

# PROJET APPLICATION

Projet application pour une alimentation plus saine

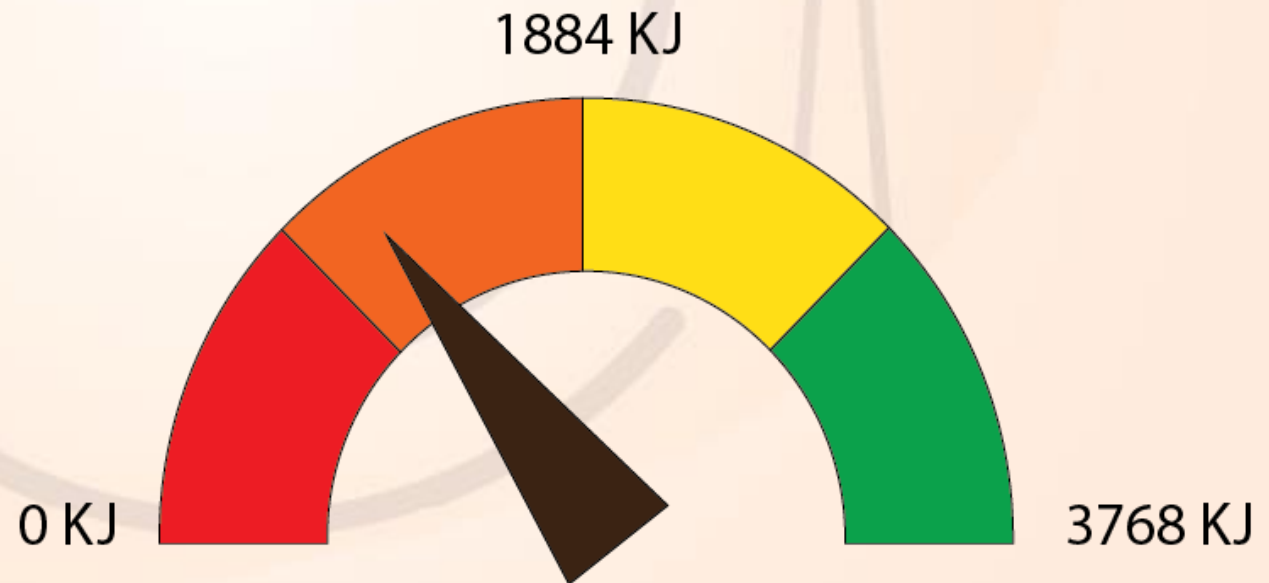
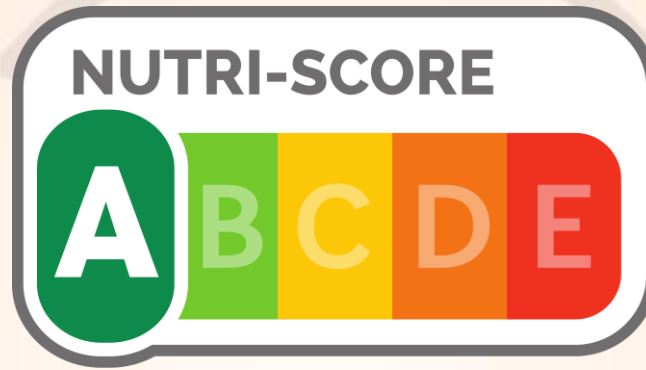
# RAPPEL PROBLÉMATIQUE



**Santé Publique France lance un appel à projets pour trouver des idées innovantes en lien avec l'alimentation**

# IDÉE APPLICATION

- Compteur de Joules
- Fixer les limites des macronutriments à limiter
- Favoriser les produits avec bon nutri-score et une bonne classification NOVA



# Rappel recommandations

Les éléments à éviter:

- Sel
- Sucre
- Graisses saturées
- Énergie

Les éléments à favoriser:

- Fibres
- Protéines
- Légumes – Fruits et Noix

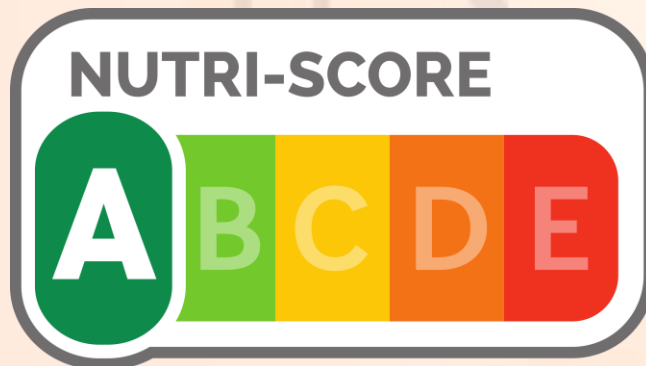
Les éléments à éviter:

Les éléments ultra-transformés

Les éléments à favoriser:

Les éléments avec le moins de transformation

NOVA



# NETTOYAGE

# Présentation du jeu de données

Contient des données sur  
les produits  
Lignes: 253678  
Colonnes: 197

## Variables Classification:

- Code
- Nom du produit
- Nutri-Score
- Score nutrition
- Classification NOVA
- Ingrédients
- URL

## Variables Macronutriments pour 100g:

- Energie en kcal
- Lipides
- Graisses saturées
- Acides gras
- Glucides
- Sucres
- Fibres
- Protéines
- Fruits/Légumes et Noix
- Sel
- Sodium
- Alcool

# Suppressions

## Le pré-traitement:

- Produits sans code
- Produits sans nom
- Produits sans données présentes dans les variables : Lipides, Glucides et Protéines

## Les cas particuliers:

- Produits dont les sucres sont supérieurs aux glucides
- Produits dont les acides gras et les graisses saturées sont supérieurs aux Lipides
- Suppression des produits dont la somme des Lipides, Glucides et Protéines est supérieur à 100g

## Les traitements des Outliers:

- Passage des valeurs négatives en valeurs absolues
- Lignes dont des macronutriments sont supérieurs à 100g

# Imputations

## Imputation par 0

- Les additifs
- Les Fruits/Légumes et Noix
- L'alcool
- Les fibres

## Imputation par le calcul

Calcul pour 100g:

$$\text{Energy(kcal)} = 9 * \text{Lipides} + 4 * \text{Glucides} + 4 * \text{Protéines}$$

$$\text{Energy(kJ)} = 4,1868 * \text{Energy(kcal)}$$

L'energie ne peut pas dépasser les 3768 kJ



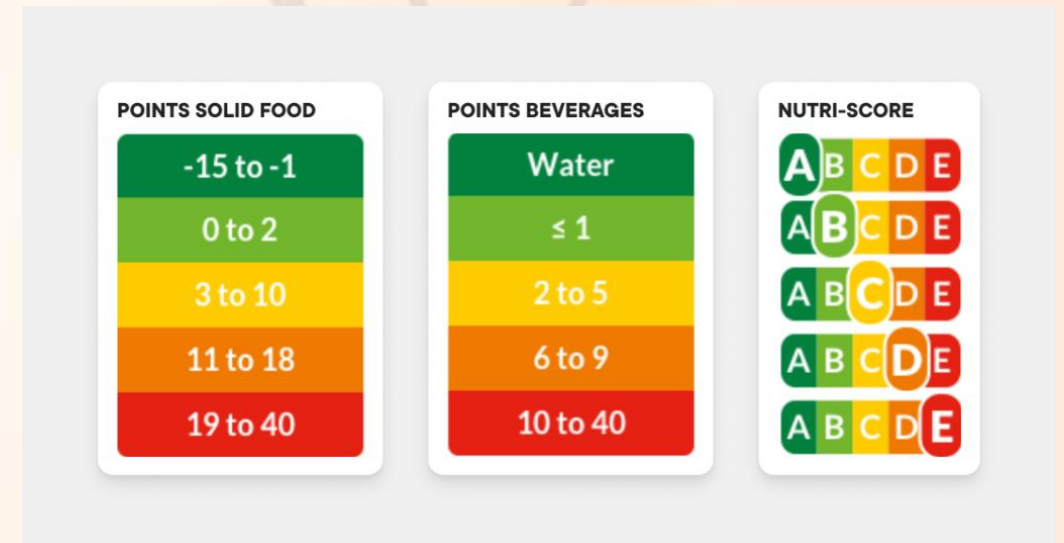
## Imputation via le KNN Imputer


- Score Nutrition
- Energie
- Lipides
- Protéines
- Glucides
- Graisses saturées
- Acides gras
- Sel
- Sodium
- Sucres

## Imputation via l'Iterative Imputer

- Classification NOVA puis conversion en entier

## Imputation par attribution



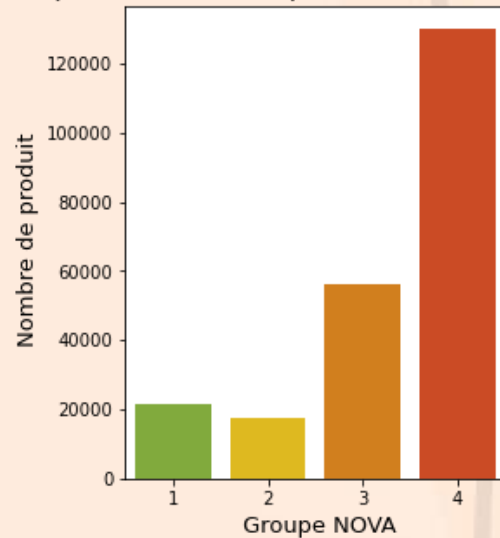
A faint, stylized illustration of a white plate with a fork and a spoon, centered on a light orange background. The plate and cutlery are rendered in a light grey or beige tone, creating a subtle watermark effect.

# **ANALYSE EXPLORATOIRE**

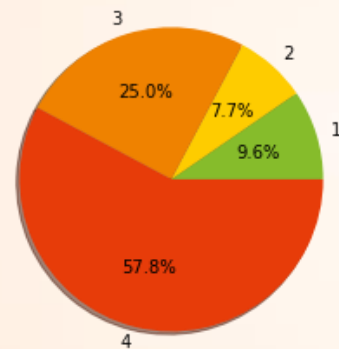
# ANALYSE UNIVARIÉE

# Variables qualitatives

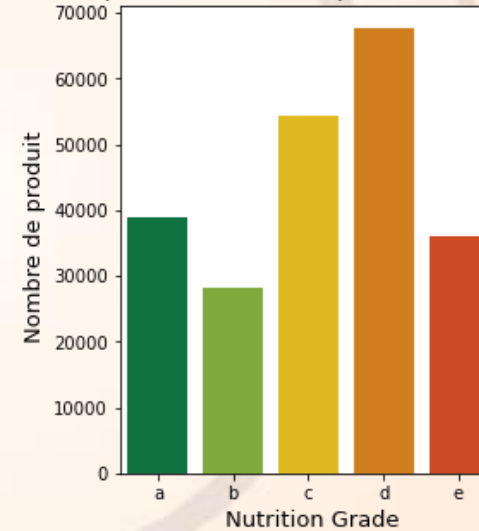
Répartition Produit par classification NOVA



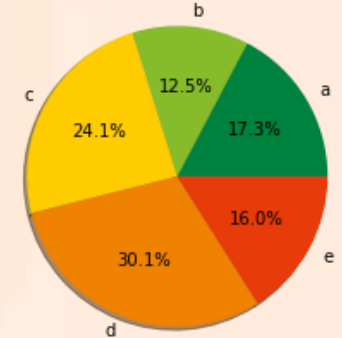
Répartition Produit par classification NOVA



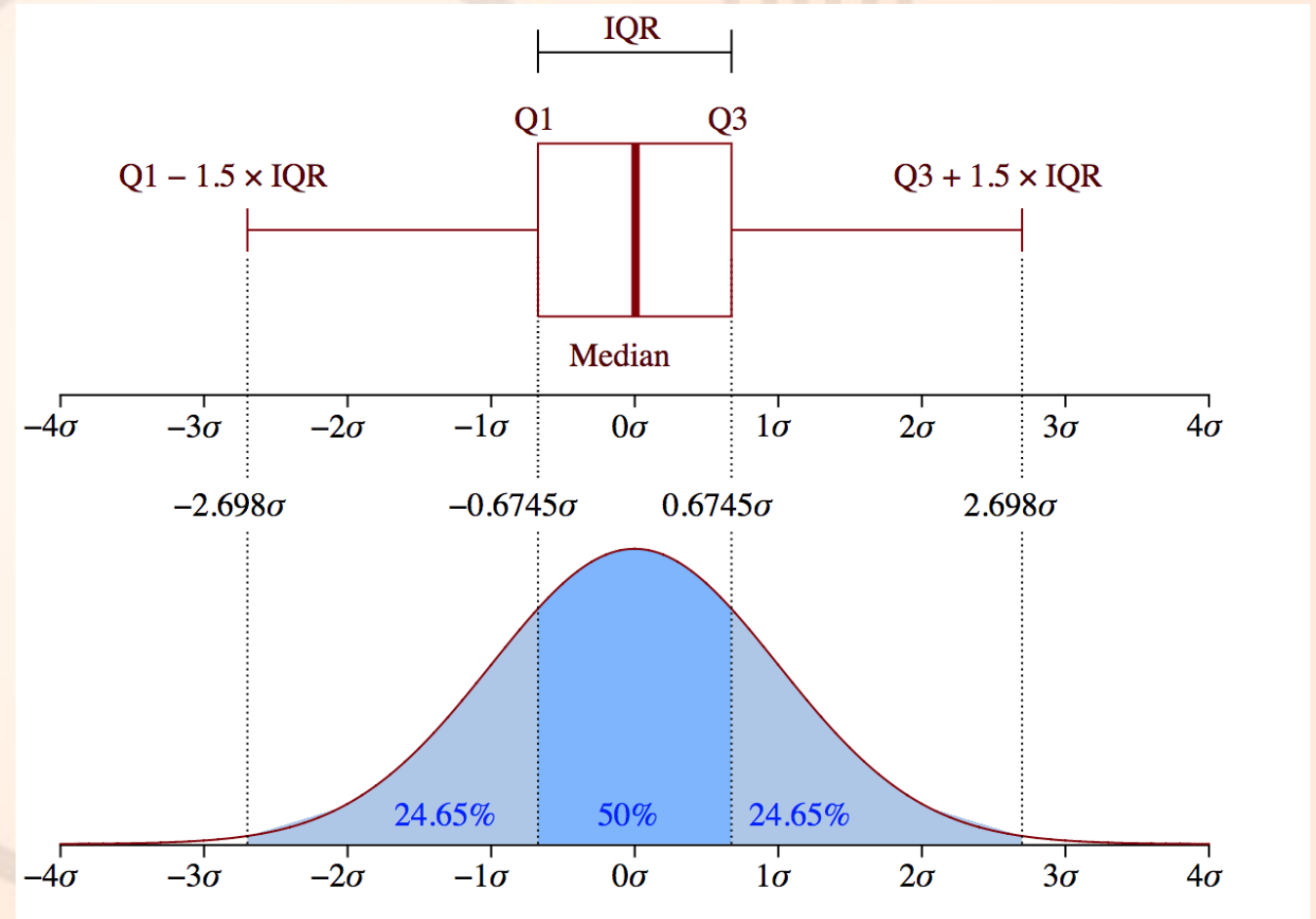
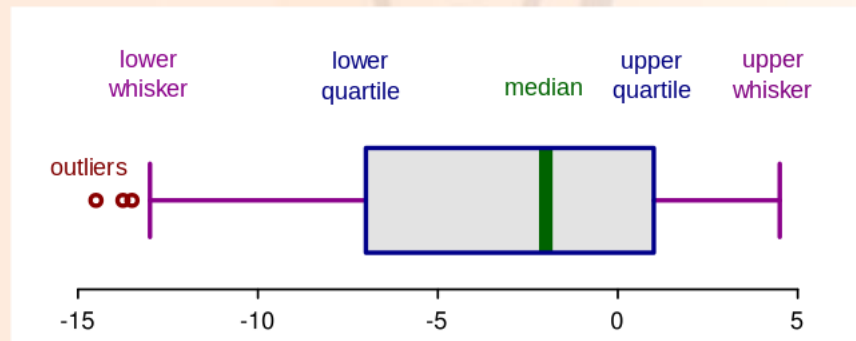
Répartition Produit par Nutri-Score

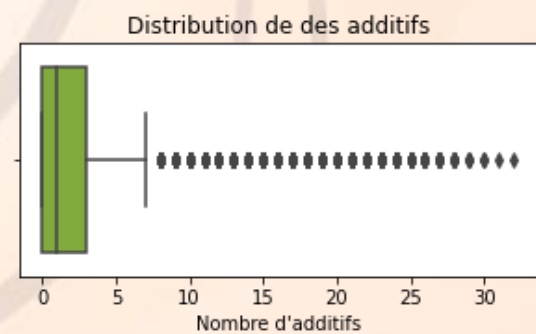
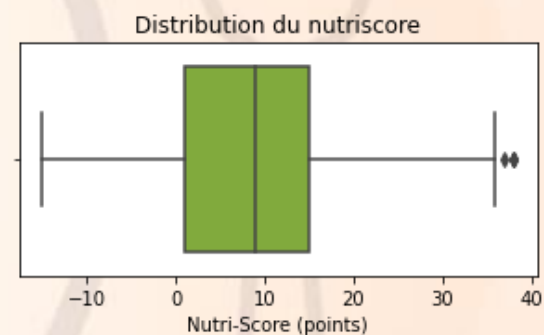
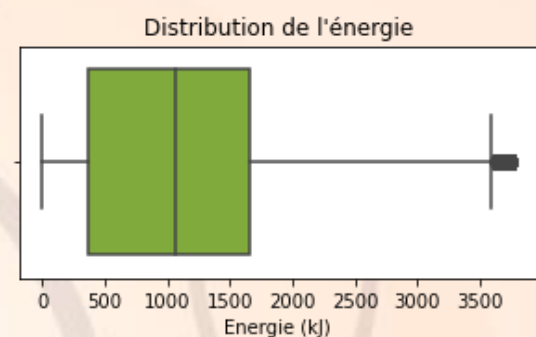
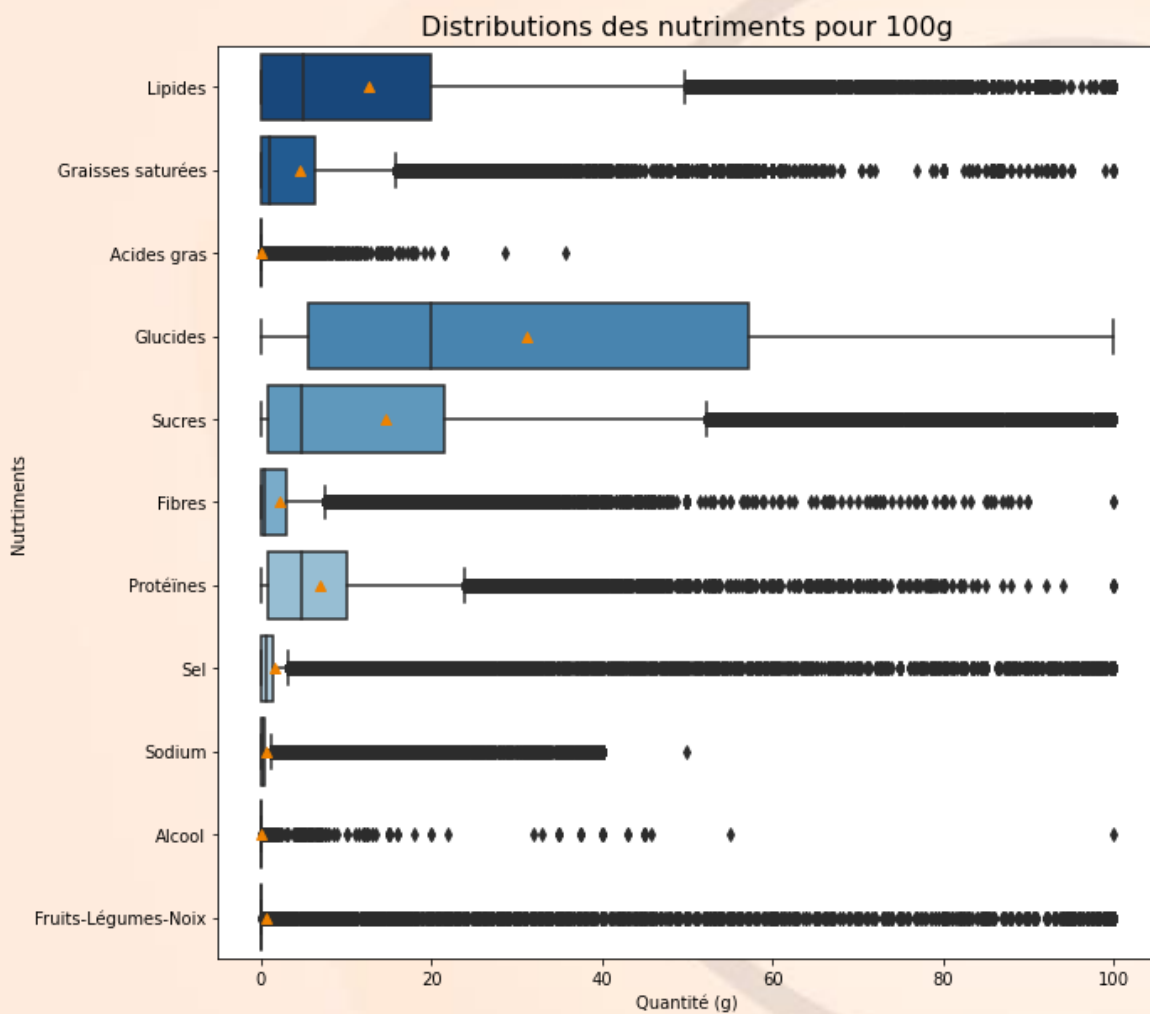


Répartition Produit par Nutri-Score



# Variables quantitatives





# ANALYSE BIVARIÉE

Tableau de corrélation de Pearson

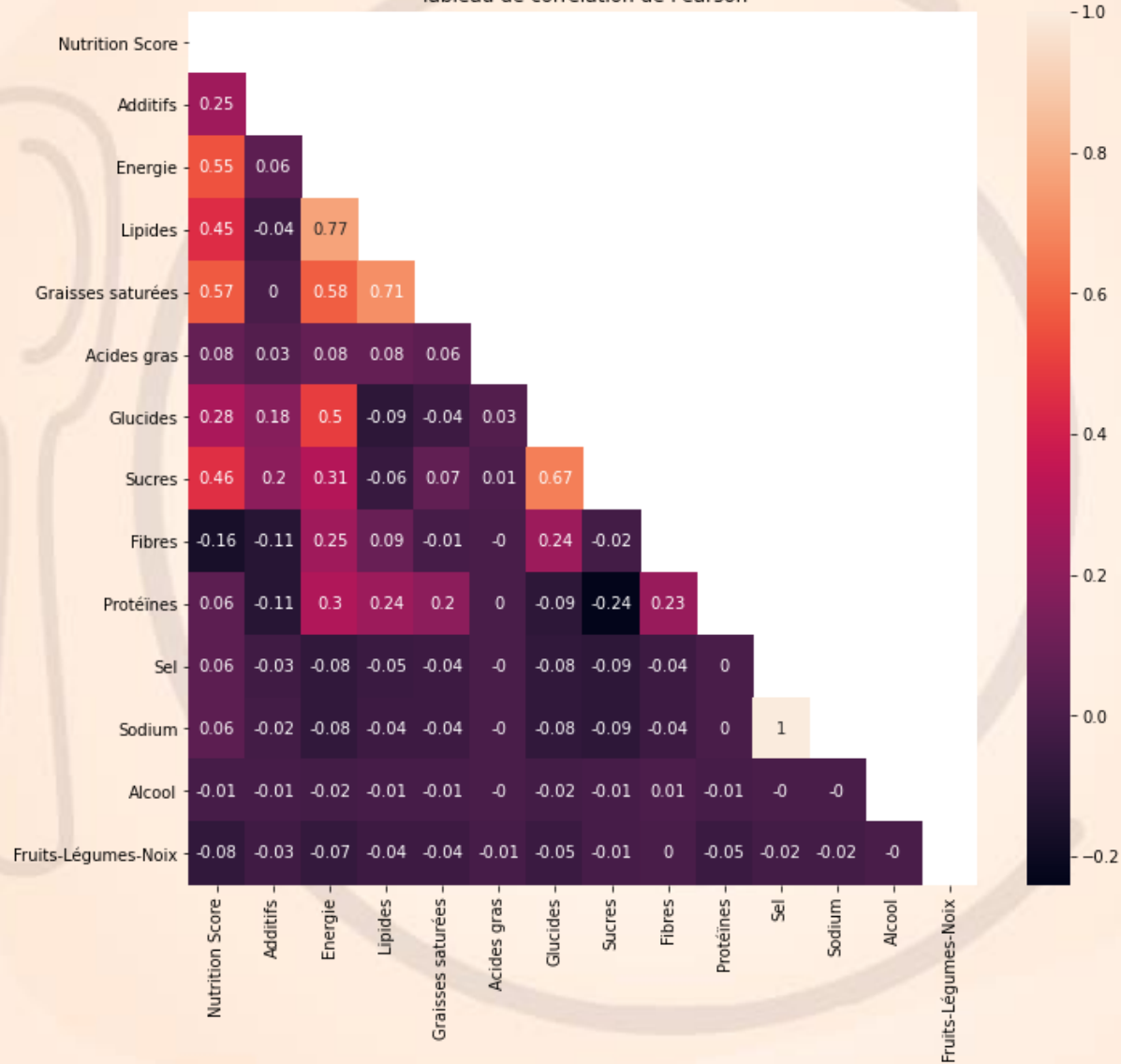




Tableau de corrélation de Pearson  
Variables concernant l'énergie

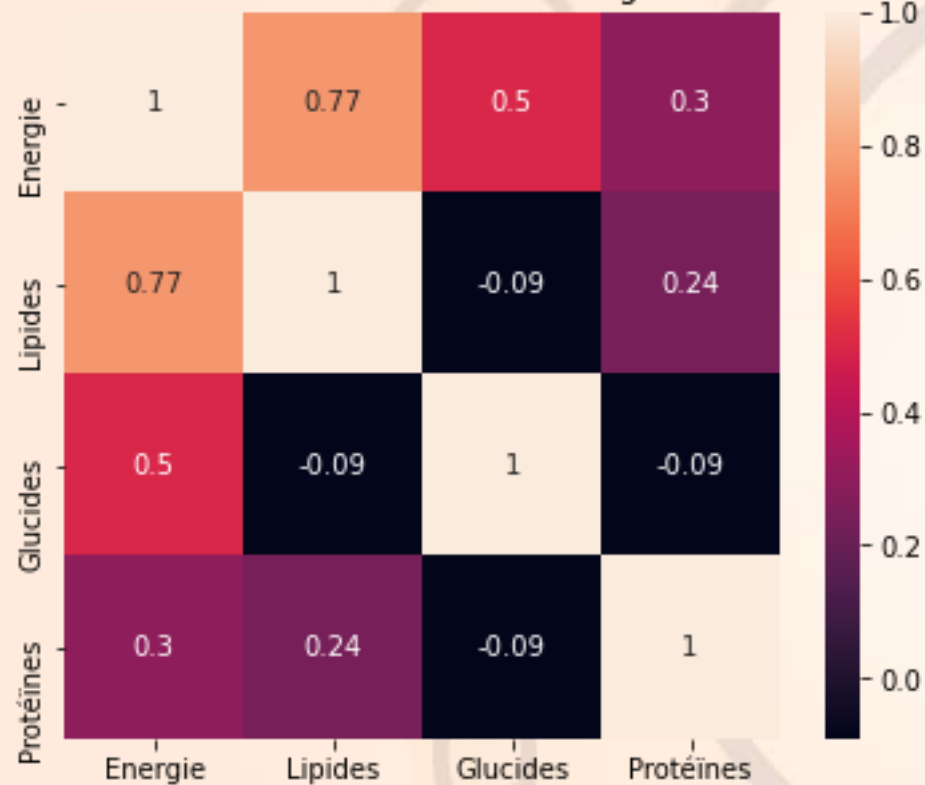
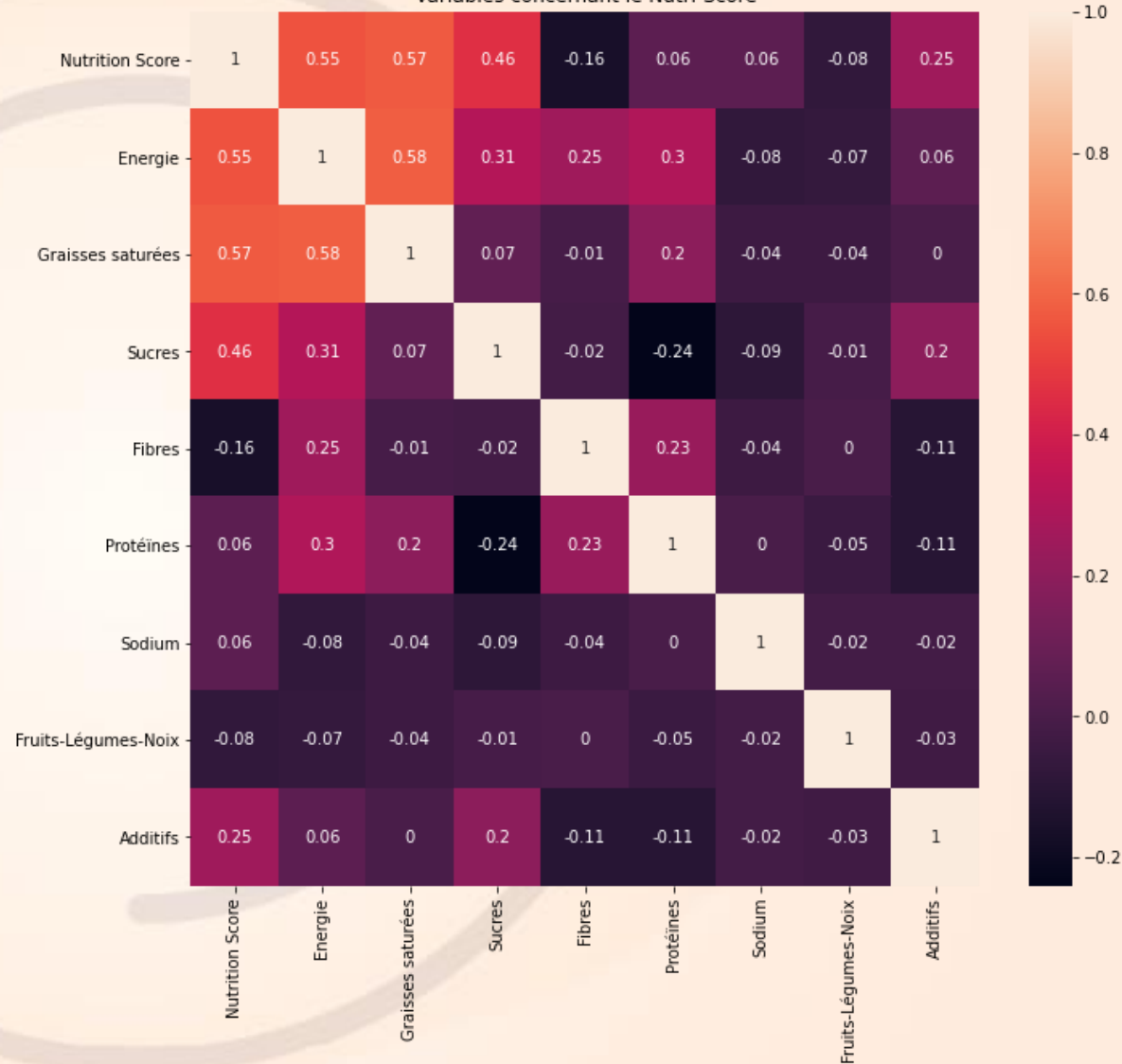


Tableau de corrélation de Pearson  
Variables concernant le Nutri-Score



# ANOVA

## Principe:

Tester si une variable qualitative et une variable quantitatives sont indépendantes ou s'il existe un lien

## Hypothèse de validation:

- $H_0$  : Les variables cibles et le Nutri-Score sont indépendantes
- $H_a$ : Il existe un lien

## Conclusion:

$p\_value < 0,05$  on rejette  $H_0$  soit il existe un lien entre le Nutri-Score et les variables cibles

## Variables Cibles

- |                     |                       |
|---------------------|-----------------------|
| - Score Nutrition   | - Fibres              |
| - Energie           | - Protéines           |
| - Graisses saturées | - Sodium              |
| - Sucres            | - Fruits-Légumes-Noix |

## Observations:

- Forte corrélation de la variable Score nutrition
- Bonne corrélation des variables Graisses saturées, Energie, et Sucres

# KHI-CARRE

## Principe:

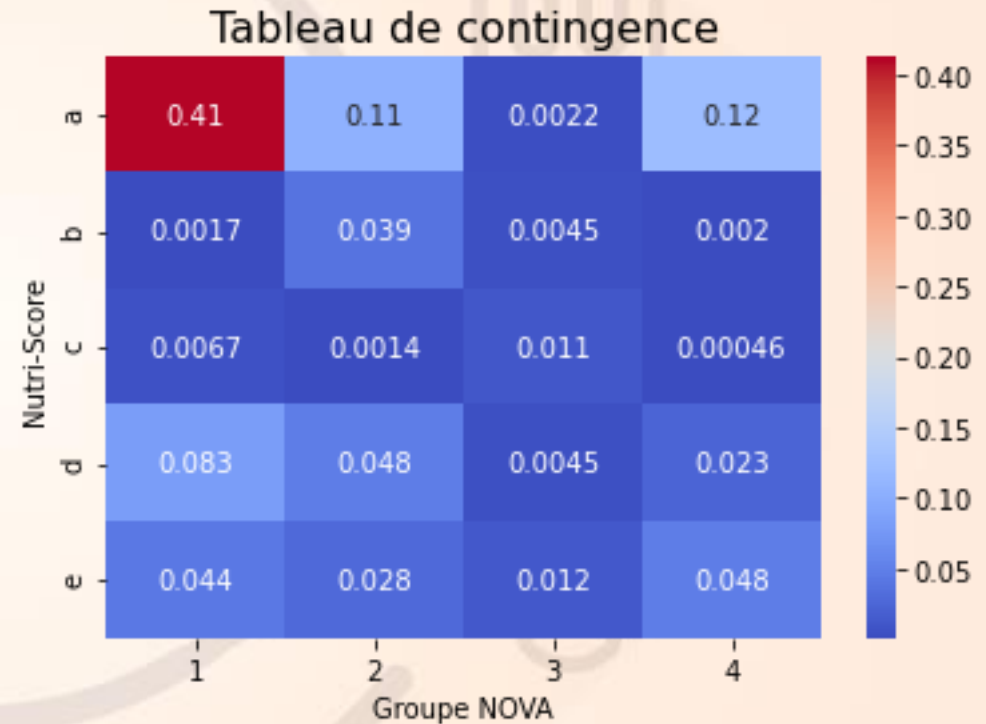
Tester si les variables qualitatives sont indépendantes ou s'il existe un lien

## Hypothèse de validation:

- $H_0$  : Les variables Nutri-Score et Groupe NOVA sont indépendantes
- $H_a$ : Il existe un lien

## Conclusion:

$p\_value < 0,05$  on rejette  $H_0$  soit il existe un lien entre les 2 variables

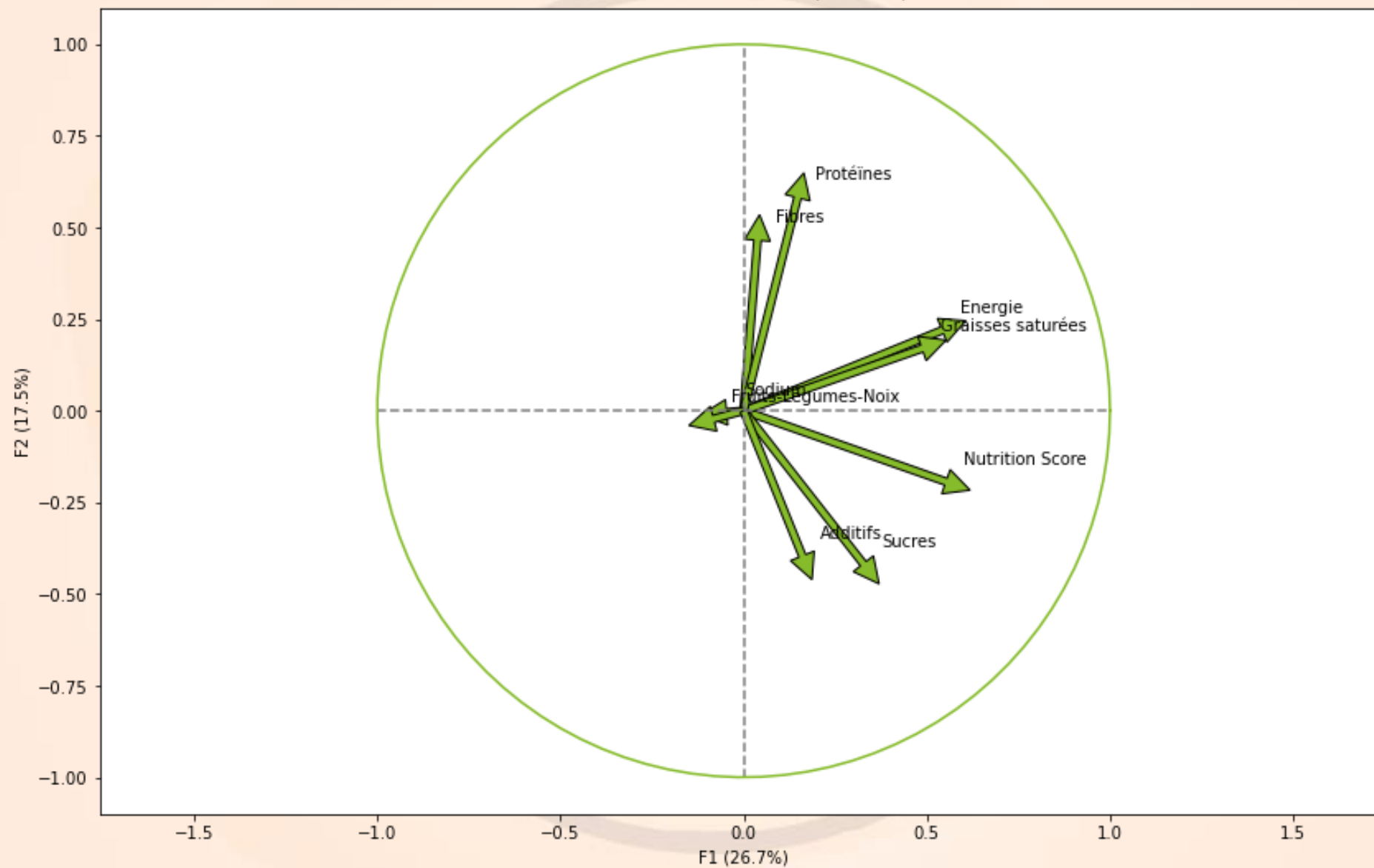


# ANALYSE MULTIVARIÉE

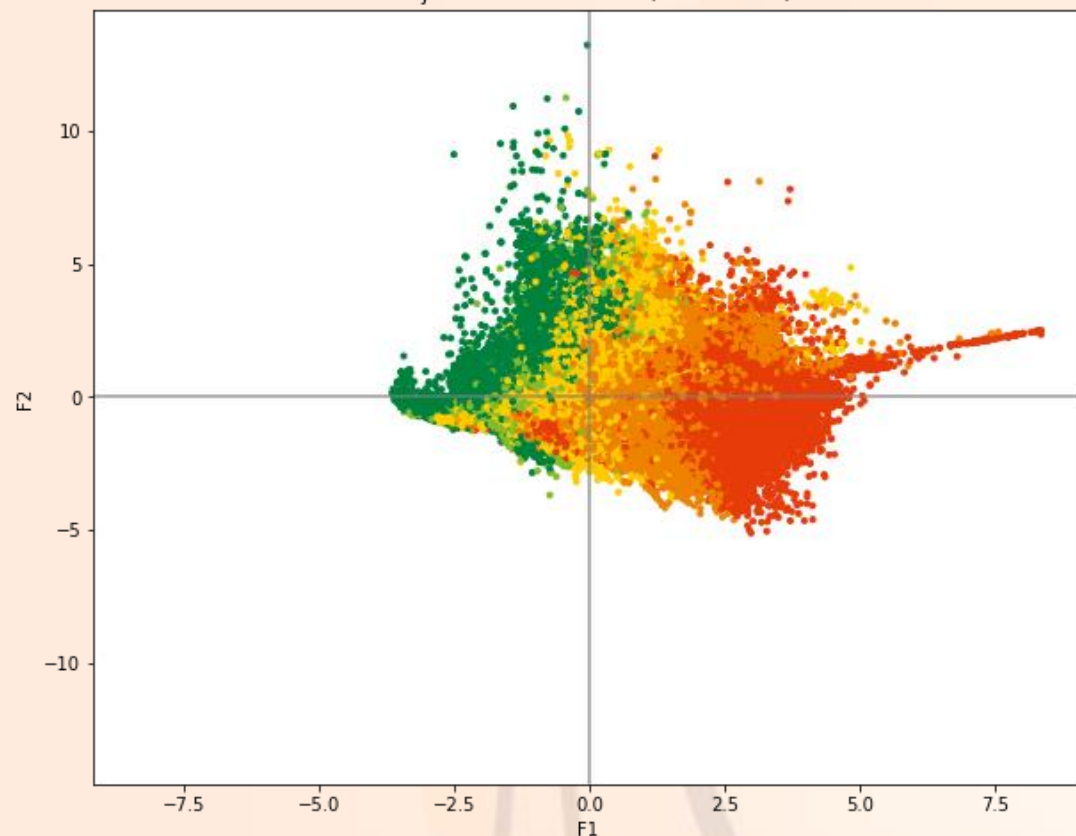
# ANALYSE DES COMPOSANTES MULTIPLES

- Permet de grouper des variables corrélées en variables synthétiques  
On peut observer les variables et leurs axes d'inertie à travers un cercle de corrélation
- Observer la variabilité entre les individus  
On peut observer les variabilité des individus à travers un nuage d'individus

Cercle des corrélations (F1 et F2)



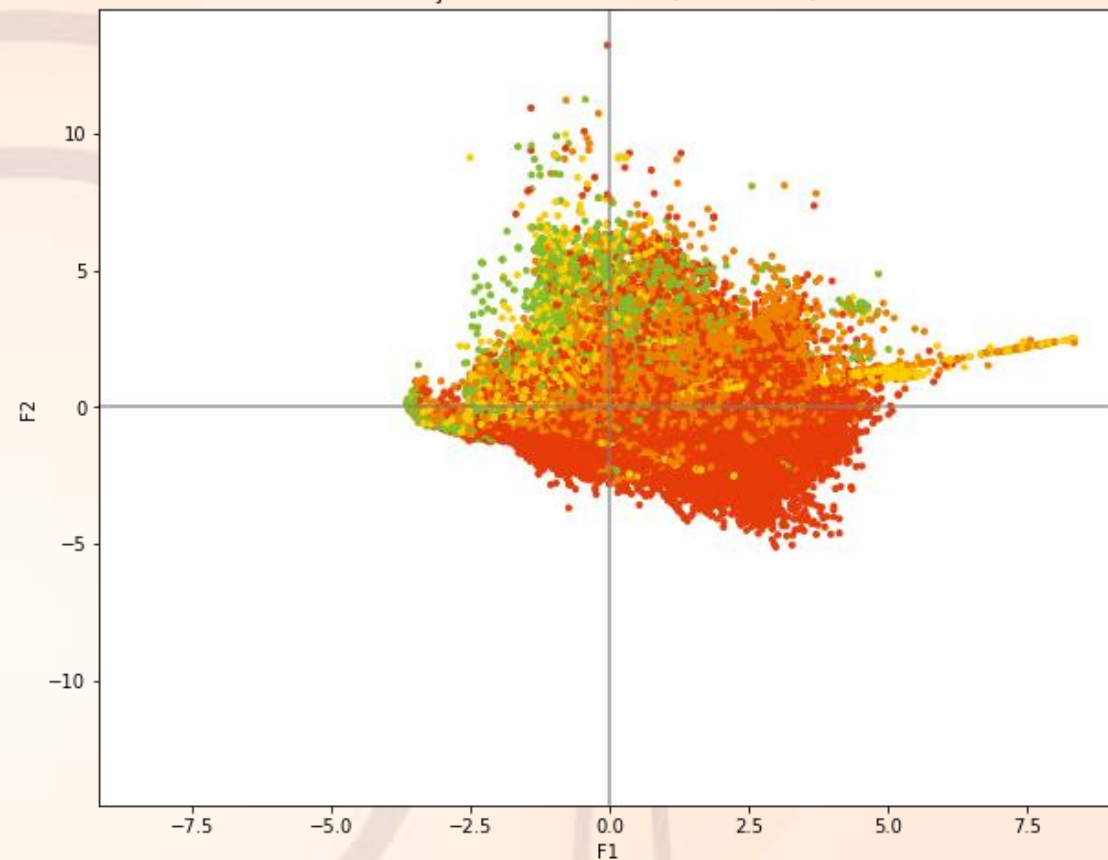
Projection des individus (sur F1 et F2)



NUTRI-SCORE



Projection des individus (sur F1 et F2)



NOVA



# APPLICATION



## Avantages :

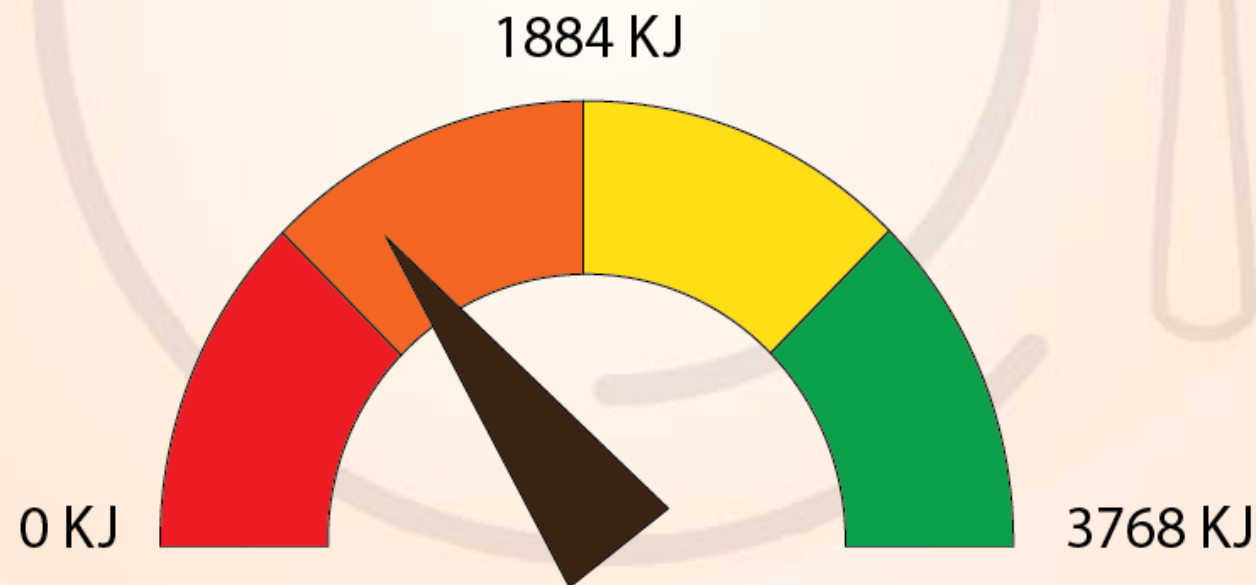
- Contrôle de l'énergie
- Limitations des macronutriments à éviter
- Limitations des additifs
- Propositions des aliments les mieux classées

## Public visé :

- Les sportifs
- Les personnes souhaitant surveiller leurs consommations de macronutriments
- Les personnes qui font un rééquilibrage alimentaire

## Avenir :

- Sélections des régimes spéciaux
- En accord avec l'évolution des classifications



# MERCI POUR VOTRE ATTENTION

- Sources:

<https://quoidansmonassiette.fr/comment-est-calcule-le-nutri-score-logo-nutritionnel/>

Mémoire DESU « Nutrition et Santé » - « EFFET DE LA SENSIBILISATION AU DEGRE DE TRANSFORMATION DES ALIMENTS SUR LA COMPOSITION CORPORELLE » - JOHAN DEBRABANT