



bottleneck
fine wine spirit

Missions

- ▶ Nettoyage des données
- ▶ Rapprochement des données de l'ERP et du CMS
- ▶ Calcul du chiffre d'affaire par produits
- ▶ Calcul du chiffre d'affaire total
- ▶ Analyse des prix des produits

Ressources

- ▶ Données venant de l'ERP
- ▶ Données venant du CMS
- ▶ Table de liaison

Outils

- Langage de programmation : Python



- Logiciel : Jupyter



Nettoyage des données

Première vu des données WEB

RangeIndex: 1513 entries, 0 to 1512

Data columns (total 28 columns):

#	Column	Non-Null Count	Dtype
0	sku	1428 non-null	object
1	virtual	1513 non-null	int64
2	downloadable	1513 non-null	int64
3	rating_count	1513 non-null	int64
4	average_rating	1430 non-null	float64
5	total_sales	1430 non-null	float64
6	tax_status	716 non-null	object
7	tax_class	0 non-null	float64
8	post_author	1430 non-null	float64
9	post_date	1430 non-null	object
10	post_date_gmt	1430 non-null	object
11	post_content	0 non-null	float64
12	post_title	1430 non-null	object
13	post_excerpt	716 non-null	object
14	post_status	1430 non-null	object
15	comment_status	1430 non-null	object
16	ping_status	1430 non-null	object
17	post_password	0 non-null	float64
18	post_name	1430 non-null	object
19	post_modified	1430 non-null	object
20	post_modified_gmt	1430 non-null	object
21	post_content_filtered	0 non-null	float64
22	post_parent	1430 non-null	float64
23	guid	1430 non-null	object
24	menu_order	1430 non-null	float64
25	post_type	1430 non-null	object
26	post_mime_type	714 non-null	object
27	comment_count	1430 non-null	float64

dtypes: float64(10), int64(3), object(15)

	sku	virtual	downloadable	rating_count	average_rating	total_sales	tax_status	tax_class	post_author	post_date	...	post_name	post_modified	p
0	bon-cadeau-25-euros	0	0	0	0.0	10.0	taxable	NaN	1.0	2018-06-01 13:53:46	...	bon-cadeau-de-25-euros	2018-06-01 14:13:57	
1	15298	0	0	0	0.0	6.0	taxable	NaN	2.0	2018-02-08 12:58:52	...	pierre-jean-villa-saint-joseph-preface-2018	2019-12-30 09:30:29	
2	15296	0	0	0	0.0	0.0	taxable	NaN	2.0	2018-02-08 13:49:41	...	pierre-jean-villa-saint-joseph-bide-2017	2019-12-21 09:00:17	
3	15300	0	0	0	0.0	0.0	taxable	NaN	2.0	2018-02-08 14:08:36	...	pierre-jean-villa-croze-hermitage-accroche-coe...	2020-06-26 18:15:03	
4	19814	0	0	0	0.0	3.0	taxable	NaN	2.0	2018-02-09 14:01:05	...	pierre-jean-villa-igp-gamine-2018	2020-01-04 16:36:01	
...	
1508	16135	0	0	0	0.0	5.0	NaN	NaN	2.0	2020-04-25 13:22:38	...	mouthes-le-bihan-aime-chai-2015	2020-08-26 17:35:03	
1509	15891	0	0	0	0.0	0.0	NaN	NaN	2.0	2020-05-02 14:53:40	...	camlin-larredy-a-jurancon-sec-la-virada-2018	2020-08-26 17:35:02	
1510	15887	0	0	0	0.0	0.0	NaN	NaN	2.0	2020-05-02 15:00:54	...	jamet-cote-rotie-fructus-voluptas-2018	2020-08-14 18:15:03	
1511	13127-1	0	0	0	0.0	0.0	NaN	NaN	2.0	2020-06-09 15:42:04	...	clos-du-mont-olivier-chateauneuf-du-pape-2007-2	2020-07-20 17:09:06	
1512	16230	0	0	0	0.0	0.0	NaN	NaN	2.0	2020-07-20 11:00:00	...	domaine-saint-nicolas-fiefs-vendeens-blanc-les...	2020-08-13 10:45:03	

Constats :

- ▶ La première colonne du DataFrame ne porte pas le bon nom
- ▶ Erreurs lexical (bon-cadeau-25-euros et 13127-1)
- ▶ Beaucoup de valeurs nulles (NaN) et de zéro
- ▶ Des dates au format datetime

Nettoyage des données

Correction du nom de la colonne erroné

#	Column	Non-Null Count	Dtype
0	sku	1428 non-null	object



```
# Renommer la colonne 'sku' en 'id_web'  
dataweb = dataweb.rename(columns={'sku': 'id_web'})
```



#	Column	Non-Null Count	Dtype
0	id_web	1428 non-null	object

Nettoyage des données

Traitement des valeurs nulles (NaN)

```
print(dataweb.isnull().sum())
```

id_web	85
virtual	0
downloadable	0
rating_count	0
average_rating	83
total_sales	83
tax_status	797
tax_class	1513
post_author	83
post_date	83
post_date_gmt	83
post_content	1513
post_title	83
post_excerpt	797
post_status	83
comment_status	83
ping_status	83
post_password	1513
post_name	83
post_modified	83
post_modified_gmt	83
post_content_filtered	1513
post_parent	83
guid	83
menu_order	83
post_type	83
post_mime_type	799
comment_count	83

- ▶ Beaucoup de colonnes comporte plusieurs lignes vide (83 lignes)
- ▶ 4 Colonnes comportant que des valeurs nulles (NaN)
- ▶ 3 Colonnes comportant beaucoup de valeurs nulle (NaN)

Nettoyage des données

Traitement des valeurs nulles (NaN)

```
# Remplacement des 0 par de NaN
```

```
dataweb = dataweb.replace(0, np.nan)  
print(len(dataweb.index), 'lignes avant la suppression des lignes NaN')  
dataweb.head()
```

1513 lignes avant la suppression des lignes NaN

	id_web	virtual	downloadable	rating_count	average_rating	total_sales	tax_status	tax_class	post_author	post_date	...	post_name	post_modified
0	bon-cadeau-25-euros	NaN	NaN	NaN	NaN	10.0	taxable	NaN	1.0	2018-06-01 13:53:46	...	bon-cadeau-de-25-euros	2018-06-01 14:13:57
1	15298	NaN	NaN	NaN	NaN	6.0	taxable	NaN	2.0	2018-02-08 12:58:52	...	pierre-jean-villa-saint-joseph-preface-2018	2019-12-30 09:30:29
2	15296	NaN	NaN	NaN	NaN	NaN	taxable	NaN	2.0	2018-02-08 13:49:41	...	pierre-jean-villa-saint-joseph-tilde-2017	2019-12-21 09:00:17
3	15300	NaN	NaN	NaN	NaN	NaN	taxable	NaN	2.0	2018-02-08 14:08:36	...	pierre-jean-villa-croze-hermitage-accroche-coe...	2020-06-26 18:15:03
4	19814	NaN	NaN	NaN	NaN	3.0	taxable	NaN	2.0	2018-02-09 14:01:05	...	pierre-jean-villa-igp-gamine-2018	2020-01-04 16:36:01

```
# Suppression des lignes de NaN
```

```
dataweb = dataweb.dropna(how = 'all')  
print(len(dataweb.index), 'lignes après la suppression des lignes NaN')
```

1430 lignes après la suppression des lignes NaN

- Remplacement de zéro par des valeur (NaN)
- Suppression des lignes contenant que des valeurs nulles (NaN)
- Passage de 1513 lignes à 1430 lignes dans le DataFrame

Nettoyage des données

Suppression des valeurs nulles sur la clé

```
dataweb[dataweb['id_web'].isnull()]
```

	id_web	virtual	downloadable	rating_count	average_rating	total_sales	tax_status	tax_class	post_author	post_date	...	post_name	post_modified
470	NaN	NaN	NaN	NaN	NaN	NaN	taxable	NaN	2.0	2018-07-31 12:07:23	...	pierre-jean-villa-cote-rotie-fongeant-2017	2019-11-02 13:24:15
471	NaN	NaN	NaN	NaN	NaN	NaN	taxable	NaN	2.0	2018-08-08 11:23:43	...	pierre-jean-villa-condrieu-suspendu-2018	2019-11-02 13:24:01

```
# Suppression des produits n'ayant pas de 'id_web'  
dataweb = dataweb.dropna(subset=['id_web'])
```

- Recherche des 'id_web' nulles (NaN)
- Suppression des 'id_web' nulles (NaN)

Nettoyage des données

Sélection des données intéressantes

```
dataweb = dataweb[['id_web', 'total_sales', 'post_date', 'post_title', 'post_excerpt', 'post_name', 'post_type']]
dataweb
```

	id_web	total_sales	post_date	post_title	post_excerpt	post_name	post_type
0	bon-cadeau-25-euros	10.0	2018-06-01 13:53:46	Bon cadeau de 25€	Parlons ...	bon-cadeau-de-25-euros	product
1	15298	6.0	2018-02-08 12:58:52	Pierre Jean Villa Saint-Joseph Préface 2018	Il exhale un nez subtil, racé avec du poivre f...	pierre-jean-villa-saint-joseph-preface-2018	product
2	15296	NaN	2018-02-08 13:49:41	Pierre Jean Villa Saint-Joseph Rouge Tildé 2017	Les vieilles vignes lui apportent une rare pro...	pierre-jean-villa-saint-joseph-tilde-2017	product
3	15300	NaN	2018-02-08 14:08:36	Pierre Jean Villa Crozes-Hermitage Accroche Co...	Dentelle de fruit de jeunes syrah, aux tanins ...	pierre-jean-villa-croze-hermitage-accroche-coe...	product
4	19814	3.0	2018-02-09 14:01:05	Pierre Jean Villa IGP Collines Rhodaniennes Ga...	Gamine représente tout le fruité et la gourman...	pierre-jean-villa-igp-gamine-2018	product
...
1508	16135	5.0	2020-04-25 13:22:38	Mouthes Le Bihan Côtes de Duras L'Aimé Chai 2015	NaN	mouthes-le-bihan-aime-chai-2015	attachment
1509	15891	NaN	2020-05-02 14:53:40	Camin Larredya Jurançon Sec La Virada 2018	NaN	camin-larredya-jurancon-sec-la-virada-2018	attachment
1510	15887	NaN	2020-05-02 15:00:54	Domaine Jamet Côte Rôtie Fructus Voluptas 2018	NaN	jamet-cote-rotie-fructus-voluptas-2018	attachment
1511	13127-1	NaN	2020-06-09 15:42:04	Clos du Mont-Olivet Châteauneuf-du-Pape 2007	NaN	clos-du-mont-olivier-chateauneuf-du-pape-2007-2	attachment
1512	16230	NaN	2020-07-20 11:00:00	Domaine Saint-Nicolas Vin de France Blanc Les ...	NaN	domaine-saint-nicolas-fiefs-vendeens-blanc-les...	attachment

1428 rows x 7 columns

- Reconstruction du DataFrame avec les données intéressante pour notre étude

Nettoyage des données

Suppression des doublons sur la clé

```
dataweb.loc[dataweb['id_web'].duplicated(keep=False),:]
```

	id_web	total_sales	post_date	post_title	post_excerpt	post_name	post_type
0	bon-cadeau-25-euros	10.0	2018-06-01 13:53:46	Bon cadeau de 25€	Parlons ...	bon-cadeau-de-25-euros	product
1	15298	6.0	2018-02-08 12:58:52	Pierre Jean Villa Saint-Joseph Préface 2018	Il exhale un nez subtil, racé avec du poivre f...	pierre-jean-villa-saint-joseph-preface-2018	product
2	15296	NaN	2018-02-08 13:49:41	Pierre Jean Villa Saint-Joseph Rouge Tildé 2017	Les vieilles vignes lui apportent une rare pro...	pierre-jean-villa-saint-joseph-tilde-2017	product
3	15300	NaN	2018-02-08 14:08:36	Pierre Jean Villa Crozes-Hermitage Accroche Co...	Dentelle de fruit de jeunes syrah, aux tanins ...	pierre-jean-villa-croze-hermitage-accroche-coe...	product
4	19814	3.0	2018-02-09 14:01:05	Pierre Jean Villa IGP Collines Rhodaniennes Ga...	Gamine représente tout le fruité et la gourman...	pierre-jean-villa-igp-gamine-2018	product
...
1508	16135	5.0	2020-04-25 13:22:38	Mouthes Le Bihan Côtes de Duras L'Aimé Chai 2015	NaN	mouthes-le-bihan-aime-chai-2015	attachment
1509	15891	NaN	2020-05-02 14:53:40	Camin Larredya Jurançon Sec La Virada 2018	NaN	camin-larredya-jurancon-sec-la-virada-2018	attachment
1510	15887	NaN	2020-05-02 15:00:54	Domaine Jamet Côte Rôtie Fructus Voluptas 2018	NaN	jamet-cote-rotie-fructus-voluptas-2018	attachment
1511	13127-1	NaN	2020-06-09 15:42:04	Clos du Mont-Olivet Châteauneuf-du-Pape 2007	NaN	clos-du-mont-olivier-chateauneuf-du-pape-2007-2	attachment
1512	16230	NaN	2020-07-20 11:00:00	Domaine Saint-Nicolas Vin de France Blanc Les ...	NaN	domaine-saint-nicolas-fiefs-vendeens-blanc-les...	attachment

1428 rows x 7 columns

```
# Suppression des doublons
```

```
dataweb = dataweb.drop_duplicates(subset=['id_web'])
dataweb
```

- Tout les produits ont des doublons
- Chaque produit est attacher à une pièce joint (image,etc...)

Nettoyage des données

Vérification des erreurs lexical

```
dataweb[dataweb['id_web'] == '13127-1']
```

	id_web	total_sales	post_date	post_title	post_excerpt	post_name	post_type
797	13127-1	NaN	2020-06-09 15:42:04	Clos du Mont-Olivet Châteauneuf-du-Pape 2007	Nez gracieux, très élégant avec une touche flo...	clos-du-mont-oliv-et-chateauneuf-du-pape-2007-2	product

```
dataweb[dataweb['id_web'] == '13127']
```

	id_web	total_sales	post_date	post_title	post_excerpt	post_name	post_type
18	13127	23.0	2018-02-12 11:18:33	Clos du Mont-Olivet Châteauneuf-du-Pape 2012	Nez gracieux, très élégant avec une touche flo...	clos-du-mont-oliv-et-chateauneuf-du-pape-2012	product

- ▶ Ce sont les même produits avec des caractéristiques différentes

Nettoyage des données

Remplacement des valeurs nulles (Nan)

```
dataweb = dataweb.replace(np.nan, 0)
dataweb.sort_values(by=['id_web'], ascending=False)
```

	id_web	total_sales	post_date	post_title	post_excerpt	post_name	post_type
0	bon-cadeau-25-euros	10.0	2018-06-01	Bon cadeau de 25€	Parlons ...	bon-cadeau-de-25-euros	product
422	9937	4.0	2018-05-17	Marcel Richaud Cairanne Rouge Les Estrambords ...	Un vin généreux, profond et intense. Les notes...	marcel-richaud-cairanne-rouge-les-estrambords-...	product
215	9636	0.0	2018-03-22	Wemyss Malts Blended Malt Scotch Whisky The Hi...	The Hive est un assemblage de malts du Speysid...	whisky-wemyss-the-hive-12-ans	product
206	9562	0.0	2018-03-02	Champagne Larmandier-Bernier Latitude	La cuvée Latitude est très équilibrée entre fr...	champagne-larmandier-bernier-latitude	product
531	8463	0.0	2018-11-26	Domaine Huet Vouvray Le Mont Moelleux Première...	Harmonie majestueuse dans ce vin d'exception, ...	domaine-huet-vouvray-le-mont-moelleux-premiere...	product
...
257	11049	0.0	2018-04-13	Alphonse Mellot Sancerre Rouge En Grands Champ...	Superbe robe carmin profond. Un très grand ter...	alphonse-mellot-sancerre-rouge-en-grands-champ...	product
281	10814	0.0	2018-04-17	Thierry Germain Saumur-Champigny Outre Terre 2013	La robe est rouge vif. Le nez est très floral,...	thierry-germain-saumur-champigny-outre-terre-2013	product
321	10775	0.0	2018-04-17	Albert Mann Pinot Gris Vendanges Tardives Alte...	Belle bouteille dotée d'une grande complexité ...	albert-mann-pinot-gris-vendanges-tardives-alte...	product
255	10459	0.0	2018-04-13	Alphonse Mellot Sancerre Rouge Génération XIX ...	Robe d'un superbe rubis, dense et profond - As...	alphonse-mellot-sancerre-rouge-generation-xix-...	product
628	10014	0.0	2019-04-04	Darnley's London Dry Gin Original	Un gin léger, floral et aromati...	darnleys-london-dry-gin-original	product

714 rows x 7 columns

- Remplacer les valeur nulles (Nan) de la colonne 'total_sales' par des zéro

Nettoyage des données

Première vu des données ERP

```
RangeIndex: 825 entries, 0 to 824  
Data columns (total 5 columns):  
#   Column          Non-Null Count  Dtype  
---  ---  
0    product_id      825 non-null    int64  
1    onsale_web       825 non-null    int64  
2    price           825 non-null    float64  
3    stock_quantity  825 non-null    int64  
4    stock_status     825 non-null    object  
dtypes: float64(1), int64(3), object(1)
```

	product_id	onsale_web	price	stock_quantity	stock_status
0	3847	1	24.2	0	outofstock
1	3849	1	34.3	0	outofstock
2	3850	1	20.8	0	outofstock
3	4032	1	14.1	0	outofstock
4	4039	1	46.0	0	outofstock
...
820	7203	0	45.0	30	instock
821	7204	0	45.0	9	instock
822	7247	1	54.8	23	instock
823	7329	0	26.5	14	instock
824	7338	1	16.3	45	instock

825 rows × 5 columns

Constats :

- ▶ Aucune valeur nulles (NaN)
- ▶ Pas d'erreur lexical

Nettoyage des données

Vérification des valeurs nulles (NaN) et des doublons

```
print(dataerp.isnull().sum())
```

```
product_id      0  
onsale_web      0  
price           0  
stock_quantity  0  
stock_status    0  
dtype: int64
```

► Aucune valeur nulles (NaN)

```
dataerp.loc[dataerp['product_id'].duplicated(keep=False),:]
```

```
product_id  onsale_web  price  stock_quantity  stock_status
```

► Aucun doublon

Nettoyage des données

Recherche et correction des erreurs de saisie

```
# Recherche erreurs de saisie 'instock'  
  
instock = dataerp[dataerp['stock_status'] == 'instock']  
errorinstock = instock[instock['stock_quantity'] == 0]  
  
errorinstock
```

	product_id	onsale_web	price	stock_quantity	stock_status
443	4954	1	25.0	0	instock

```
# Recherche erreurs de saisie 'outofstock'  
  
outofstock = dataerp[dataerp['stock_status'] == 'outofstock']  
erroroutofstock = outofstock[outofstock['stock_quantity'] != 0]  
  
erroroutofstock
```

	product_id	onsale_web	price	stock_quantity	stock_status
--	------------	------------	-------	----------------	--------------

- 1 erreur détecter au niveau du '*stock_quantity*' ou du '*stock_status*'

```
dataerp['stock_status'] = np.where(dataerp['stock_quantity'] == 0, 'outofstock', 'instock')
```

- Correction des stocks

Nettoyage des données

Première vu des données de liaison

```
RangeIndex: 825 entries, 0 to 824  
Data columns (total 2 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   product_id  825 non-null    int64  
1   id_web      734 non-null    object  
dtypes: int64(1), object(1)
```

	product_id	id_web
0	3847	15298
1	3849	15296
2	3850	15300
3	4032	19814
4	4039	19815
...
820	7203	NaN
821	7204	NaN
822	7247	13127-1
823	7329	14680-1
824	7338	16230

825 rows x 2 columns

- ▶ Aucune valeur nulles (NaN) dans la colonne 'product_id'
- ▶ Des valeurs nulles (NaN) dans la colonne 'id_web'

Nettoyage des données

Vérification et traitement des valeurs nulles (NaN) et des doublons

```
# Recherche des valeurs null (NaN)

print(dataliaison.isnull().sum())

product_id      0
id_web          91
dtype: int64
```

```
# Suppression des valeurs null (NaN)

dataliaison = dataliaison.dropna()
```

```
# Recherche des doublons
```

```
dataliaison.loc[dataliaison['id_web'].duplicated(keep=False),:]
```

```
product_id  id_web
```

► Suppression des 91 valeurs nulles (NaN)

► Aucun doublon

Rapprochement des données

Premier merge

- Premier merge entre les données de liaison et les données ERP

```
dataLE = pd.merge(data liaison, dataerp, how="inner", on=["product_id"])  
dataLE
```

	product_id	id_web	onsale_web	price	stock_quantity	stock_status
0	3847	15298	1	24.2	0	outofstock
1	3849	15296	1	34.3	0	outofstock
2	3850	15300	1	20.8	0	outofstock
3	4032	19814	1	14.1	0	outofstock
4	4039	19815	1	46.0	0	outofstock
...
729	7023	15891	1	27.5	15	instock
730	7025	15887	1	69.0	2	instock
731	7247	13127-1	1	54.8	23	instock
732	7329	14680-1	0	26.5	14	instock
733	7338	16230	1	16.3	45	instock

734 rows × 6 columns

Rapprochement des données

Deuxième merge

- Deuxième merge entre le premier merge et les données WEB

```
data = pd.merge(dataLE, dataweb, how="inner", on=["id_web"])
data
```

	product_id	id_web	onsale_web	price	stock_quantity	stock_status	total_sales	post_date	post_title	post_excerpt	post_name	post_type
0	3847	15298	1	24.2	0	outofstock	6.0	2018-02-08	Pierre Jean Villa Saint-Joseph Préface 2018	Il exhale un nez subtil, racé avec du poivre f...	pierre-jean-villa-saint-joseph-preface-2018	product
1	3849	15296	1	34.3	0	outofstock	0.0	2018-02-08	Pierre Jean Villa Saint-Joseph Rouge Tildé 2017	Les vieilles vignes lui apportent une rare pro...	pierre-jean-villa-saint-joseph-tilde-2017	product
2	3850	15300	1	20.8	0	outofstock	0.0	2018-02-08	Pierre Jean Villa Crozes-Hermitage Accroche Co...	Dentelle de fruit de jeunes syrah, aux tanins ...	pierre-jean-villa-croze-hermitage-accroche-coe...	product
3	4032	19814	1	14.1	0	outofstock	3.0	2018-02-09	Pierre Jean Villa IGP Collines Rhodaniennes Ga...	Gamine représente tout le fruité et la gourman...	pierre-jean-villa-igp-gamine-2018	product
4	4039	19815	1	46.0	0	outofstock	0.0	2018-02-12	Pierre Jean Villa Côte Rôtie Carmina 2017	Le côte rôtie Carmina monte en puissance mais ...	pierre-jean-villa-cote-rotie-carmina-2017	product
...
709	6930	16135	1	8.4	83	instock	5.0	2020-04-25	Mouthes Le Bihan Côtes de Duras L'Aimé Chai 2015	Belle robe jeune à dominante rubis soutenue. N...	mouthes-le-bihan-aime-chai-2015	product
710	7023	15891	1	27.5	15	instock	0.0	2020-05-02	Camin Larredya Jurançon Sec La Virada 2018	L'exotisme du nez est complété par d'élégantes...	camin-larredya-jurancon-sec-la-virada-2018	product
711	7025	15887	1	69.0	2	instock	0.0	2020-05-02	Domaine Jamet Côte Rôtie Fructus Voluptas 2018	Pour cette cuvée, Jean-Paul Jamet recherche un...	jamet-cote-rotie-fructus-voluptas-2018	product
712	7247	13127-1	1	54.8	23	instock	0.0	2020-06-09	Clos du Mont-Olivet Châteauneuf-du-Pape 2007	Nez gracieux, très élégant avec une touche flo...	clos-du-mont-oliv-et-chateauneuf-du-pape-2007-2	product
713	7338	16230	1	16.3	45	instock	0.0	2020-07-20	Domaine Saint-Nicolas Vin de France Blanc Les ...	Issu d'un assemblage de chenin blanc et de cha...	domaine-saint-nicolas-tiefs-vendeens-blanc-les...	product

714 rows x 12 columns

Analyse des données

Calcul du CA par produit et du CA total

```
data = data.assign(CA = data['price']*data['total_sales'])  
data[['product_id', 'id_web', 'price', 'total_sales', 'post_title', 'CA']]
```

	product_id	id_web	price	total_sales	post_title	CA
0	3847	15298	24.2	6.0	Pierre Jean Villa Saint-Joseph Préface 2018	145.2
1	3849	15296	34.3	0.0	Pierre Jean Villa Saint-Joseph Rouge Tildé 2017	0.0
2	3850	15300	20.8	0.0	Pierre Jean Villa Crozes-Hermitage Accroche Co...	0.0
3	4032	19814	14.1	3.0	Pierre Jean Villa IGP Collines Rhodaniennes Ga...	42.3
4	4039	19815	46.0	0.0	Pierre Jean Villa Côte Rôtie Carmina 2017	0.0
...
709	6930	16135	8.4	5.0	Mouthes Le Bihan Côtes de Duras L'Aimé Chai 2015	42.0
710	7023	15891	27.5	0.0	Camin Larredya Jurançon Sec La Virada 2018	0.0
711	7025	15887	69.0	0.0	Domaine Jamet Côte Rôtie Fructus Voluptas 2018	0.0
712	7247	13127-1	54.8	0.0	Clos du Mont-Olivet Châteauneuf-du-Pape 2007	0.0
713	7338	16230	16.3	0.0	Domaine Saint-Nicolas Vin de France Blanc Les ...	0.0

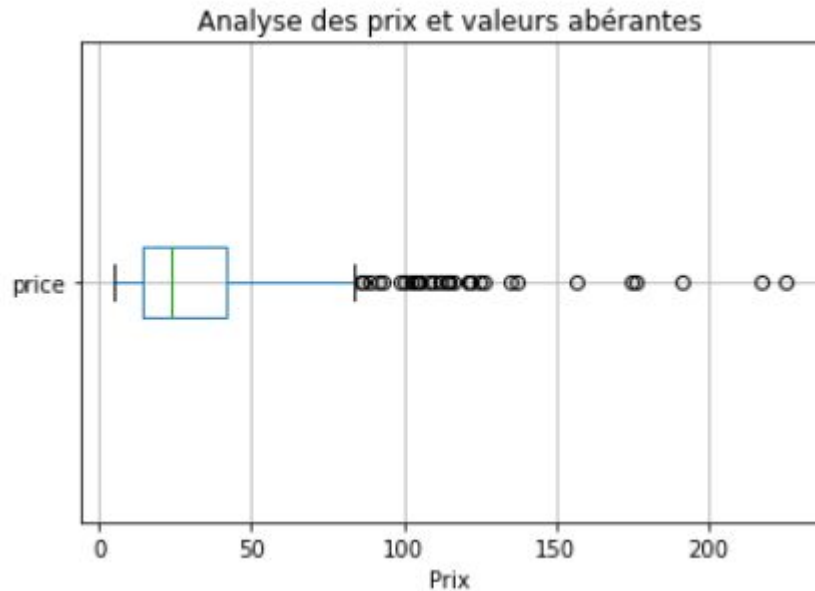
714 rows × 6 columns

Chiffre d'affaire total
70568,6 €

- Ajout d'une colonne 'CA' qui présente le chiffre d'affaire de chaque produits

Analyse des données

Analyse des prix des produits



Moyenne des prix : 32.49
Médiane : 23.55
Q1 = 14
Q3 = 42
IQ = 28
min = 5.2
max = 84.0

- ▶ 50% des prix sont compris entre 14€ et 42€
- ▶ 75% des prix sont compris entre 5,2€ et 42€
- ▶ Les Outlier sont supérieur à 84€

Analyses supplémentaire

Etudes des ventes

► Top 5 des produits les plus vendu

product_id	id_web	onsale_web	price	stock_quantity	stock_status	total_sales	post_date	post_title	post_excerpt	post_name	post_type	CA	
194	4334	7818	1	49.0	0	outofstock	96.0	2018-03-01	Champagne Gosset Grand Blanc de Blancs	La bulle fine et presente se dirige vers la su...	champagne-gosset-grand-blanc-de-blanc	product	4704.0
71	4144	1662	1	49.0	11	instock	87.0	2018-02-13	Champagne Gosset Grand Rosé	Le nez exprime sans détour la fraise, dans sa ...	champagne-gosset-grand-rose	product	4263.0
30	4068	16416	1	16.6	157	instock	62.0	2018-02-12	Gilles Robin Crozes-Hermitage Rouge Papillon 2019	Avec ses arômes de cassis, de myrtille, de cer...	gilles-robin-crozes-hermitage-papillon-2019	product	1029.2
119	4200	16295	1	5.8	190	instock	46.0	2018-02-15	Moulin de Gassac IGP Pays d'Hérault Guilhem Ro...	Rosé très agréable, intense, floral et minéral...	moulin-de-gassac-igp-pays-dherault-guilhem-ros...	product	266.8
97	4172	16210	1	5.7	167	instock	43.0	2018-02-14	Maurel Pays d'Oc Chardonnay 2019	Un très joli petit chardonnay qui nous révèle ...	maurel-pays-oc-chardonnay-2019	product	245.1

Analyses supplémentaire

Etudes des ventes

► Top 5 des produits avec le plus gros CA

	product_id	id_web	onsale_web	price	stock_quantity	stock_status	total_sales	post_date	post_title	post_excerpt	post_name	post_type	CA
194	4334	7818	1	49.0	0	outofstock	96.0	2018-03-01	Champagne Gosset Grand Blanc de Blancs	La bulle fine et presente se dirige vers la su...	champagne-gosset-grand-blanc-de-blanc	product	4704.0
71	4144	1662	1	49.0	11	instock	87.0	2018-02-13	Champagne Gosset Grand Rosé	Le nez exprime sans détour la fraise, dans sa ...	champagne-gosset-grand-rose	product	4263.0
218	4402	3510	1	176.0	8	instock	13.0	2018-03-22	Cognac Frapin VIP XO	La cuvée VIP XO à été enrichie d'eaux-de-vie t...	cognac-frapin-vip-xo	product	2288.0
70	4142	11641	1	53.0	8	instock	30.0	2018-02-13	Champagne Gosset Grand Millésime 2006	L'attaque est ample, gourmande avec une belle ...	champagne-gosset-grand-millesime-2006	product	1590.0
69	4141	304	1	39.0	1	instock	40.0	2018-02-13	Champagne Gosset Grande Réserve	Le nez, ouvert et expressif, évoque les fruits...	gosset-champagne-grande-reserve	product	1560.0

Synthèse sur le nettoyage

Fichier WEB :

- Beaucoup de données manquantes
- Des colonnes non-utilisé
- Des doublon pour chaque produits
- Des erreurs lexical à revoir

Fichier ERP :

- Faire attention au cohésion entre la quantité et le statut des stocks

Fichier Liaison :

- Quelques valeurs nulles

Synthèse sur l'analyse

- ▶ Chiffre d'affaire de **70568,6**
- ▶ **50%** des prix sont compris entre **14€** et **42€**
- ▶ Les Outlier sont des **valeurs Extrême**
- ▶ Le produit ayant le plus gros CA a été vendu dans sa **totalité**