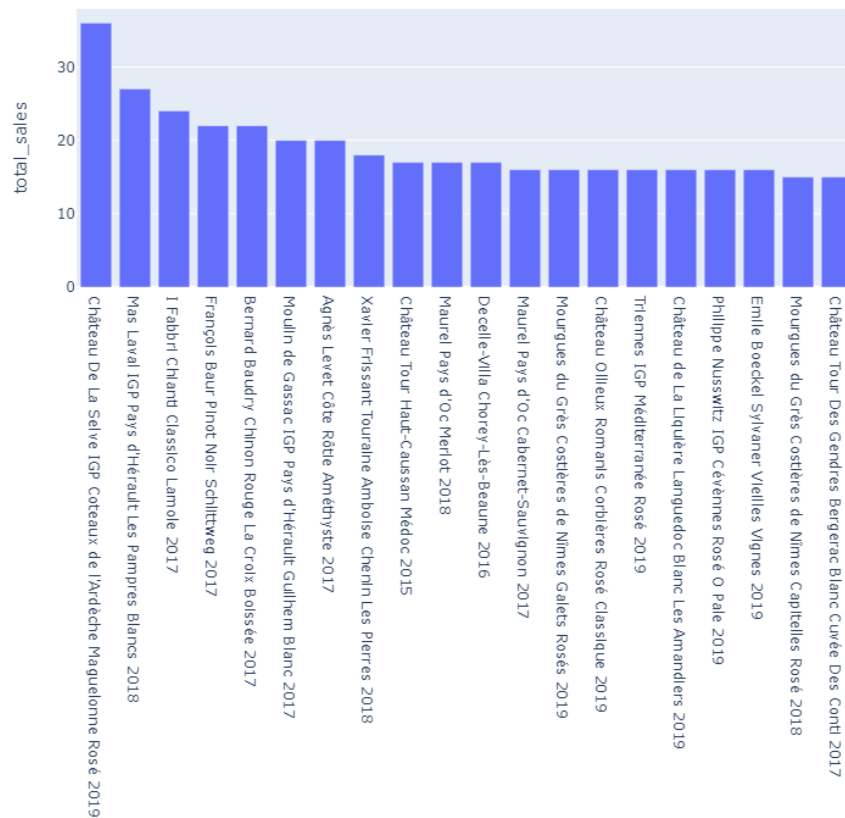


Analyse des ventes



**154 articles génèrent
80 % des quantités vendues**

Top 20 des articles par quantité vendue



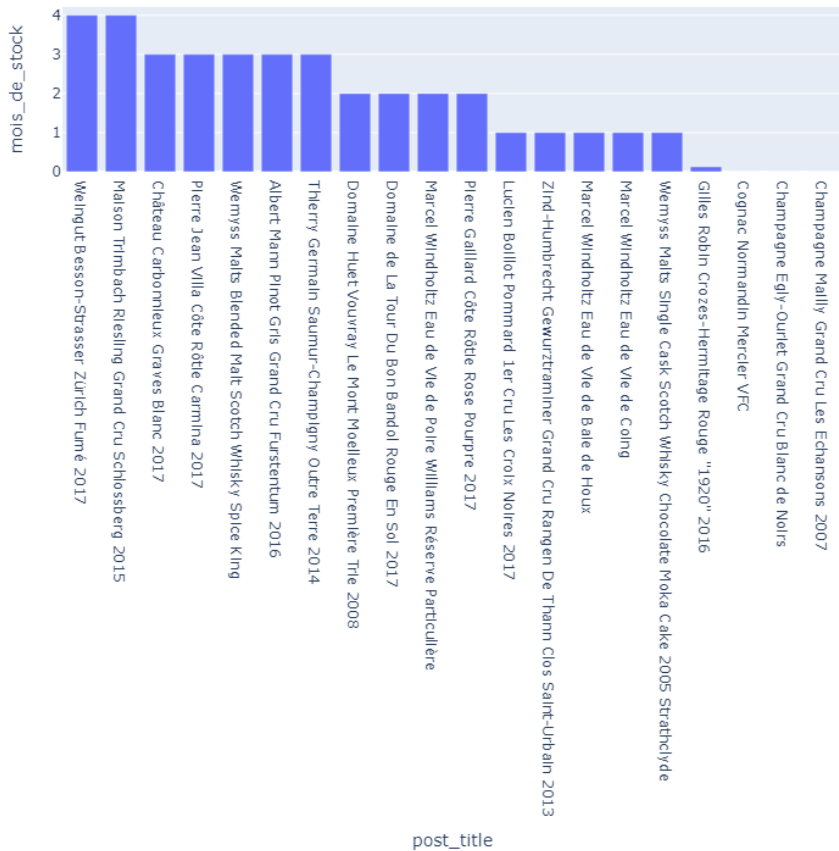
post_title

Analyse des stocks



**Moins de 20 articles ont une
marge de maintien de stock**

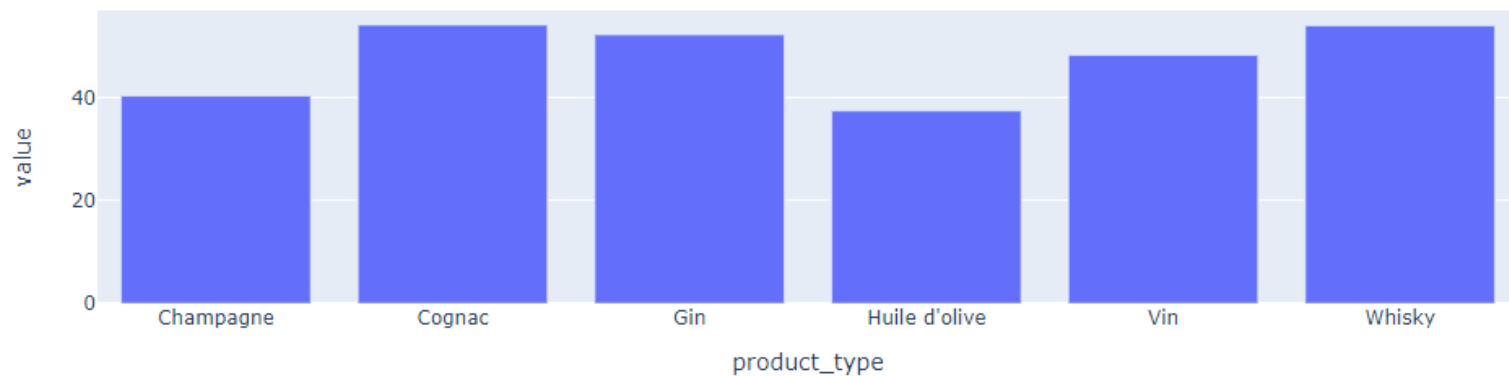
Flop 20 des articles ayant le plus de mois de stock



Analyse du taux de marge



Taux de marge par produit



Marge + importante sur alcools forts

Analyse correlations



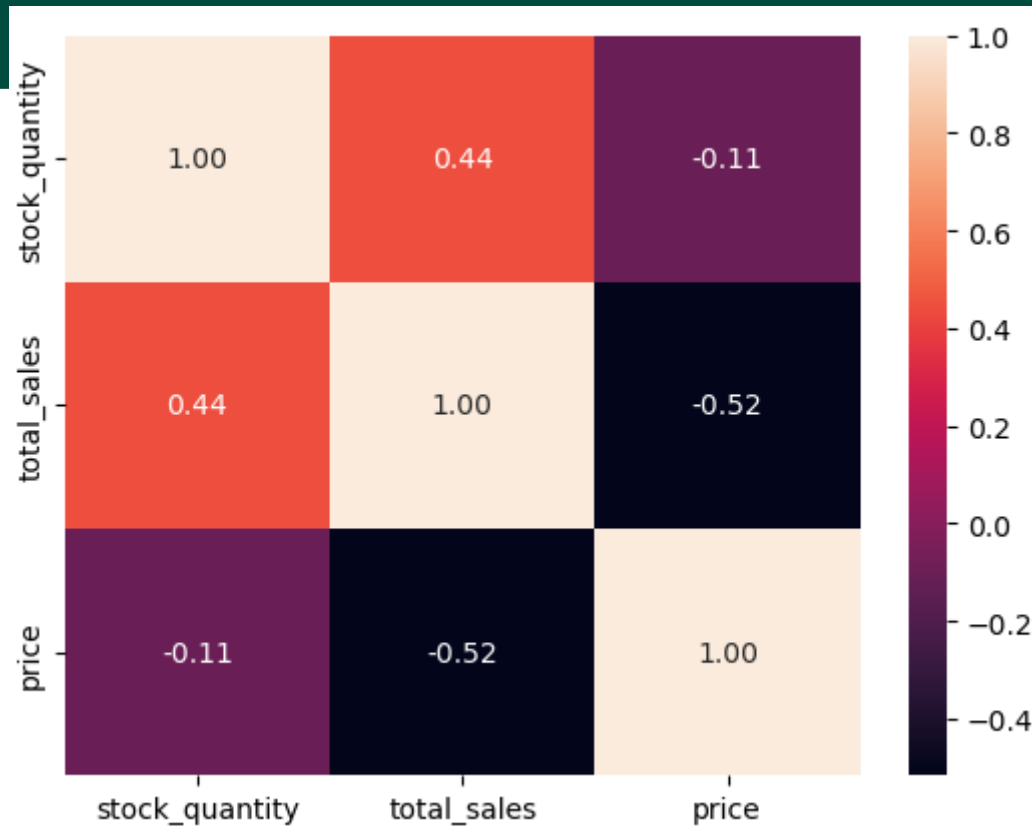
**Prix + bas pour les
stocks élevés**



**Ventes + importantes
sur les stocks élevés**



**Articles les + chers
se vendent moins**



Actions pour la suite



Ajout d'une variable taux_tva



Corrections des anomalies relevées



Commande auprès des fournisseurs

Point sur les compétences apprises

- **Qu'est-ce qui s'est bien passé pour vous dans ce travail de nettoyage ?**
- **Qu'est-ce que vous avez trouvé le plus difficile ?**
- **Sur quelles tâches est-ce que vous pensez avoir besoin de plus d'entraînement ?**



Merci de votre attention !

Manon DARGET

Business Intelligence Analyst

MARS 2024