



Aloe Vera



Table of contents



01 The datasets

02 Cleaning

03 Analysis





01



The datasets

Sales
Customers
Salespeople

The datasets



Sales

primary_key	order_reference	code_cip_id	salesrep_id	article_code	article_name	quantity	unit_price	discount	net_revenue	order_date
19560	2105202234S0001	2260084	26	Vide	Vide	373	1705.23	38%	1057.24	22/05/2020
19561	2105202234S0001	2260084	26	5046	Nutri-Pulpe 1000ml - 8.026.380.050.468	12	24.9	38%	185.26	22/05/2020
19562	2105202234S0001	2260084	26	5047	Nutri-Pulpe 500ml - 8.026.380.050.475	12	16.9	38%	125.74	22/05/2020
19563	2105202234S0001	2260084	26	5048	Nutri-Pulpe Pêche Blanche 1000ml - 8.026.380.050.482	6	24.9	38%	92.63	22/05/2020
19564	2105202234S0001	2260084	26	5049	Nutri-Pulpe Pêche Blanche 500ml - 8.026.380.050.499	6	16.9	38%	62.87	22/05/2020
19565	2105202234S0001	2260084	26	5050	Pur Gel Feuille 250ml - 8.026.380.050.505	36	16.9	38%	377.21	22/05/2020
19566	2105202234S0001	2260084	26	5051	Pur Gel Flacon 100ml - 8.026.380.050.512	24	9.9	38%	147.31	22/05/2020
19567	2105202234S0001	2260084	26	5052	Aloe-Oil Huile Réparatrice 50ml - 8.026.380.050.529	12	8.9	38%	66.22	22/05/2020
19568	2105202234S0001	2260084	26	5064	Cœur d'Aloe Cube	60		0%		22/05/2020
19569	2105202234S0001	2260084	26	5065	Totem Double Face	1		0%		22/05/2020
19570	2105202234S0001	2260084	26	5150	Pop-Up Coeur D'Aloe Cube Offre Consommateur	1		0%		22/05/2020
19571	2105202234S0001	2260084	26	5153	Présentoir de Sol Gamme Aloe	1		0%		22/05/2020
19572	2105202234S0001	2260084	26	5154	Grand Aloe Vera Cartonné de Linéaire ou Vitrine	2		0%		22/05/2020
19573	2105202234S0001	2260084	26	5156	Brochure Gamme Aloe Vera	50		0%		22/05/2020
19574	2105202234S0001	2260084	26	5177	Echantillon Nutri-Pulpe	50		0%		22/05/2020
19575	2105202234S0001	2260084	26	5586	Echantillon Pur Gel	50		0%		22/05/2020
19576	2105202234S0001	2260084	26	5869	Echantillon Nutri-Pulpe Goût Pêche Blanche	50		0%		22/05/2020
19577	2005201635S0001	2213034	26	Vide	Vide	397	2050.82	38%	1271.51	22/05/2020
19578	2005201635S0001	2213034	26	5046	Nutri-Pulpe 1000ml - 8.026.380.050.468	12	24.9	38%	185.26	22/05/2020
19579	2005201635S0001	2213034	26	5047	Nutri-Pulpe 500ml - 8.026.380.050.475	12	16.9	38%	125.74	22/05/2020
19580	2005201635S0001	2213034	26	5048	Nutri-Pulpe Pêche Blanche 1000ml - 8.026.380.050.482	12	24.9	38%	185.26	22/05/2020
19581	2005201635S0001	2213034	26	5049	Nutri-Pulpe Pêche Blanche 500ml - 8.026.380.050.499	12	16.9	38%	125.74	22/05/2020
19582	2005201635S0001	2213034	26	5050	Pur Gel Feuille 250ml - 8.026.380.050.505	24	16.9	38%	251.47	22/05/2020
19583	2005201635S0001	2213034	26	5051	Pur Gel Flacon 100ml - 8.026.380.050.512	24	9.9	38%	147.31	22/05/2020
19584	2005201635S0001	2213034	26	5052	Aloe-Oil Huile Réparatrice 50ml - 8.026.380.050.529	24	8.9	38%	132.43	22/05/2020
19585	2005201635S0001	2213034	26	5053	Aloe-Oil Crème Réparatrice 150ml - 8.026.380.050.536	12	15.9	38%	118.3	22/05/2020

Information about sales including:

Product names

Date and total amount of the sale

Code of the salesperson who made the sale

CIP code of the customer who made the purchase (personal code)

The datasets



Customers



uga	code_cip_id	pharmacies_name	appellation_titulaire	owner_last_name	owner_first_name	commune	postal_code
01BOU	2267205	PHARMACY NUM1	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2257464	PHARMACY NUM2	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2243163	PHARMACY NUM3	M	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055070	PHARMACY NUM4	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055086	PHARMACY NUM5	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055091	PHARMACY NUM6	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055112	PHARMACY NUM7	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055149	PHARMACY NUM8	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2055175	PHARMACY NUM9	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2003379	PHARMACY NUM10	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2004604	PHARMACY NUM11	M	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2000435	PHARMACY NUM12	M	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2000461	PHARMACY NUM13	M	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2000016	PHARMACY NUM14	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2002855	PHARMACY NUM15	M	Nom de Famille	Prénom	Nom Retiré	1010
01BOU	2000257	PHARMACY NUM16	Mme	Nom de Famille	Prénom	Nom Retiré	1010
69VFS	2055505	PHARMACY NUM17	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2054941	PHARMACY NUM18	M	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2000084	PHARMACY NUM19	M	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2000215	PHARMACY NUM20	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2055573	PHARMACY NUM21	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2055589	PHARMACY NUM22	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2055615	PHARMACY NUM23	M	Nom de Famille	Prénom	Nom Retiré	1010
01NAN	2282710	PHARMACY NUM24	M	Nom de Famille	Prénom	Nom Retiré	1010
01GEX	2005357	PHARMACY NUM25	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01GEX	2055636	PHARMACY NUM26	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01GEX	2055641	PHARMACY NUM27	M	Nom de Famille	Prénom	Nom Retiré	1010
01GEX	2263200	PHARMACY NUM28	Mme	Nom de Famille	Prénom	Nom Retiré	1010
01BEL	2004130	PHARMACY NUM29	M	Nom de Famille	Prénom	Nom Retiré	1010

Information about customers including:

Name of the pharmacy

First and last name of the manager

Location of the pharmacy

CIP code (personal code)

The datasets



Salespeople

salesrep_id	last_name	first_name
1	Martin	Marie
2	Bernard	Jeanne
3	Thomas	Françoise
4	Petit	Monique
5	Robert	Catherine
6	Richard	Nathalie
7	Durand	Isabelle
8	Dubois	Jacqueline
9	Moreau	Sylvie
10	Laurent	Anne
11	Simon	Martine
12	Michel	Madeleine
13	Lefebvre	Nicole
14	Leroy	Suzanne
15	Roux	Hélène
16	David	Christine
17	Bertrand	Louise
18	Morel	Marguerite
19	Fournier	Denise
20	Girard	Christiane
21	Bonnet	Claire
22	Dupont	Pierre
23	Lambert	Michel
24	Fontaine	André
25	Rousseau	Philippe
26	Vincent	Louis
27	Muller	René
28	Lefevre	Alain

Information about salespeople including:

Their first and last name

Their salesperson ID number

02

Cleaning

Cleaning the data





Sales dataset

```
array(['Vide', 'Nutri-Pulpe 1000ml - 8.026.380.050.468',  
      'Nutri-Pulpe 500ml - 8.026.380.050.475',  
      'Nutri-Pulpe Pêche Blanche 1000ml - 8.026.380.050.482',  
      'Nutri-Pulpe Pêche Blanche 500ml - 8.026.380.050.499',  
      'Pur Gel Feuille 250ml - 8.026.380.050.505',  
      'Pur Gel Flacon 100ml - 8.026.380.050.512',  
      'Aloe-Oil Huile Réparatrice 50ml - 8.026.380.050.529',  
      "Cœur d'Aloe Cube", 'Totem Double Face',  
      'Pop-Up Cœur D'Aloe Cube Offre Consommateur',  
      'Présentoir de Sol Gamme Aloe',  
      'Grand Aloe Vera Cartonné de Linéaire ou Vitrine',  
      'Brochure Gamme Aloe Vera', 'Echantillon Nutri-Pulpe',  
      'Echantillon Pur Gel',  
      'Echantillon Nutri-Pulpe Goût Pêche Blanche',  
      'Aloe-Oil Crème Réparatrice 150ml - 8.026.380.050.536',  
      'Présentoir de Comptoir Aloe-Oil',  
      'Réglette De Linéaire Aloe Vera',  
      'Présentoir de Comptoir Gamme Aloe', "Cœur d'aloe cube",  
      'Réglette de linéaire Aloe Vera', 'Totem double face',  
      'Grand aloe vera cartonné de linéaire ou vitrine',  
      'Pur Gel Feuille - 250ml - 8.026.380.050.505',  
      'Pur Gel Flacon - 100ml - 8.026.380.050.512',  
      'Aloe-Oil Huile Réparatrice - 50ml - 8.026.380.050.529',  
      'Aloe-Oil Crème Réparatrice - 150ml - 8.026.380.050.536',  
      "Pop-up cœur d'aloe cube offre consommateur",  
      'Pur gel feuille - 250ml', 'Pur gel flacon - 100ml',
```

Some product names contain numbers.

Some special characters interfere with the analysis because products containing them are not recognized properly: it's impossible to group them to perform a general analysis of the product in question.



Sales dataset

Cleaning of special characters

```
#Removing special characters from the products' names and replacing them with the right ones  
sales["article_name"] = sales["article_name"].apply(lambda x : x.replace("Ã©", "é").replace("Ãª", "ê").replace("Ãª", "è"))
```

Removal of numbers from product names

```
#Using a loop to go through the column, find the EAN number, split the string in two and keep the first part  
  
new_values = []  
for i, v in sales["article_name"].items():  
    if "- 8.02" in v:  
        v = v.replace("- 8.02", "removeafterthat")  
        sep = "removeafterthat"  
        v = v.split(sep, 1)[0]  
        new_values.append(v)  
    elif "- 8.02" not in v:  
        v = v  
        new_values.append(v)  
  
sales["article_name"] = new_values
```



Sales dataset

Result of the cleaning

```
'Nutri-pulpe Pêche Blanche 1000ml',  
'Nutri-pulpe Pêche Blanche 500ml', 'Pur Gel Feuille 250ml',  
'Pur Gel Flacon 100ml', 'Aloe-oil Huile Réparatrice 50ml',  
"Cœur D'aloe Cube", 'Totem Double Face',  
'Pop-up Coeur D'aloe Cube Offre Consommateur',  
'Présentoir De Sol Gamme Aloe',  
'Grand Aloe Vera Cartonné De Linéaire Ou Vitrine',  
'Brochure Gamme Aloe Vera', 'Echantillon Nutri-pulpe',  
'Echantillon Pur Gel',  
'Echantillon Nutri-pulpe Goût Pêche Blanche',  
'Aloe-oil Crème Réparatrice 150ml',  
'Présentoir De Comptoir Aloe-oil',  
'Réglette De Linéaire Aloe Vera',  
'Présentoir De Comptoir Gamme Aloe',  
"Pop-up Cœur D'aloe Cube Offre Consommateur",  
'Présentoir De Comptoir Nutri-pulpe & Pur Gel',  
'Présentoir De Comptoir Gamme Aloe-oil',  
'Brochure De Gamme Aloe Vera Consommateur',  
'Présentoir De Sol Aloe-oil',  
'Echantillon Aloe-oil Huile Réparatrice',  
'Echantillon Aloe-oil Crème Réparatrice',  
'Aloe-oil Huile Reparatrice 50ml',  
'Nutri-pulpe Goût Pêche Blanche 500ml',
```



03

Analysis

Excerpt from the requested
sales analysis

Merging the tables to perform the analysis

The CIP code allows to merge the "customers" and "sales" tables

primary_key	order_reference	code_cip_id	salesrep_id	article_code	article_name	quantity	unit_price	discount	net_revenue	order_date	Unnamed: 11	uga	pharmacies_name
19561	2105202234S0001	2260084	26	5046	Nutri-pulpe 1000ml	12	24.9	38%	185.26	2020-05-22	NaN	83SMA	PHARMACY NUM17193
19562	2105202234S0001	2260084	26	5047	Nutri-pulpe 500ml	12	16.9	38%	125.74	2020-05-22	NaN	83SMA	PHARMACY NUM17193
19563	2105202234S0001	2260084	26	5048	Nutri-pulpe Pêche Blanche 1000ml	6	24.9	38%	92.63	2020-05-22	NaN	83SMA	PHARMACY NUM17193
19564	2105202234S0001	2260084	26	5049	Nutri-pulpe Pêche Blanche 500ml	6	16.9	38%	62.87	2020-05-22	NaN	83SMA	PHARMACY NUM17193
19565	2105202234S0001	2260084	26	5050	Pur Gel Feuille 250ml	36	16.9	38%	377.21	2020-05-22	NaN	83SMA	PHARMACY NUM17193
...
10459	2605220959S0003	2051212	29	5046	Nutri-pulpe 1000ml	6	24.9	38%	92.63	2022-05-27	NaN	95PON	PHARMACY NUM19934
10460	2605220959S0003	2051212	29	5047	Nutri-pulpe 500ml	6	16.9	38%	62.87	2022-05-27	NaN	95PON	PHARMACY NUM19934
10461	2605220959S0003	2051212	29	5048	Nutri-pulpe Pêche Blanche 1000ml	6	24.9	38%	92.63	2022-05-27	NaN	95PON	PHARMACY NUM19934
10462	2605220959S0003	2051212	29	5049	Nutri-pulpe Pêche Blanche 500ml	6	16.9	38%	62.87	2022-05-27	NaN	95PON	PHARMACY NUM19934

The merge results in an extra column named "unnamed" which needs to be removed

This merge enables a deeper analysis of customer purchasing habits, the seasonality of these purchases, and an examination of which clients order the most/least and which products they prefer

Merging the tables to perform the analysis

The salesperson ID allows merging the "sales" and "salespeople" tables

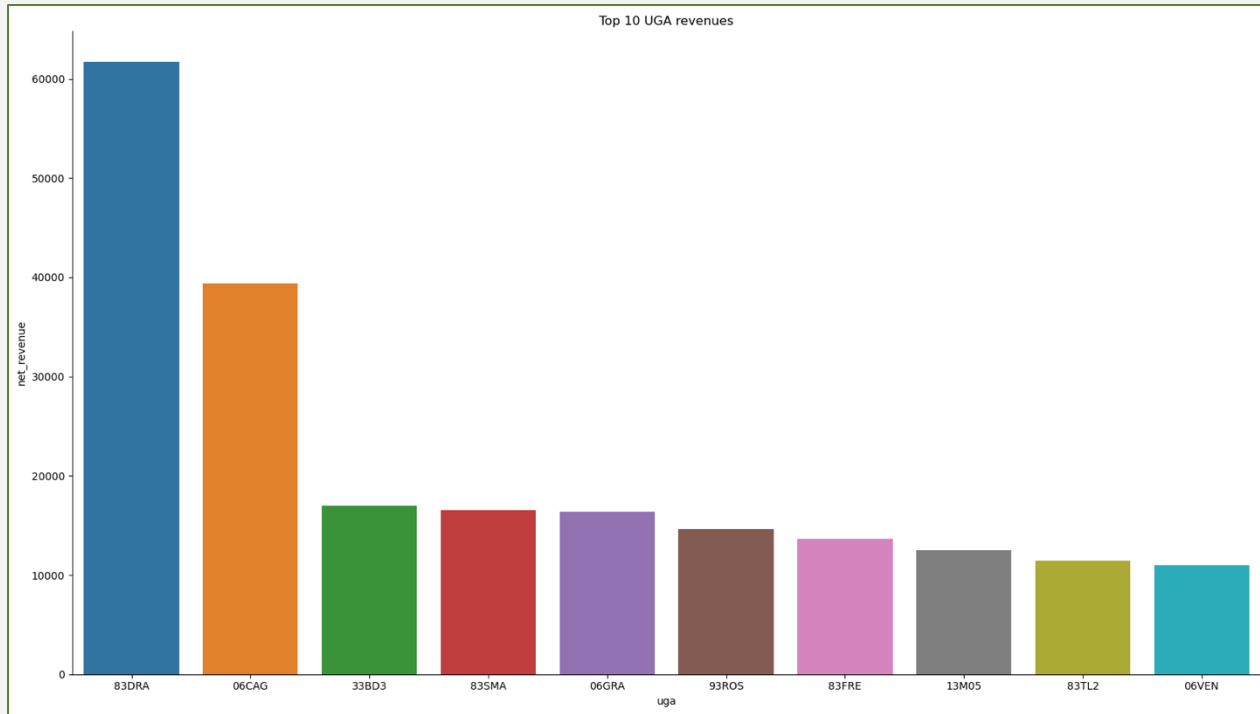
```
#Left join because some sales reps have not sold any products yet, need to identify them
salesreps_sales = salesreps.merge(sales, how = "left")
```

salesreps_sales													
	salesrep_id	last_name	first_name	primary_key	order_reference	code_cip_id	article_code	article_name	quantity	unit_price	discount	net_revenue	order_date
0	1	Martin	Marie	16110.0	1707211740S0001	2249689	5050	Pur Gel Feuille 250ml	24.0	16.9	55%	182.52	2021-07-19
1	1	Martin	Marie	16111.0	1707211740S0001	2249689	5051	Pur Gel Flacon 100ml	24.0	9.9	55%	106.92	2021-07-19
2	1	Martin	Marie	16115.0	1707211651S0001	2003434	5051	Pur Gel Flacon 100ml	12.0	9.9	45%	65.34	2021-07-19
3	1	Martin	Marie	16121.0	1707211651S0001	2003434	5597	Pure Crème 400ml	12.0	14.9	45%	98.34	2021-07-19
4	1	Martin	Marie	16122.0	1707211651S0001	2003434	5598	Gyn-aloe Soin Lavant Pompe 400ml	6.0	16.1	45%	53.13	2021-07-19
...
6944	54	Dumont	Salomé	2037.0	2305221503S0019	2265606	5050	Pur Gel Feuille 250ml	24.0	16.9	55%	182.52	2022-05-24
6945	54	Dumont	Salomé	2038.0	2305221503S0019	2265606	5051	Pur Gel Flacon 100ml	24.0	9.9	55%	106.92	2022-05-24
6946	54	Dumont	Salomé	8508.0	2305221503S0019	2265606	5050	Pur Gel Feuille 250ml	24.0	16.9	55%	182.52	2022-05-24
6947	54	Dumont	Salomé	8509.0	2305221503S0019	2265606	5051	Pur Gel Flacon 100ml	24.0	9.9	55%	106.92	2022-05-24
6948	55	Marie	Jean-Guy	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

This merge enables a detailed analysis of the top-performing salespeople, as well as the products they sell the most

Excerpt of the results

Sales by UGA (Geographical Areas)



Excerpt of the results

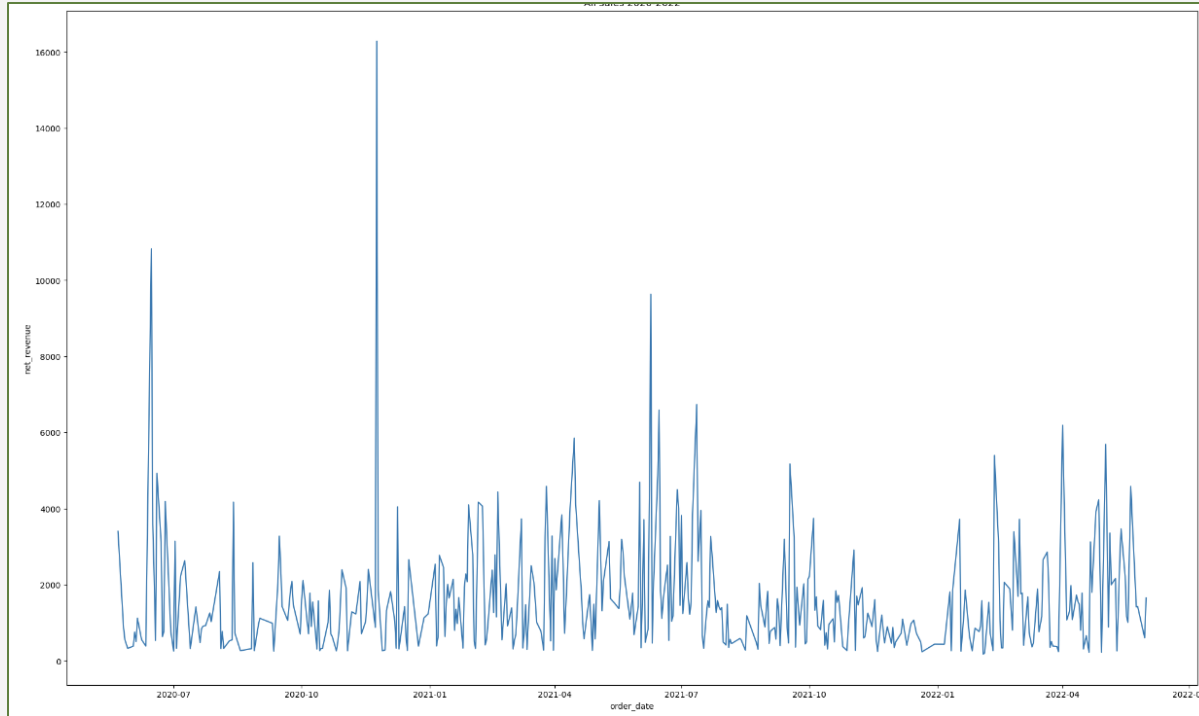
Net revenues by products

article_name	net_revenue
Pur Gel Feuille 250ml	184426.3700
Pur Gel Flacon 100ml	98695.2700
Aloe-oil Crème Réparatrice 150ml	64822.4400
Aloe-oil Huile Réparatrice 50ml	59785.2825
Nutri-pulpe 1000ml	53145.9600
Nutri-pulpe 500ml	52402.1700
Nutri-pulpe Pêche Blanche 500ml	47444.5100
Nutri-pulpe Pêche Blanche 1000ml	31723.4500
Pure Crème 400ml	27015.6600
Gyn-aloe Soin Lavant Flacon 250ml	18337.2500

All products sales over time

Excerpt of the results

Sales from May 2020 to July 2022



Excerpt of the results

Best salespeople

	net_revenue
first_name	
Louis	148182.61
Marie	119901.29
Olivia	52391.69
Pauline	49776.63
Bernard	42438.30
Jacques	35591.50
René	29187.53
Lara	28891.19
Marcel	25453.50
Georges	20730.92

Best customers

		net_revenue
code_cip_id	pharmacies_name	
2102240	PHARMACY NUM5440	16986.24
2241522	PHARMACY NUM986	13073.40
2048765	PHARMACY NUM19194	13037.88
2007231	PHARMACY NUM1764	10593.30
2042674	PHARMACY NUM17261	9475.77
2235131	PHARMACY NUM19862	7107.75
2042412	PHARMACY NUM17325	6884.95
2012778	PHARMACY NUM1003	6788.37
2242919	PHARMACY NUM17204	6706.52
2012758	PHARMACY NUM861	6343.77

