

better⁴eat

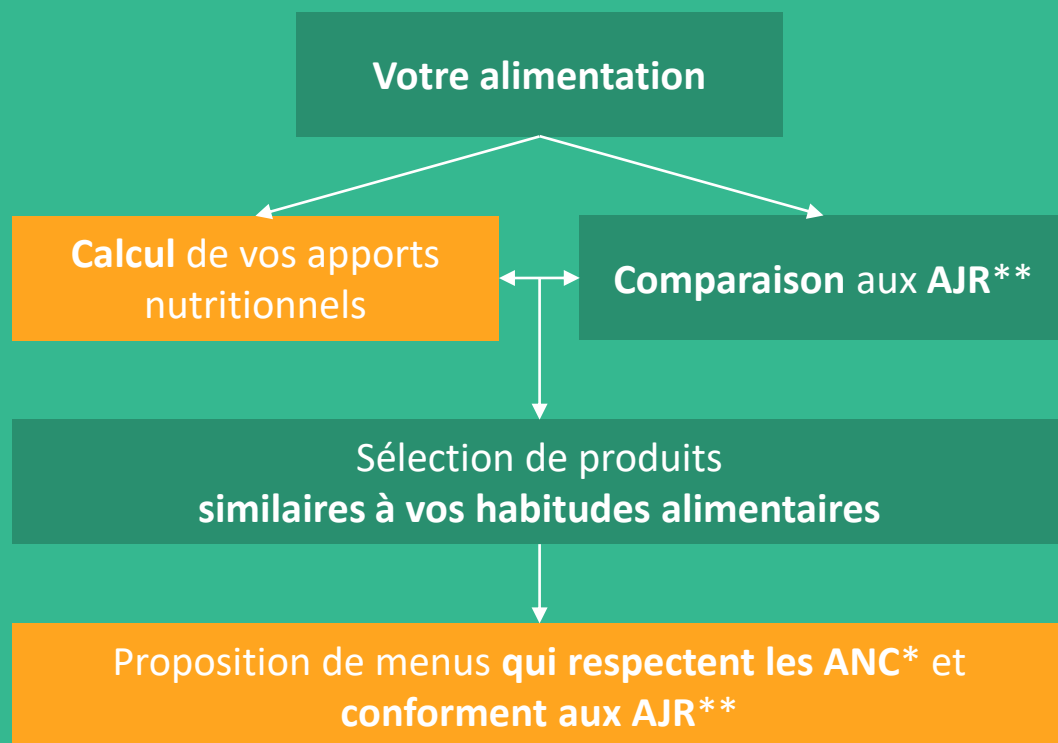
Réponse à l'APP du 22-05-22

Trouver des idées innovantes d'applications en lien avec l'alimentation



Vos repas

en équilibré



Jeu de données



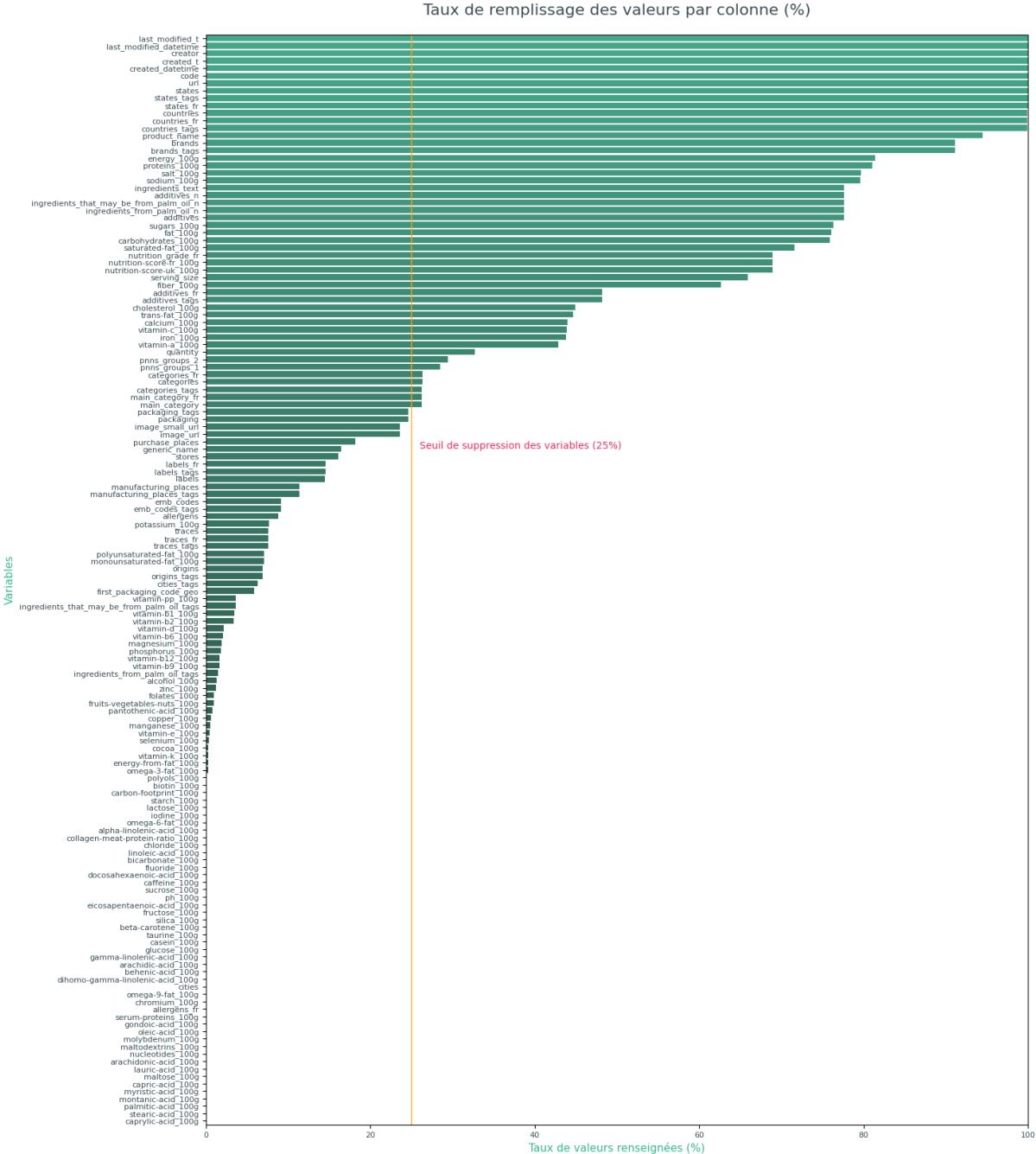
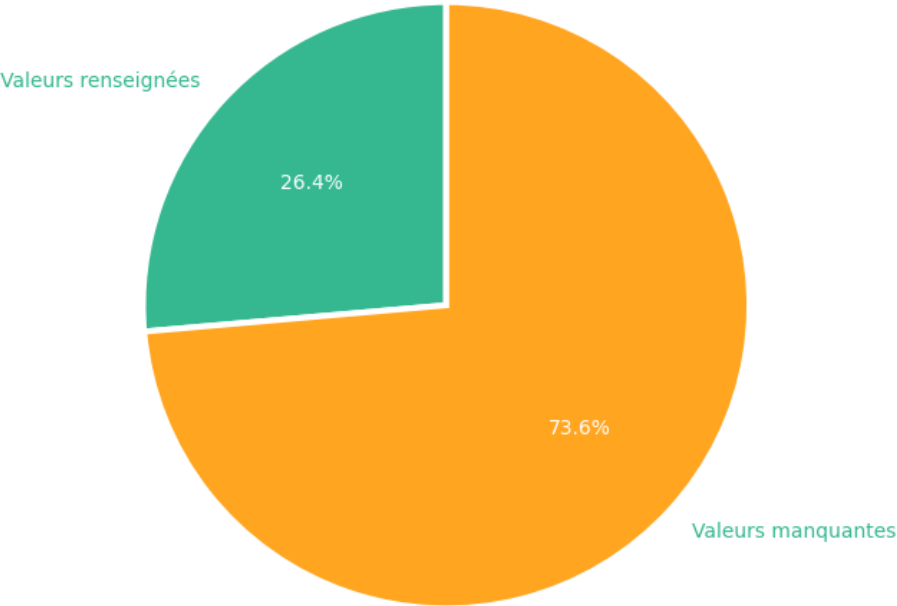
Identifiants		Marques	Constituants alimentaires			Nutri-Score	Catégories		Informations							
code	product_name	brands	allergens	additives_fr	ingredients_from_palm_oil_n	nutrition_grade_fr	pnnsgroups_1	pnnsgroups_2	energy_100g	fat_100g	saturated-fat_100g	carbohydrates_100g	sugars_100g	fiber_100g	proteins_100g	salt_100g
3350033193635	Tarte Épinard Chèvre	Monoprix	1	1	0	D	Composite foods	Pizza pies and quiche	994.0	13.8	9.480	22.4	2.07	1.540	6.03	0.862000
3292070003594	Mezze Méditerranéen	Blini	0	1	0	D	Fat and sauces	Dressings and sauces	1523.0	31.5	5.500	6.7	1.30	1.020	6.40	1.100000
8023678162315	Riznoiset	isola Bio	0	0	0	E	Beverages	Non-sugared beverages	333.0	2.7	0.400	12.5	5.50	0.900	0.80	0.080000
5413415233300	Chocolat Noir Poivre Rose	Dolfin	1	1	0	E	Sugary snacks	Chocolate products	2230.0	35.0	21.000	43.0	38.00	10.000	7.00	0.030480
3262971311258	Lait fermenté	--	0	0	0	A	Milk and dairy products	Milk and yogurt	180.0	1.1	0.800	4.9	4.90	0.320	3.40	0.120000
3760132081991	Concombre Avocat Poire	Green Smoothie ,Green Shoot	0	0	0	D	Beverages	Non-sugared beverages	139.0	0.9	0.100	4.9	4.40	3.660	0.50	0.030000
3263851740717	Gratin Dauphinois	Leader Price	1	1	0	B	Composite foods	One-dish meals	275.0	1.9	1.300	9.8	0.70	1.800	1.40	0.750000

Nettoyage



Les données brutes

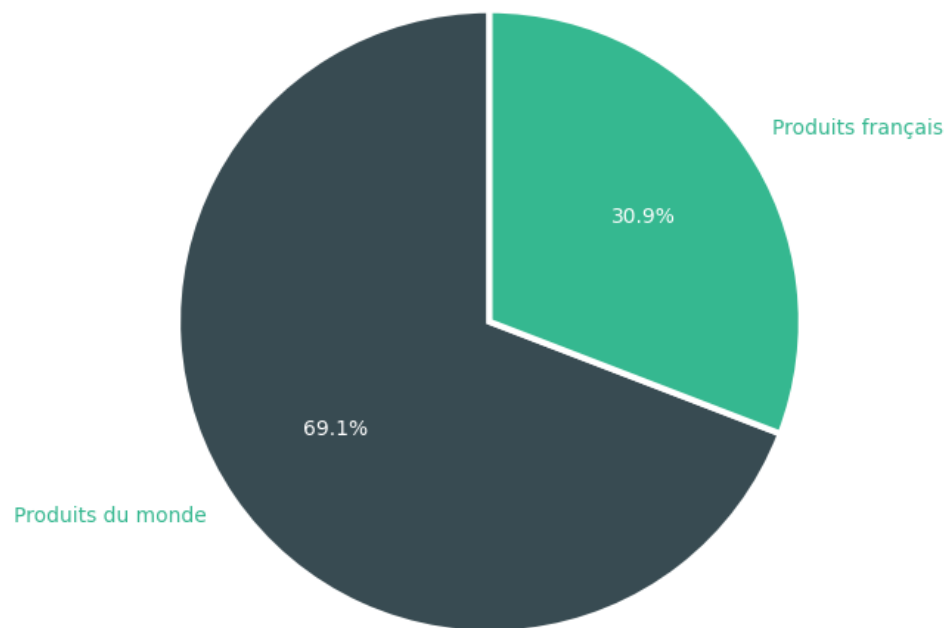
Proportion de données manquantes avant traitement
description: (320772, 146)



Des produits issus

de nos territoires

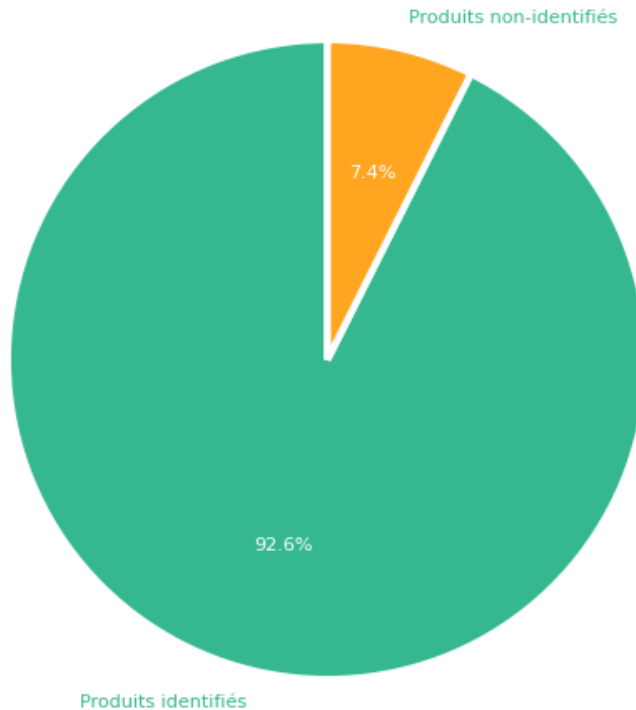
Proportion de produits issus des territoires Français



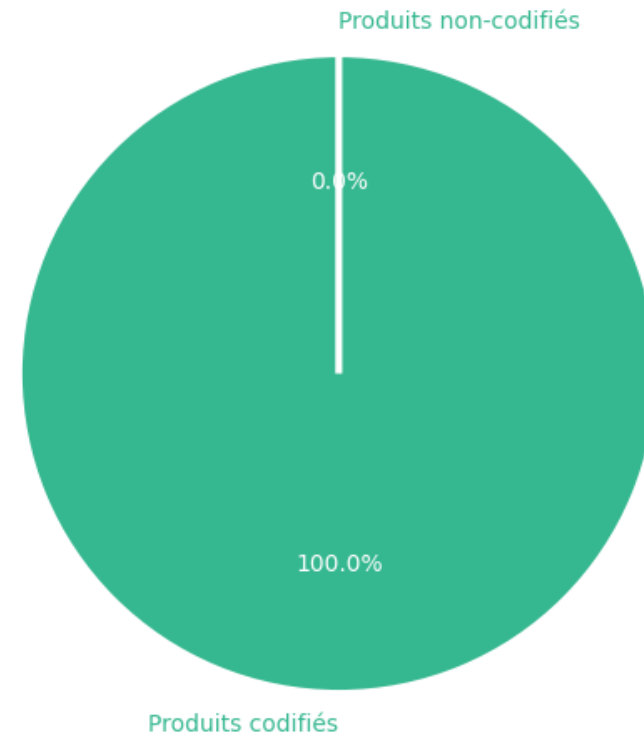
Le référencement produits

Proportion de valeurs manquantes
(identifiant produit)

Product_name

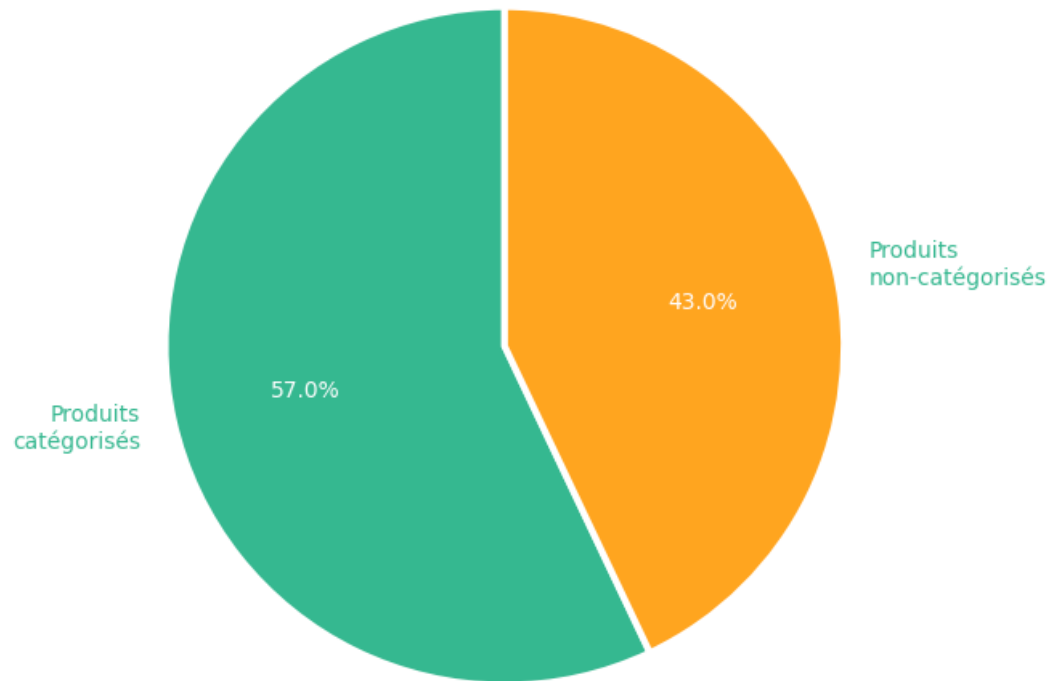


code

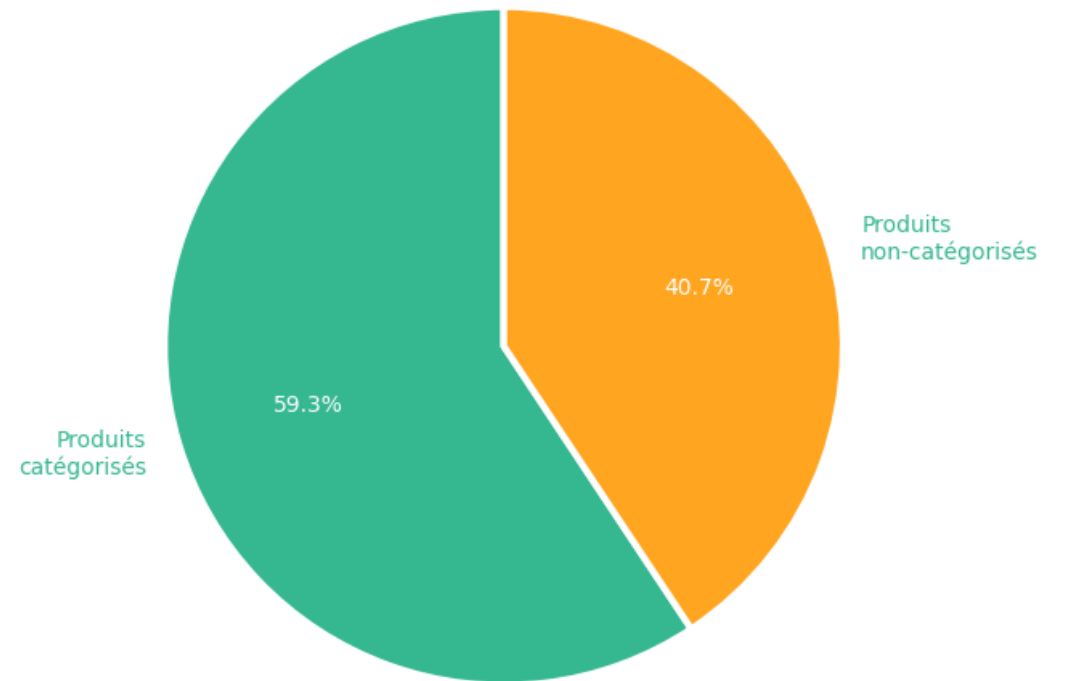


Les catégories de produits

Proportion de produits catégorisés
(pnns_groups_1)



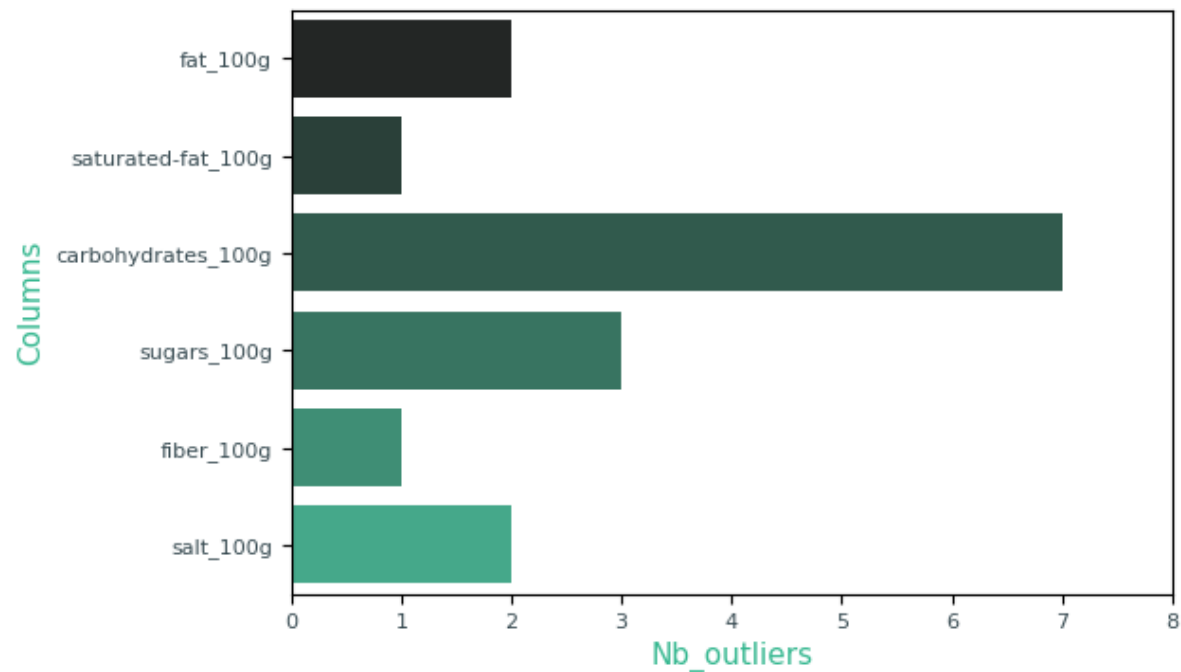
Proportion de produits catégorisés
(pnns_groups_2)



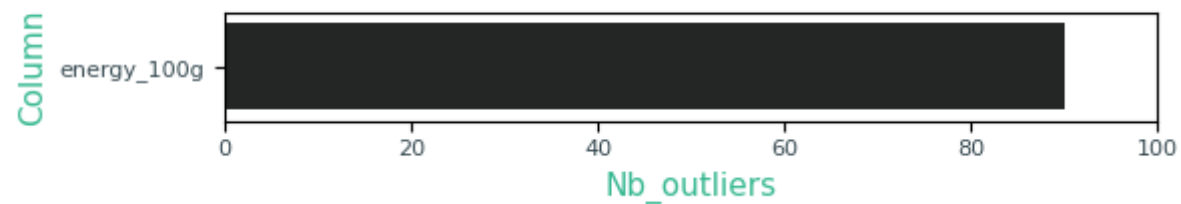
Aberrant !



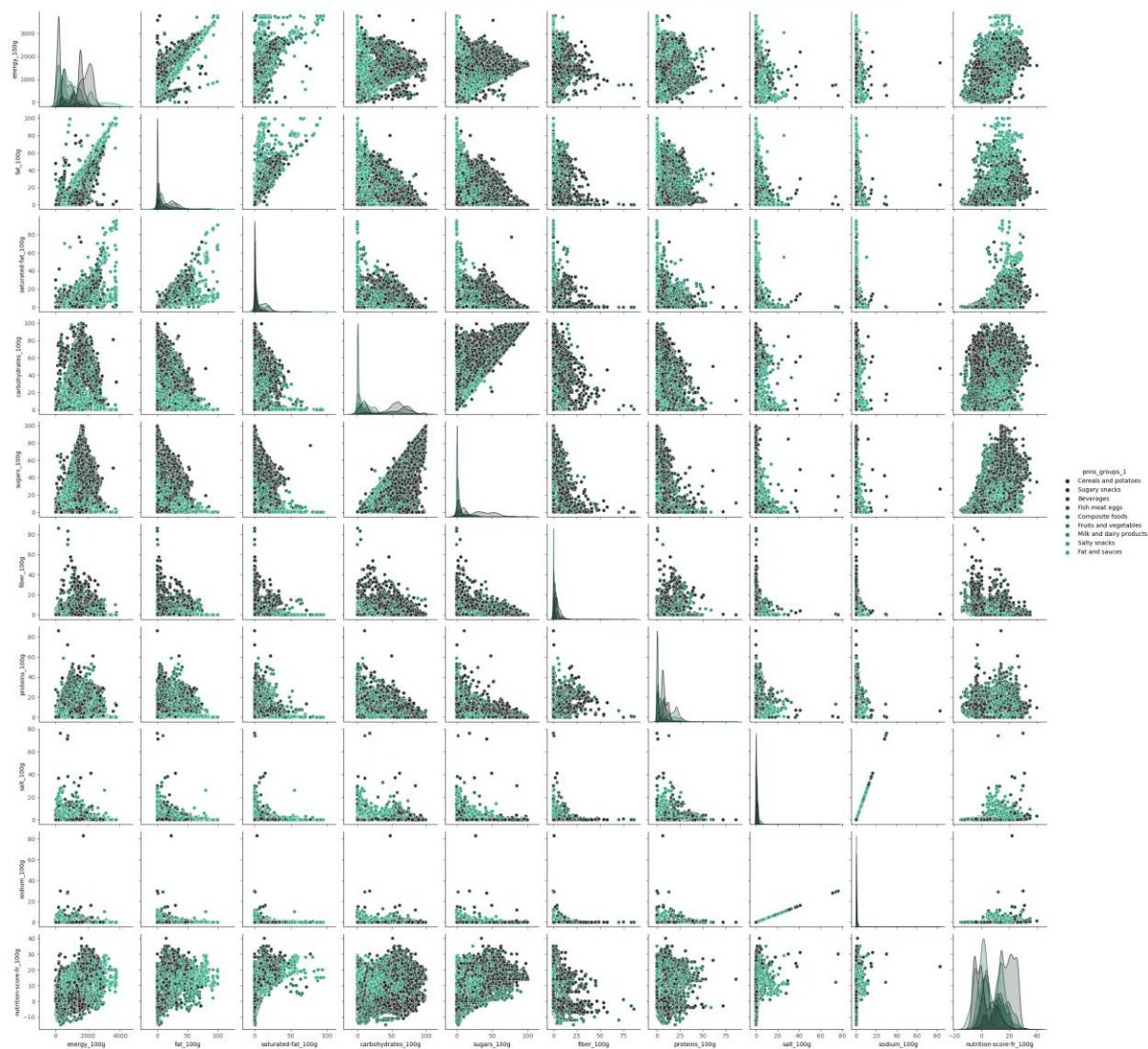
Nombre de valeurs aberrantes



Nombre de valeurs aberrantes



Pairplot des variables nutritionnelles - Groupes PNNS 1



Les relations

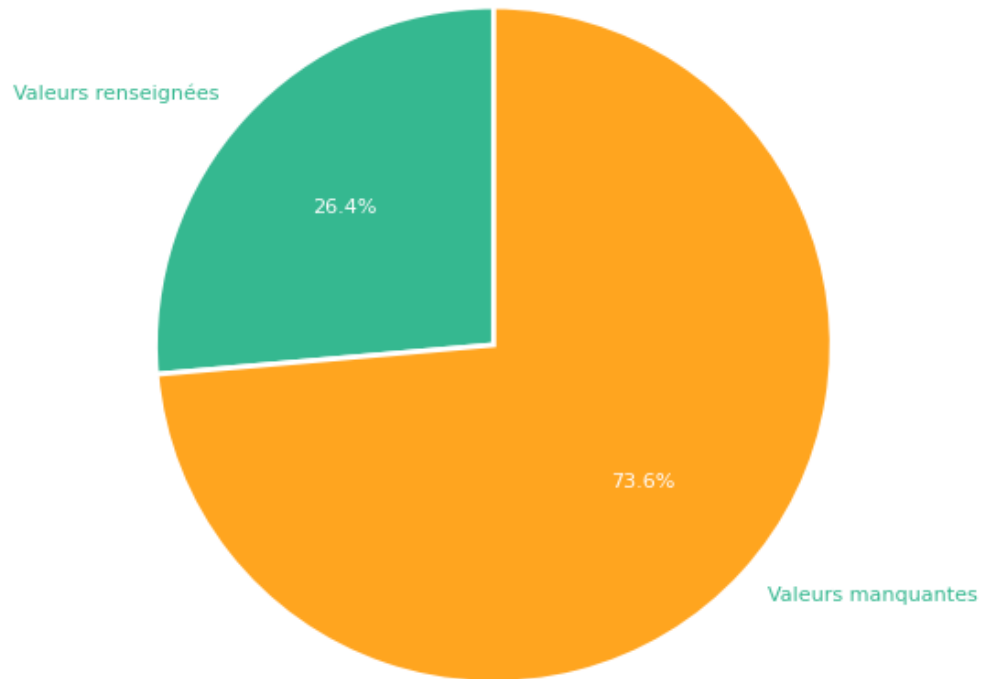
- Sel > Sodium
- Graisses > Graisses saturées
- Hydrates de carbone > Sucre



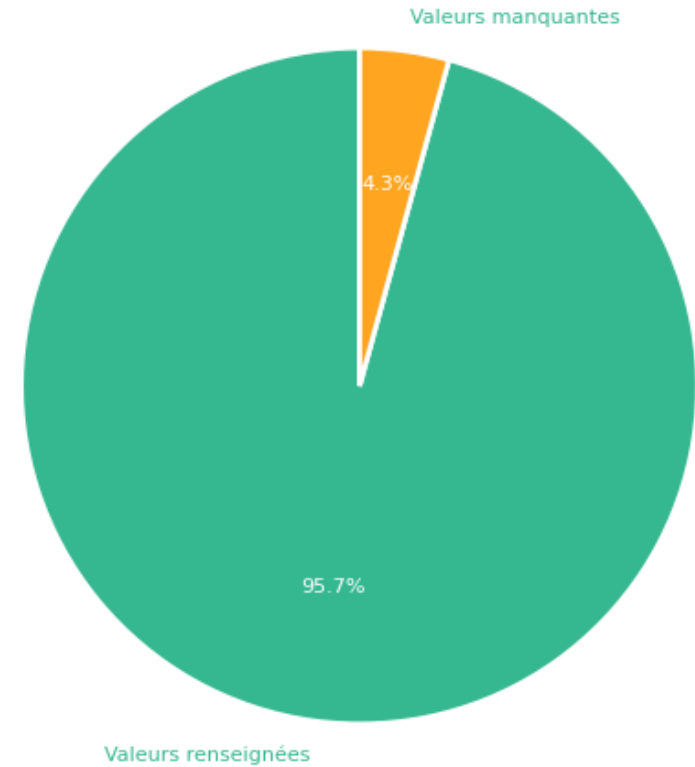
Résultat du nettoyage

Comparaison de la proportion de données manquantes

PRE-traitement (320772, 146)



POST-traitement (42451, 18)



Imputation

better⁴eat

Comparaison

Avant
imputation

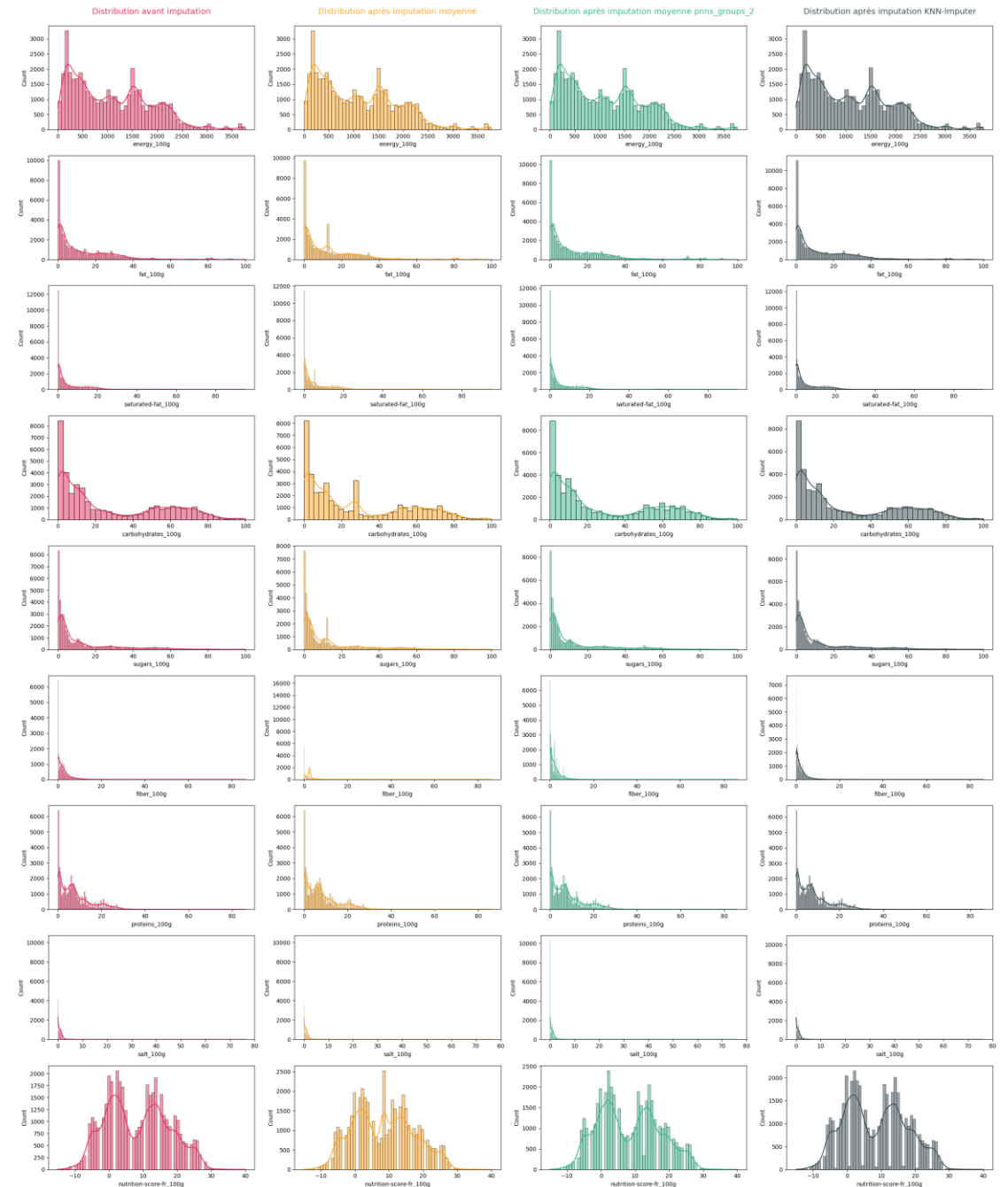
Imputation de la **moyenne**

Imputation de la **moyenne**
pnnn_groups_2

Imputation **KNN-Imputer**

Imputation via KNN-Imputer
Toutes les données sont renseignées

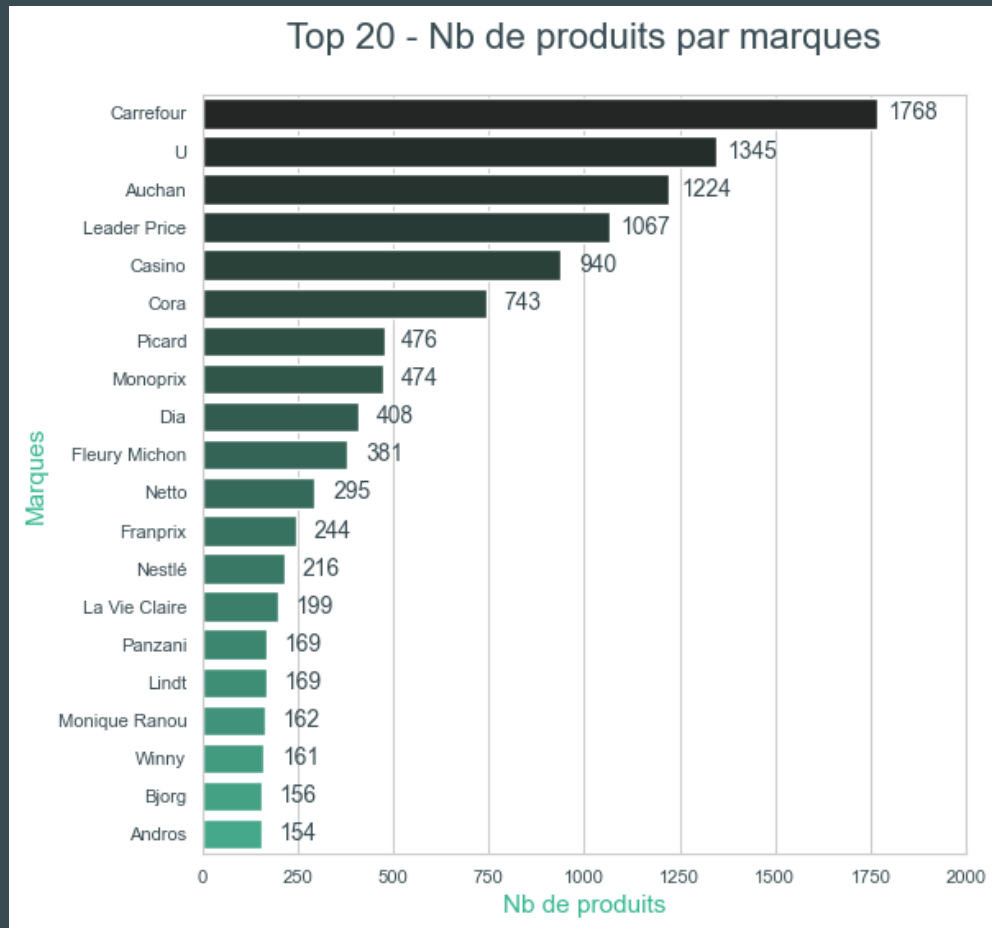
Comparaison de la distribution avant et après imputation



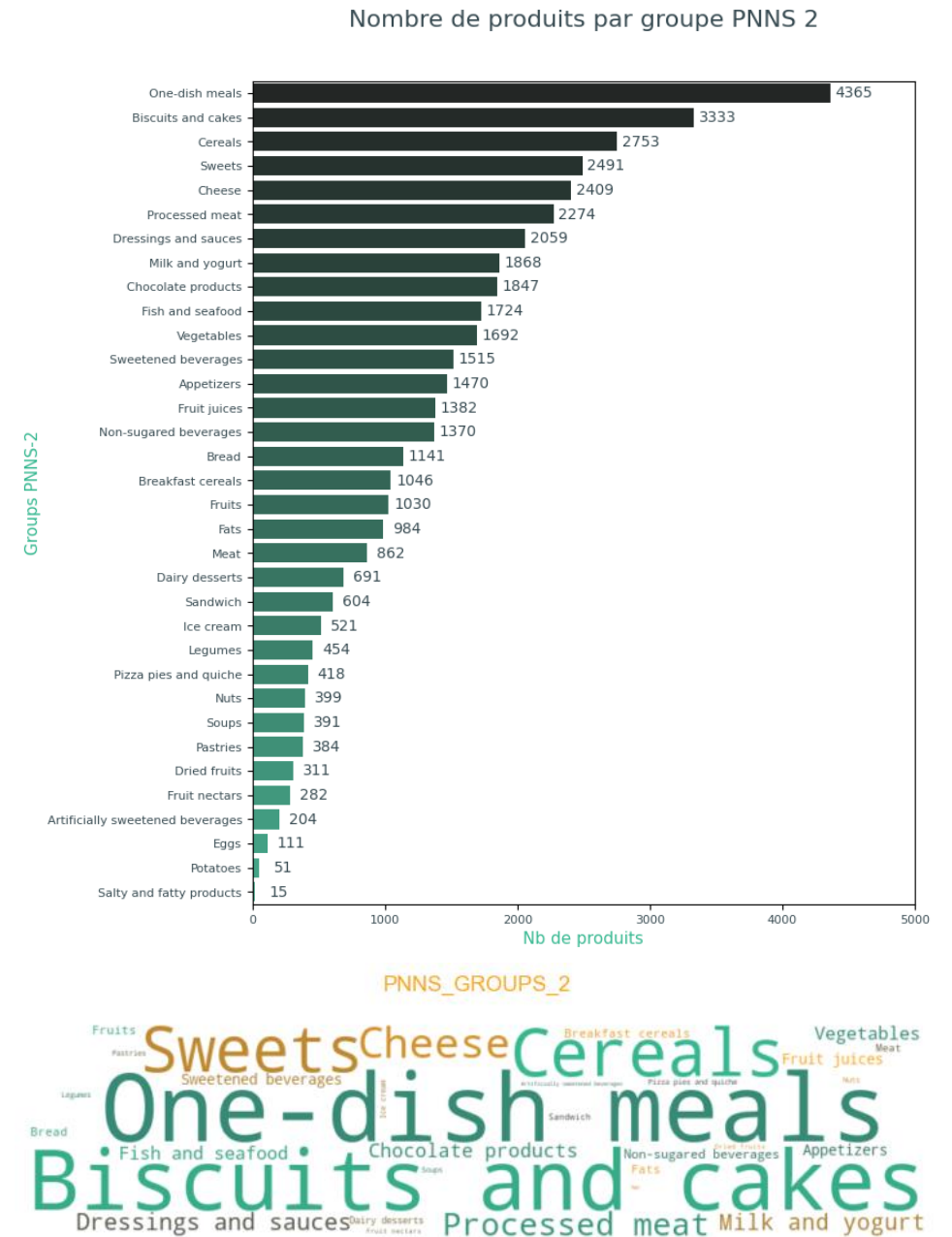
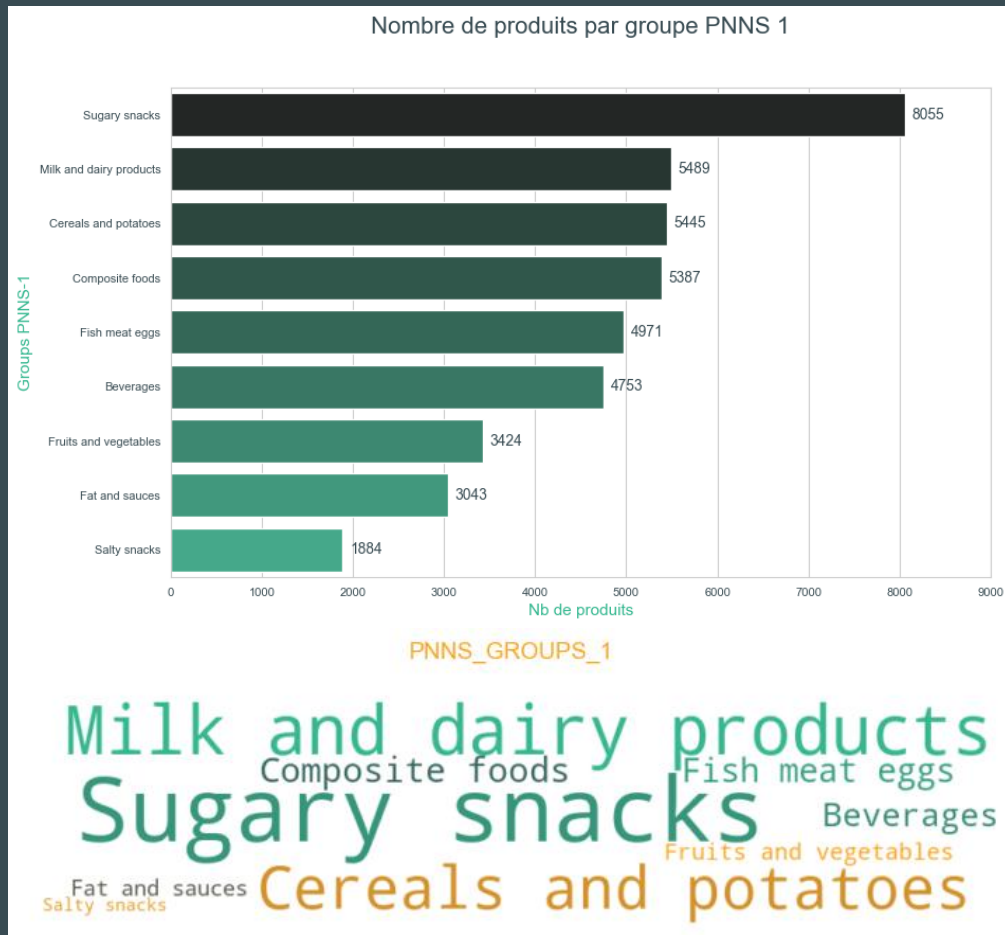
Analyse descriptive

better4eat

Top 20 – Marques

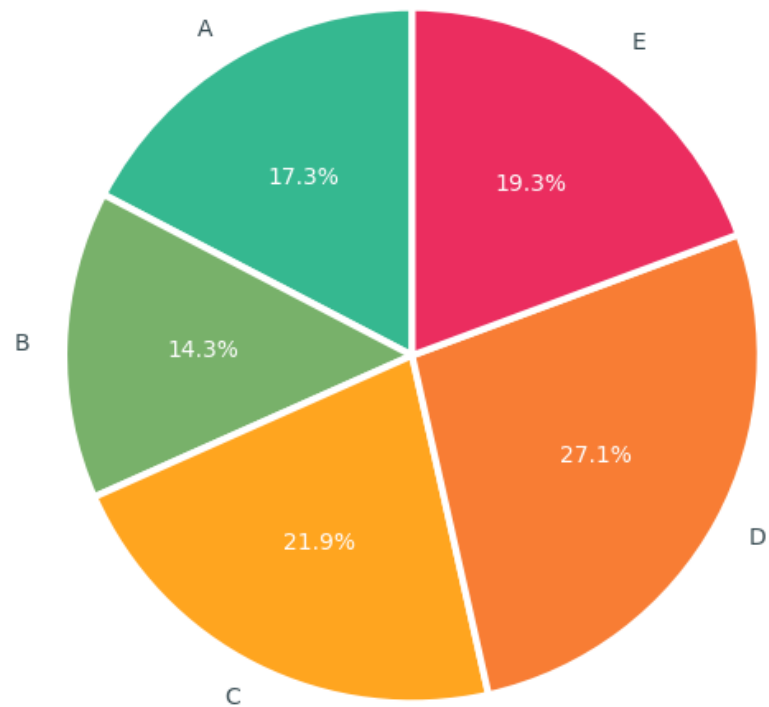


Les catégories



Nutri-Score

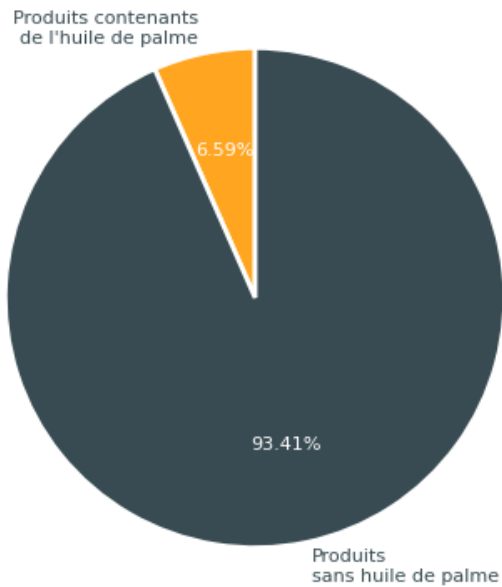
Proportion de produits par qualité nutritionnelle



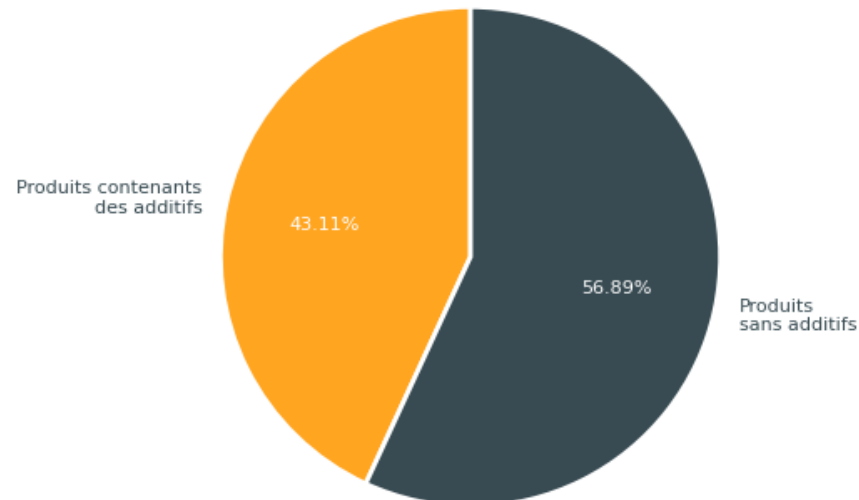
Autres constituants alimentaires

Proportion de produits contenant

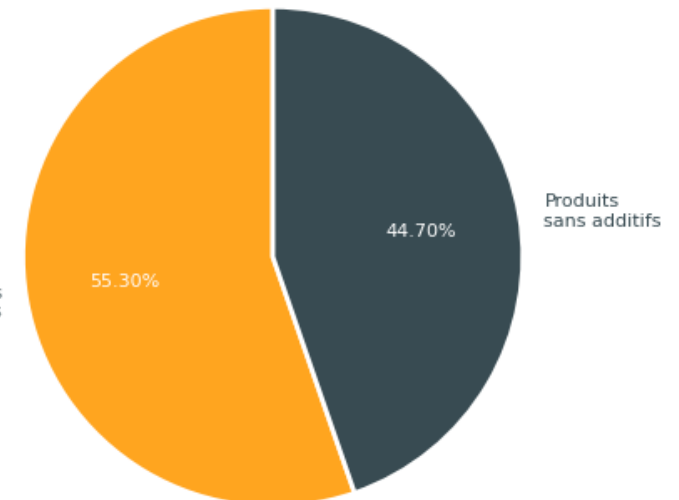
de l'huile de palme



des allergènes



des additifs

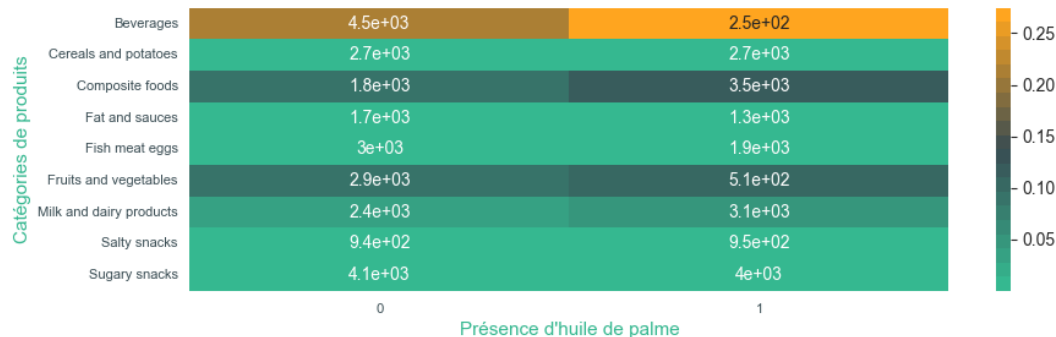


Les associations

better4eat

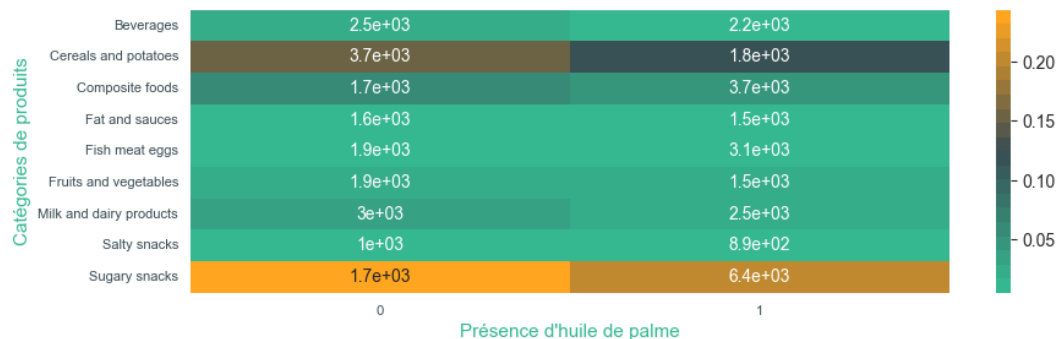
p-value: 0.0
Khi2: 5758.36

Tableau de contingence
PNNS_1/Allergens



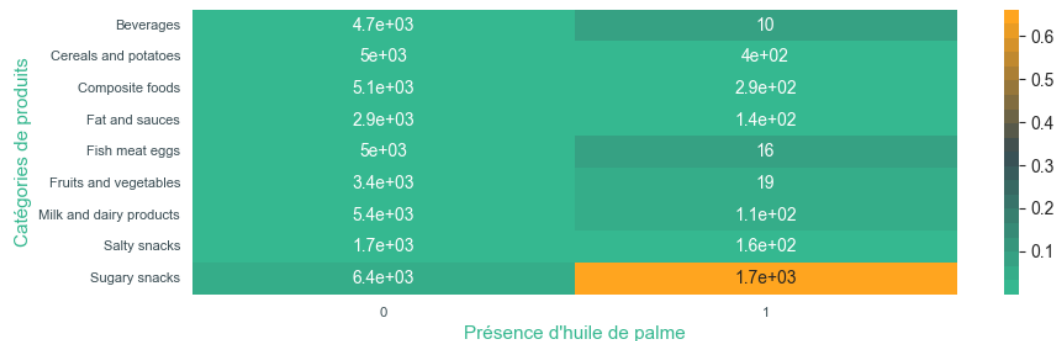
p-value: 0.0
Khi2: 4101.77

Tableau de contingence
PNNS-1/Additives_fr)

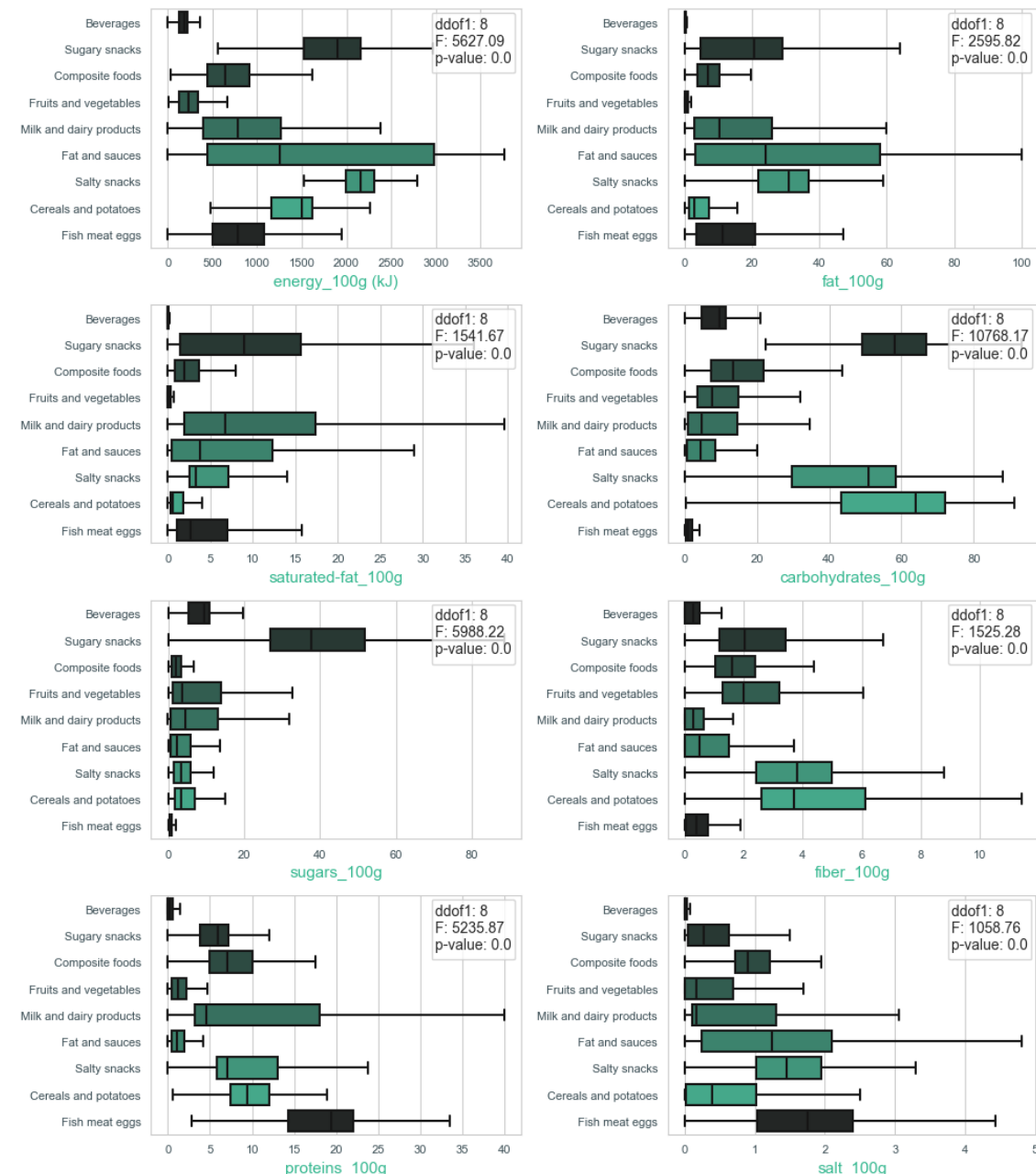


p-value: 0.0
Khi2: 3663.17

Tableau de contingence
PNNS-1/Huile de palme



Répartition de la teneur en nutriment par groupe de produits



p-value: 0.0
Khi2: 381.01



p-value: 0.0
Khi2: 2545.52



p-value: 0.0
Khi2: 2045.03



Figure 2 displays 10 horizontal box plots comparing the intake of various nutrients across five groups (A, B, C, D, E). The nutrients are arranged in two columns: energy_100g (kJ), fat_100g, saturated_fat_100g, carbohydrates_100g, sugars_100g, fiber_100g, proteins_100g, and salt_100g. Each plot shows the distribution of intake for each group, with the y-axis labeled E, D, A, B, and C from top to bottom. The x-axis represents the nutrient intake. Each plot includes a box plot with a median line, a box representing the interquartile range, and whiskers extending to the minimum and maximum values. Statistical results (ddof1, F, and p-value) are provided for each plot.

Nutrient	ddof1	F	p-value
energy_100g (kJ)	4	4100.99	0.0
fat_100g	4	3985.95	0.0
saturated_fat_100g	4	4394.21	0.0
carbohydrates_100g	4	521.31	0.0
sugars_100g	4	2398.63	0.0
fiber_100g	4	555.94	0.0
proteins_100g	4	290.68	0.0
salt_100g	4	464.24	0.0

Les corrélations

better4eat

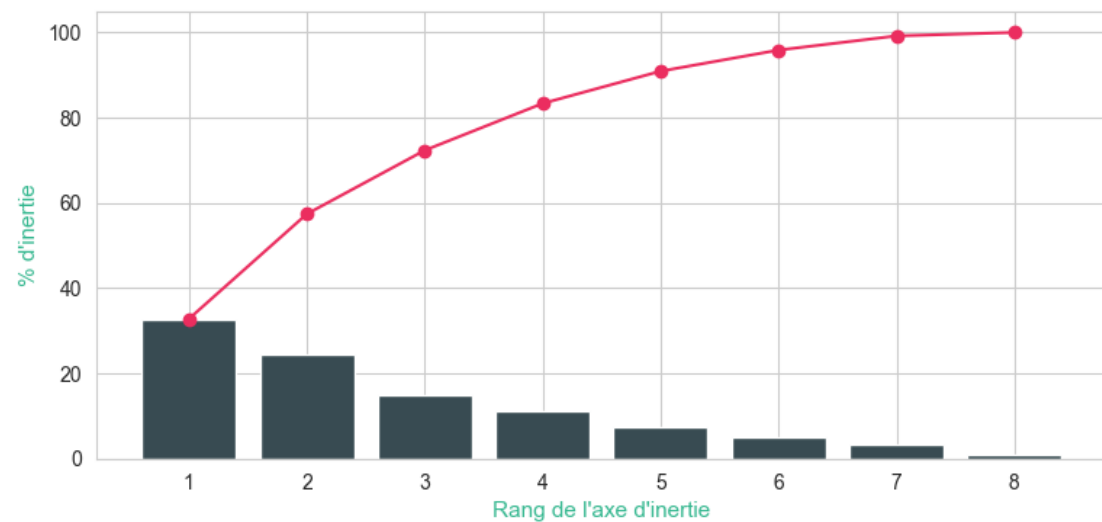
Analyse de corrélation entre les indicateurs - r de Spearman



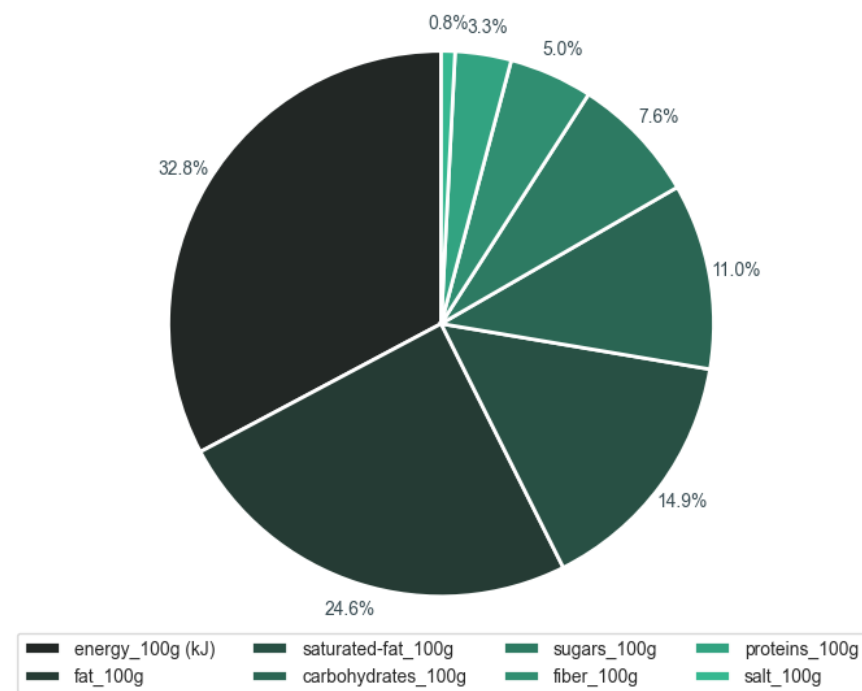
ACP

better4eat

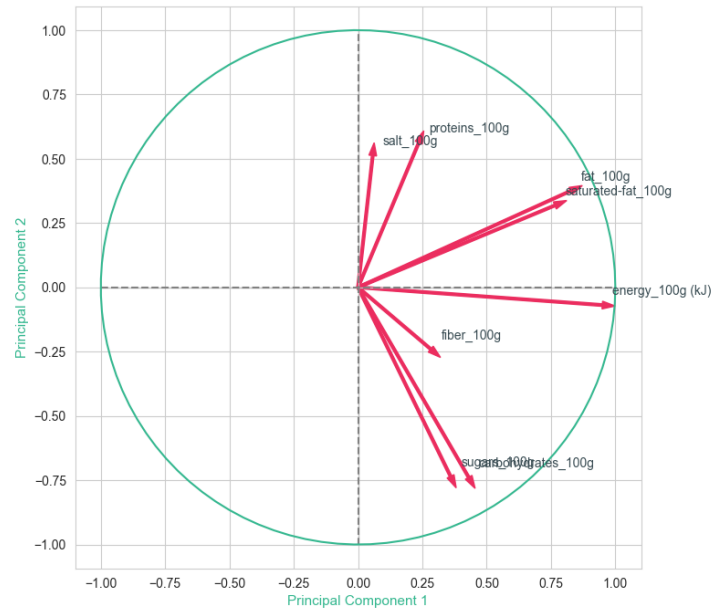
Eboulis des valeurs propres



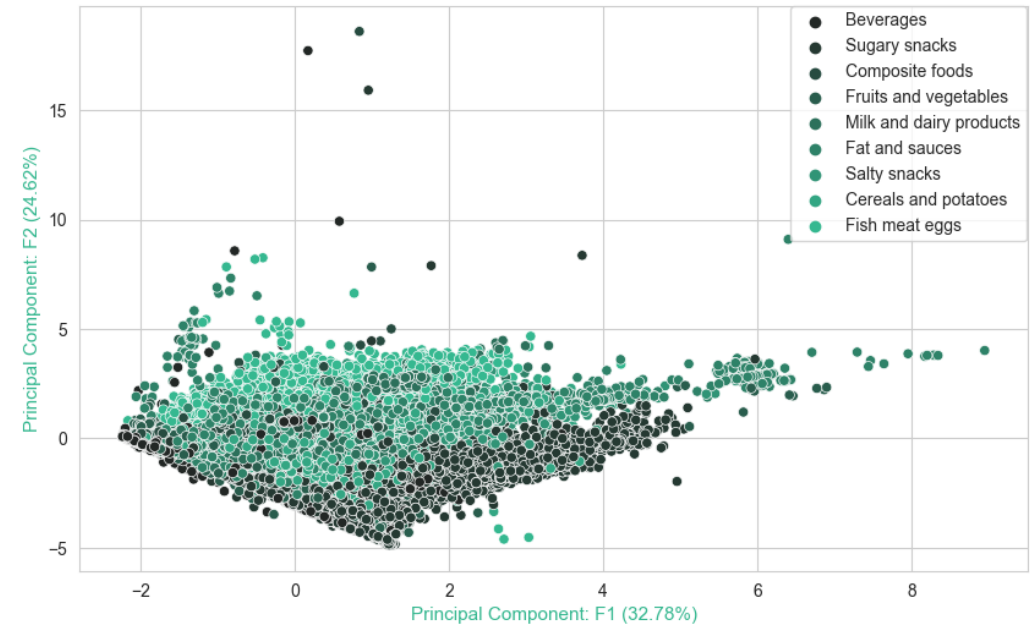
Répartition des composantes principales



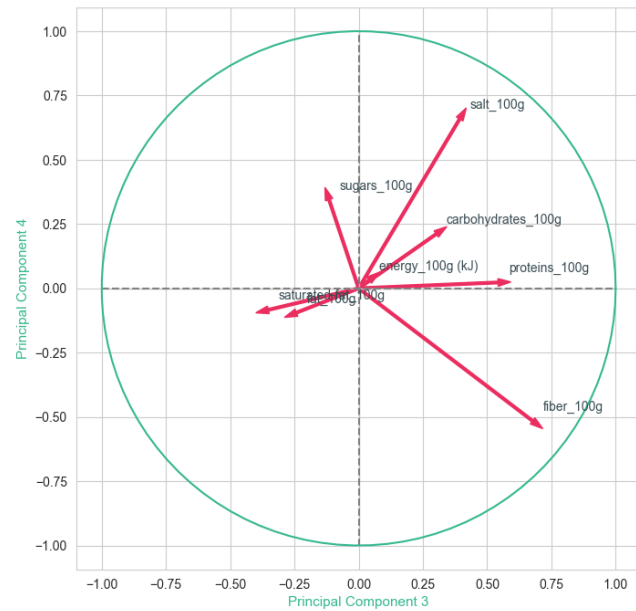
Cercle des corrélations



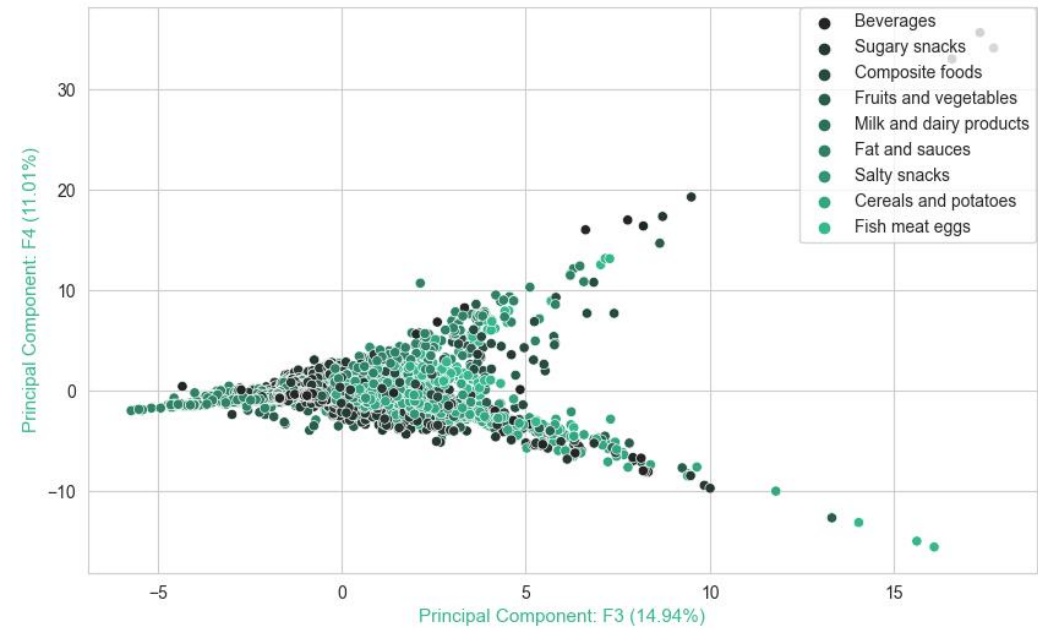
1er Plan factoriel (F1,F2): 57.39%



Cercle des corrélations



2nd Plan factoriel (F3,F4): 25.94%



Conclusion

better4eat