



Soutenance Projet OC P3 DS: Concevez une application au service de la santé publique

01/07/2021

Candidat: David CAPELLE
Mentor: Nicolas MICHEL

Evaluateur: Sara A. Benoist
Formation 100% Pôle Emploi

Plan de la soutenance

- Description de l'idée d'application Nutri-rein et principes du calcul
- Constitution du jeu de données et nettoyage
- Présentation de l'analyse exploratoire des données
 - Analyse univariée
 - Analyse multivariée
- Présentation des faits pertinents pour l'application
- Conclusion sur la faisabilité de l'application et axes d'amélioration

Présentation de l'idée d'application : *Nutri-rein*

- Calcul d'un Nutri-rein (score et grade)
 - A partir des valeurs des nutriments (sodium, sucres, glucides, graisses,...) présents dans la base Open Food Facts.
 - Comparaison de ce score au Nutri-score.
 - Calcul par rapport à l'apport en nutriments et en énergie journalier nécessaire à un patient en insuffisance rénale.
 - Calcul basé sur les expertises de néphrologues du CHU de Lyon, pour les apports attendus en nutriments.
- Un insuffisant rénal, c'est quoi ?
 - Patient dont le DFG est inférieur à $60 \text{ mL/min}1.73\text{m}^2$.
 - Le DFG mesure la quantité de liquide (sang) filtrée par le rein, sous forme d'un dosage de créatinine.
 - Cas de l'étude : insuffisance rénale modérée (DFG entre 30 et 40).

Présentation de l'idée d'application : *Nutri-rein*

Besoins en nutriments / énergie d'un insuffisant rénal

- Apport en sodium : pas plus de 3 g par jour
 - Apport en énergie : environ 146 kJ/kg poids corporel/jour.
 - Apport en protéines : pas plus de 0,8/kg poids corporel/jour.
 - Apport en fibres : au moins 30 g/jour avec 50 % de végétaux.
 - Apport en glucides : pas plus de 50 % de l'apport en énergie.
 - Apport en lipides : pas plus de 40 % de l'apport en énergie, dont au moins 50 % de graisses insaturées.
 - Apport en sucres : pas plus de 25 g par jour.
- => principe : limite des apports en sodium, énergie, glucides, lipides et sucres, favoriser les apports en protéines et fibres.**

Présentation de l'idée d'application : *Nutri-rein*

Pré-requis/ hypothèses pour le calcul

- Adulte-type de sexe masculin, de poids moyen de 80 kg, avec une activité physique moyenne.
- Calcul du score des produits en fonction de l'apport en nutriments / énergie journalier dans le cadre de 3 repas-type (petit-déjeuner, repas midi, repas soir) :
 - 4 ingrédients pour le petit-déjeuner (1 boisson chaude, 2 aliments, 1 boisson froide)
 - 4 ingrédients pour les repas midi/soir (boisson, entrée, plat, dessert)
 - Soit 12 produits dans la journée, avec le même poids, chaque élément d'un repas ne dépassant 1/12 de l'apport journalier d'un adulte-type
- Valeurs nutriments / apport énergétique pour 100 g de produit (exemples):
 - Sodium : pas plus de 0,25 g
 - Energie : pas plus de 700 kj
 - Fibres : au moins 2,5 g,...

Présentation de l'idée d'application : *Nutri-rein*

Calcul des nutriments/énergie à limiter – points négatifs (N)

Grille de calcul des points négatifs

Point N	energy_100g(kj)	sodium_100g(g)	carbohydrates_100g(g)	sugars_100g(g)	fat_100g(g)	saturated_fat(% fat)
0	<= 100 kj	<= 0.01 g	<= 1 g	<= 0 g	<= 1 g	<= 5 % fat_100g
1	> 100 kj	> 0.01 g	> 1 g	> 0.1 g	> 1 g	> 5 % fat_100g
2	> 150 kj	> 0.05 g	> 3 g	> 0.2 g	> 1.5 g	> 10 % fat_100g
3	> 200 kj	> 0.07 g	> 5 g	> 0.4 g	> 2 g	> 15 % fat_100g
4	> 300 kj	> 0.1 g	> 7 g	> 0.8 g	> 3 g	> 20 % fat_100g
5	> 350 kj	> 0.12 g	> 9 g	> 1 g	> 3.5 g	> 25 % fat_100g
6	> 400 kj	> 0.14 g	> 11 g	> 1.2 g	> 4 g	> 30 % fat_100g
7	> 450 kj	> 0.16 g	> 13 g	> 1.4 g	> 5 g	> 35 % fat_100g
8	> 500 kj	> 0.18 g	> 15 g	> 1.6 g	> 6 g	> 40 % fat_100g
9	> 600 kj	> 0.20 g	> 17 g	> 1.8 g	> 6.5 g	> 45 % fat_100g
10	> 700 kj	> 0.25 g	> 20 g	> 2 g	> 7.5 g	> 50 % fat_100g

- Note sur 10 points pour les nutriments / apport énergétiques et note sur le ratio graisses saturées / graisses.
- Total des points négatifs de 0 à 60.
- Pas de distinction entre les aliments solides et les boissons pour le calcul.

Présentation de l'idée d'application : *Nutri-rein*

Calcul des nutriments à privilégier – points positifs (P)

- Grille de calcul des points positifs

Point P	fibers_100g(g)	proteins_100g(g)
0	> 0 g	> 0 g
1	> 0.5 g	> 0.5 g
2	> 1 g	> 1 g
3	> 1.5 g	> 2 g
4	> 2 g	> 3 g
5	> 2.5 g	> 4.3 g

- Note sur 5 points pour les nutriments à privilégier pour un IRC.
- Total des points négatifs de 0 à 10.

Présentation de l'idée d'application : *Nutri-rein*

Calcul du score global et des grades

- Score global Nutri-rein = somme points N – somme points P
- Le score global Nutri-rein permet d'affecter un score de -10 (produits de bonne qualité nutritionnelle) à + 60 (produits de mauvaise qualité nutritionnelle).
- Les grades Nutri-rein sont répartis de la façon suivante :
 - Grade 'a' pour les scores de -10 à 1
 - Grade 'b' pour les scores de 2 à 25
 - Grade 'c' pour les scores de 26 à 37
 - Grade 'd' pour les scores de 38 à 48
 - Grade 'e' pour les scores de 49 à 60

Constitution du jeu de données et nettoyage:

Description du jeu de données initial

- Le jeu de données initial décrit les données nutritionnelles et énergétiques de 1 786 430 produits avec 186 variables.
- Présence de données catégorielles (nom du produit, catégories de produit, marques, points de vente,...).
- Présence de données quantitatives (composition nutritionnelle, énergie, nombre d'additifs, d'allergènes,...).
- Et aussi des méta-données (date d'entrée du produit dans la base, date de mise à jour,...).
- Certains nutriments sont décomposés en sous-composants (graisses saturées, sucres,...).
- Le dataset contient globalement près de 80 % de valeurs manquantes

Constitution du jeu de données et nettoyage:

Constitution du jeu de données pour l'application

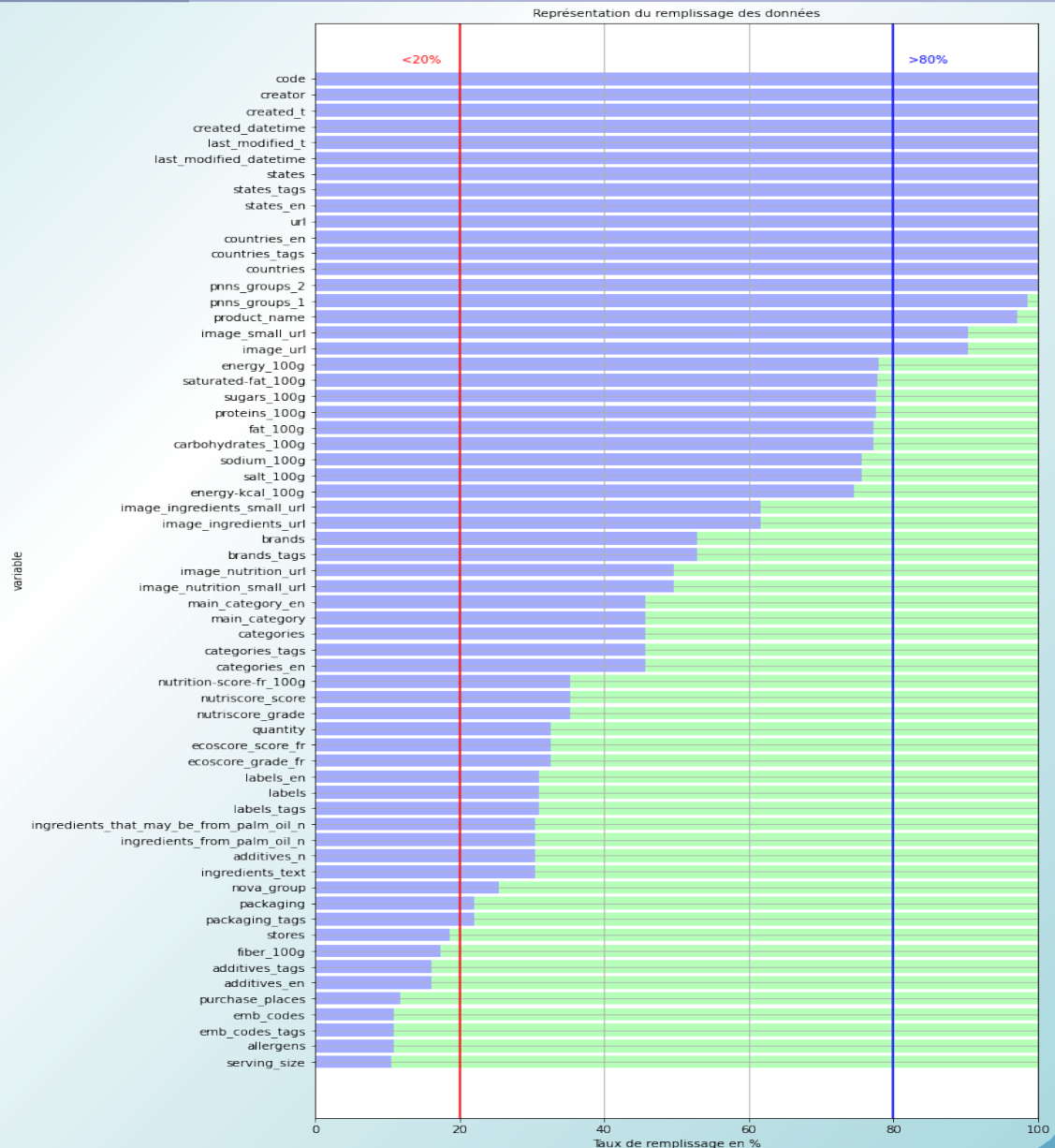
- Etape 1 : suppression des doublons => 2 lignes supprimées.
- Etape 2 : sélection des produits uniquement vendus sur le sol français (métropole et DOM/TOM).
=> suppression de produits pour retenir 801 262 produits (45 % des données).
- Etape 3 : représentation graphique du ratio de données disponibles dans le dataset par variable.
- Etape 4 : suppression des colonnes du dataset avec moins de 10 % de données disponibles. Il ne subsiste que certaines variables nutriments (voir slide suivant).

=> à l'issue de l'étape 4, le dataset contient 801 262 produits et 63 variables, avec un peu plus de 42 % de valeurs NaN.

Constitution du jeu de données et nettoyage:

Représentation du ratio de données disponibles suite étape 4

- Certains nutriments subsistent suite suppression des colonnes à l'étape 4 : énergie, graisses, sel, sodium, graisses, fibres et glucides.
- Intérêt de ces nutriments pour calcul du Nutri-rein.
- D'autres nutriments ont été écartés (calcium, potassium, phosphores).



Constitution du jeu de données et nettoyage:

Suppression des variables inutiles pour le Nutri-rein

- Etape 5 : suppression des variables inutiles pour le Nutri-rein :
 - Méta-données,
 - Emballages et codes emballage,
 - Images de produit, liens hypertextes,
 - Etat de l'observation (complet, mise à jour,...),
 - Regroupements par catégorie de produit, sauf variable `pnns_groups_2`
 - Labels, marques,
 - Lieux d'achat, de vente,
 - Présence d'allergènes et/ou d'additifs,
 - Présence d'huile de palme
 - Variables sur les scores autres que le Nutri-score
 - Quantité / portion des produits
- => à l'issue de l'étape 5, le dataset contient 801 262 produits et 23 variables, avec un peu moins de 37 % de valeurs NaN.**

Constitution du jeu de données et nettoyage:

Suppression lignes avec produit non renseigné

Traitement des valeurs manquantes

- Etape 6 : suppression des lignes pour lesquelles le nom du produit n'est pas renseigné.
=> à l'issue de l'étape 6, le dataset contient 778 191 produits et 23 variables, avec un peu moins de 36 % de valeurs NaN.
- Etape 7 : Traitement de valeurs NaN pour les variables quantitatives
Suppression des lignes avec des valeurs NaN
=> à l'issue de l'étape 7, le dataset contient 87 884 produits et 23 variables, avec un peu moins de 3,5 % de valeurs NaN.
- Etape 8 : Traitement de valeurs NaN pour les variables qualitatives
_ Imputation de la valeur 'unkwown' pour remplacer les valeurs NaN
=> à l'issue de l'étape 8, le dataset contient 87 884 produits et 23 variables, sans valeur NaN.

Constitution du jeu de données et nettoyage:

Etape 9 - Traitement des outliers – valeurs aberrantes

- Suppression des produits dont la valeur de la variable « energy-kcal_100g » est > 900 kcal, et les lignes dont la valeur de la variable « energy_100g » est > 4000 kj.
 - Suppression des produits dont la valeur des variables nutriments est > 100 g ou < à 0 g.
 - Suppression des produits dont la teneur en sel (salt_100g) dépasse 8 g (danger grave pour la santé).
 - Suppression des produits dont la teneur en sodium (sodium_100g) dépasse 3 g (danger grave pour la santé).
 - Suppression des produits dont la teneur en graisses saturées dépasse la teneur en graisses.
 - Suppression des produits dont somme nutriments > 100 g.
- => A l'issue de l'étape 9, le dataset contient 85 159 produits et 23 variables, sans valeur NaN.**

Présentation de l'analyse exploratoire:

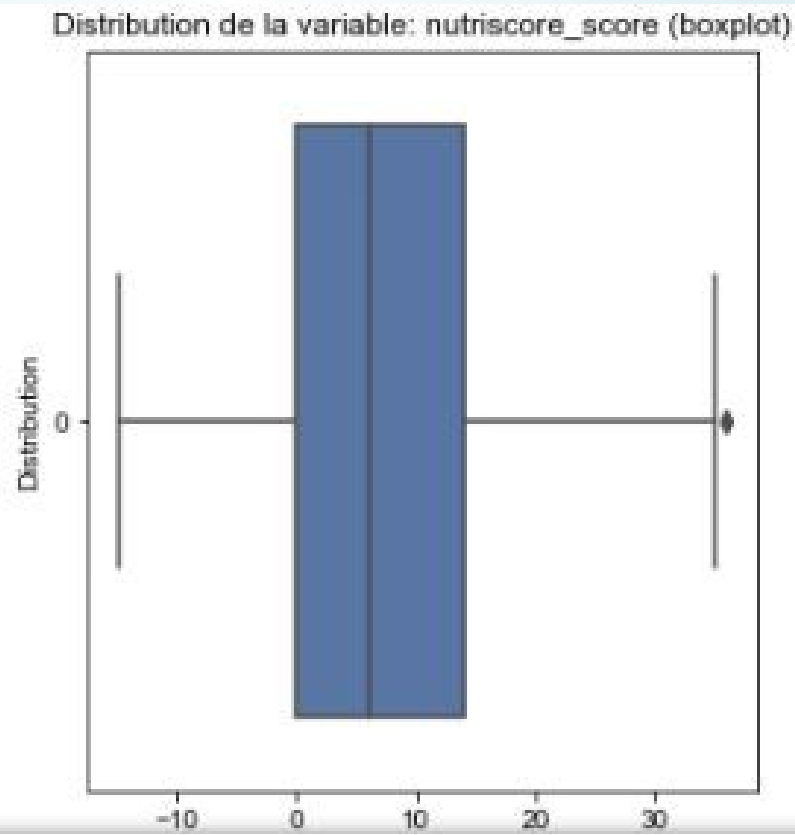
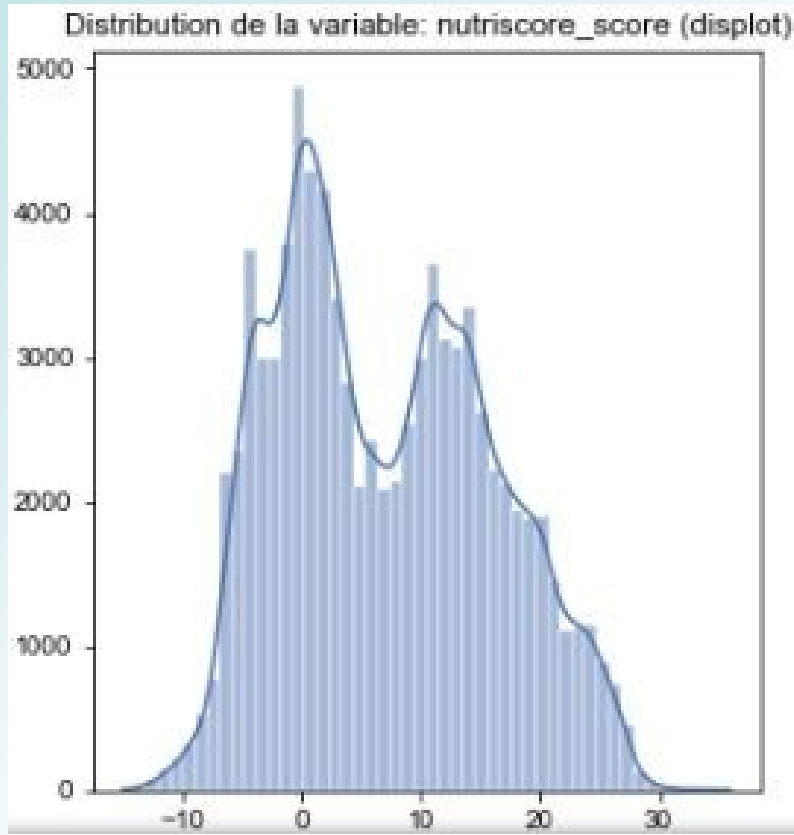
Analyse univariée – variables quantitatives

- Distribution bimodale du nutri-score (variables nutriscore_score et nutrition-score-fr_100g) représentant 2 phénomènes différents (un phénomène avec mode = 0 et l'autre avec mode = 15).
- Certaines variables semblent décrire le même phénomène dans leur distribution (variables salt_100g et sodium_100g et variables nutriscore_score et nutrition-score-fr_100g) .
=> cette similitude sera analysée dans le cadre de l'analyse multivariée.
- Beaucoup de variables ont une forme de distribution semblable, distribution asymétrique à droite (fat_100g, fiber_100g, proteins_100g,...).
- Sur toutes les variables, échantillonnage continu des données et les outliers présents sur les boxplots représentent seulement des valeurs atypiques, suite au traitement des outliers.

Présentation de l'analyse exploratoire:

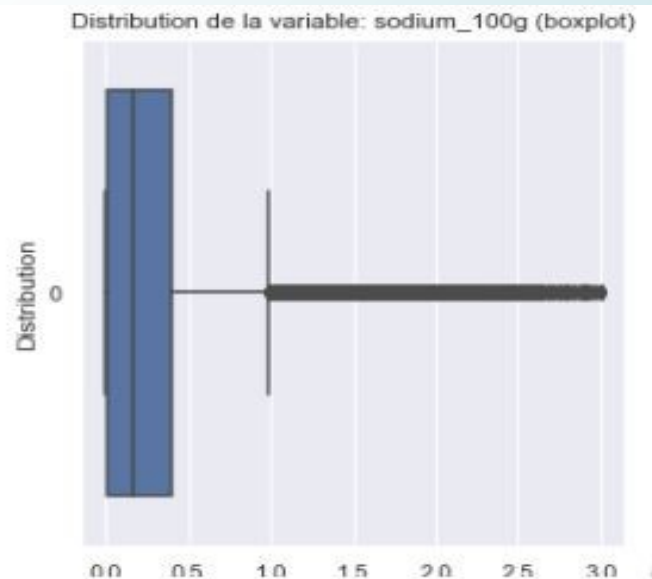
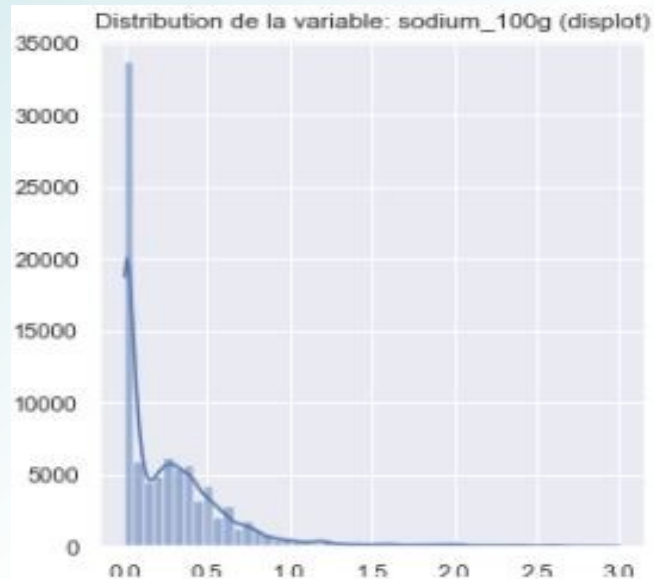
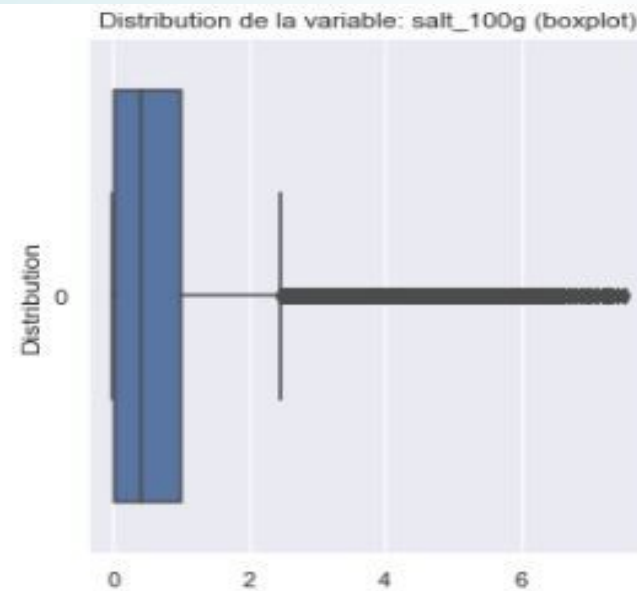
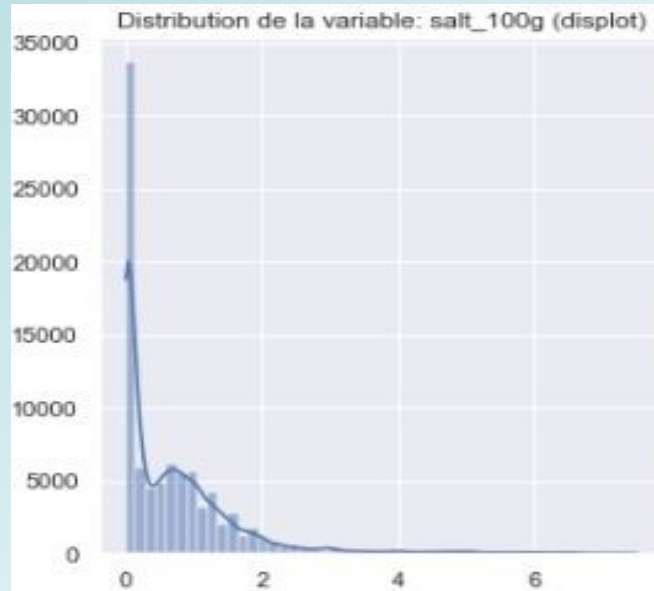
Analyse univariée – exemple distribution bimodale

- La distribution du nutri-score présente 2 phénomènes différents, un groupe de produits avec une qualité nutritionnelle correcte (mode = 0) et un autre groupe de moindre qualité (mode = 15).



Présentation de l'analyse exploratoire:

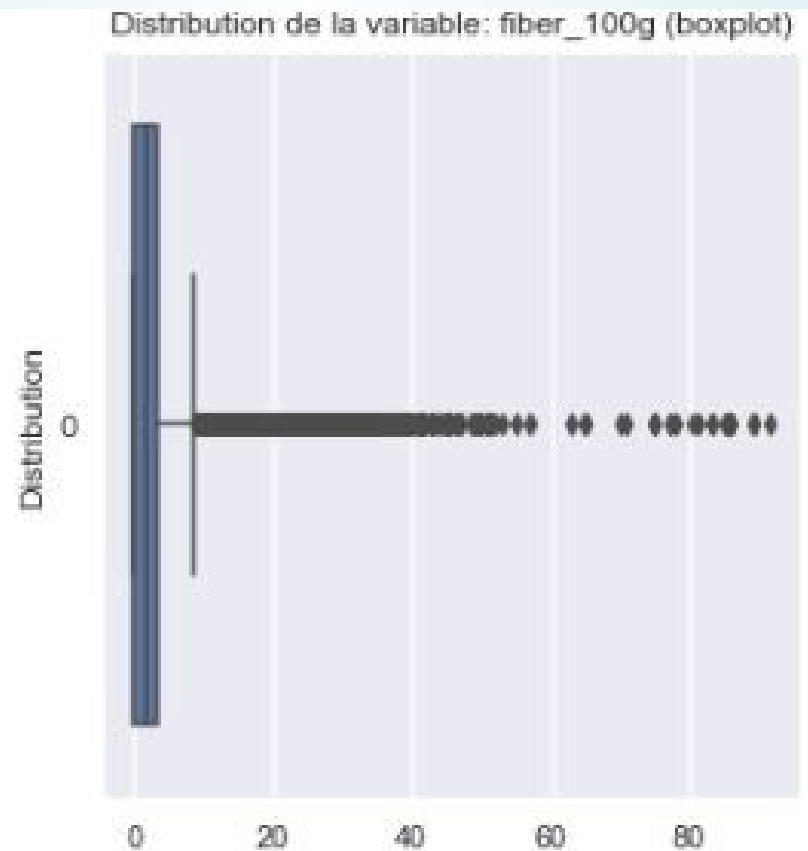
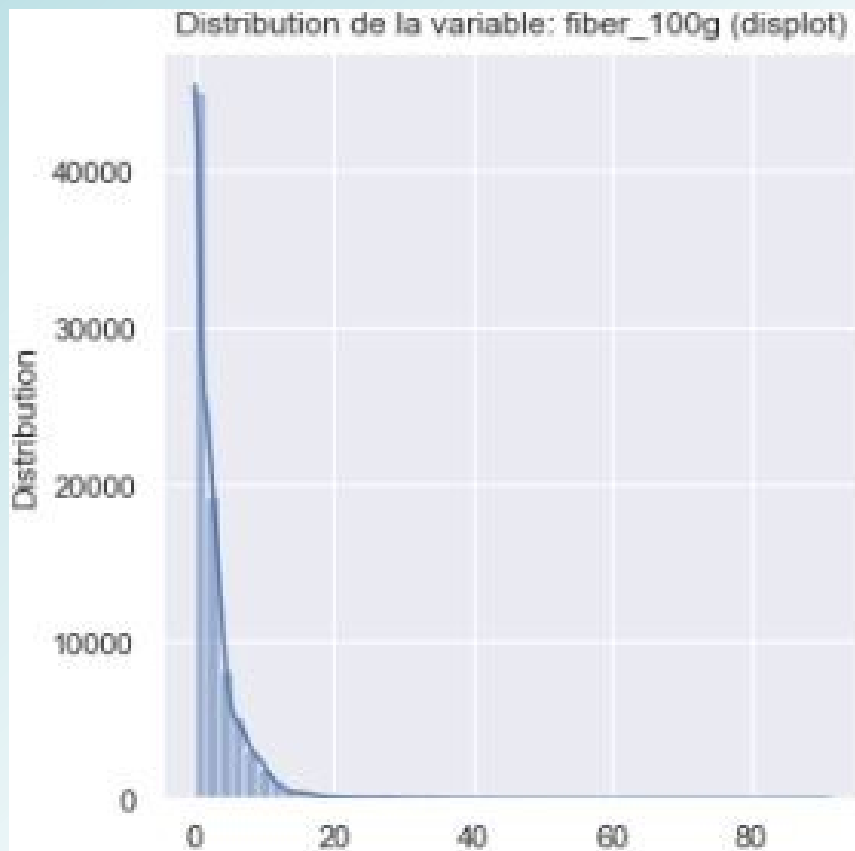
Analyse univariée – Similitude de phénomènes sel et sodium



Présentation de l'analyse exploratoire:

Analyse univariée – Forme de la distribution

- Forme de la distribution asymétrique à droite pour plusieurs variables dont la variable `fiber_100g`.



Présentation de l'analyse exploratoire:

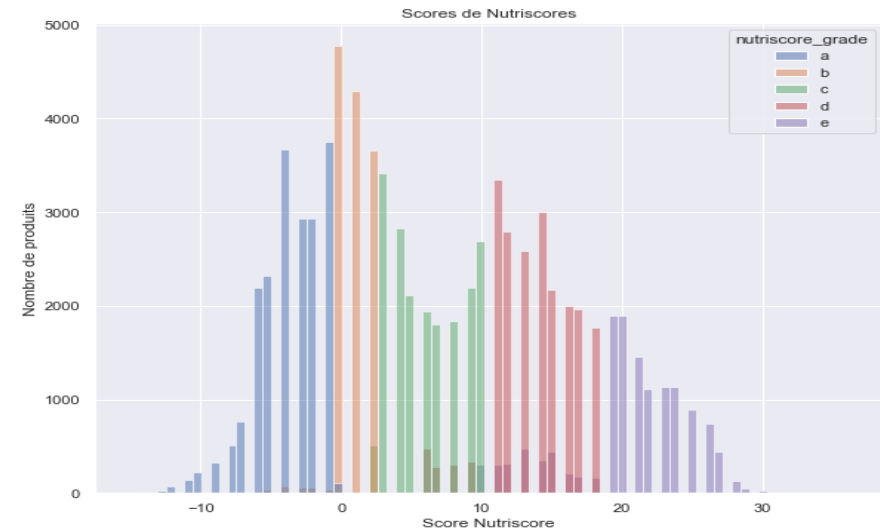
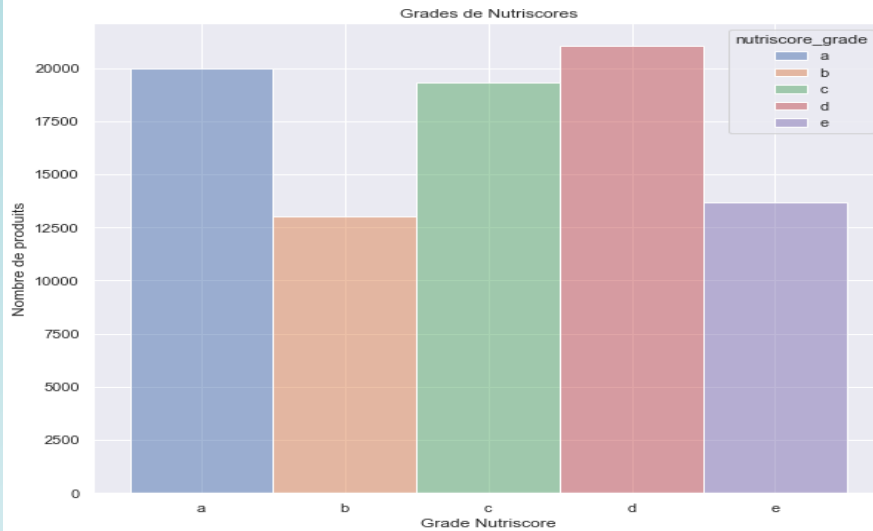
Analyse univariée – variables qualitatives

- Répartition quasiment « équi-répartie » des grades Nutri-score (avec une prépondérance du grade 'd').
- La répartition « naturelle » attendue serait plutôt une distribution normale.
- Distribution attendue : prépondérance des produits de qualité nutritionnelle moyenne (grade 'c') et peu de produits de très bonne qualité (grade 'a') et de très mauvaise qualité nutritionnelle (grade 'e') .
- Cette répartition est confirmée par la répartition des scores Nutri-score, avec une distribution bimodale.

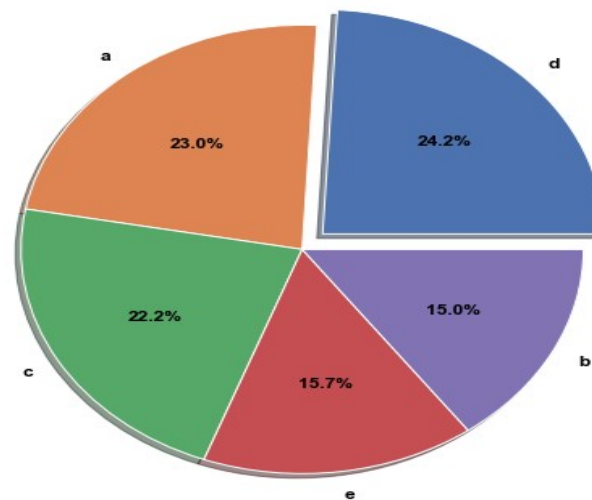
Présentation de l'analyse exploratoire:

Analyse univariée – Représentation variables qualitatives

Répartition des scores Nutriscore et de leurs grades



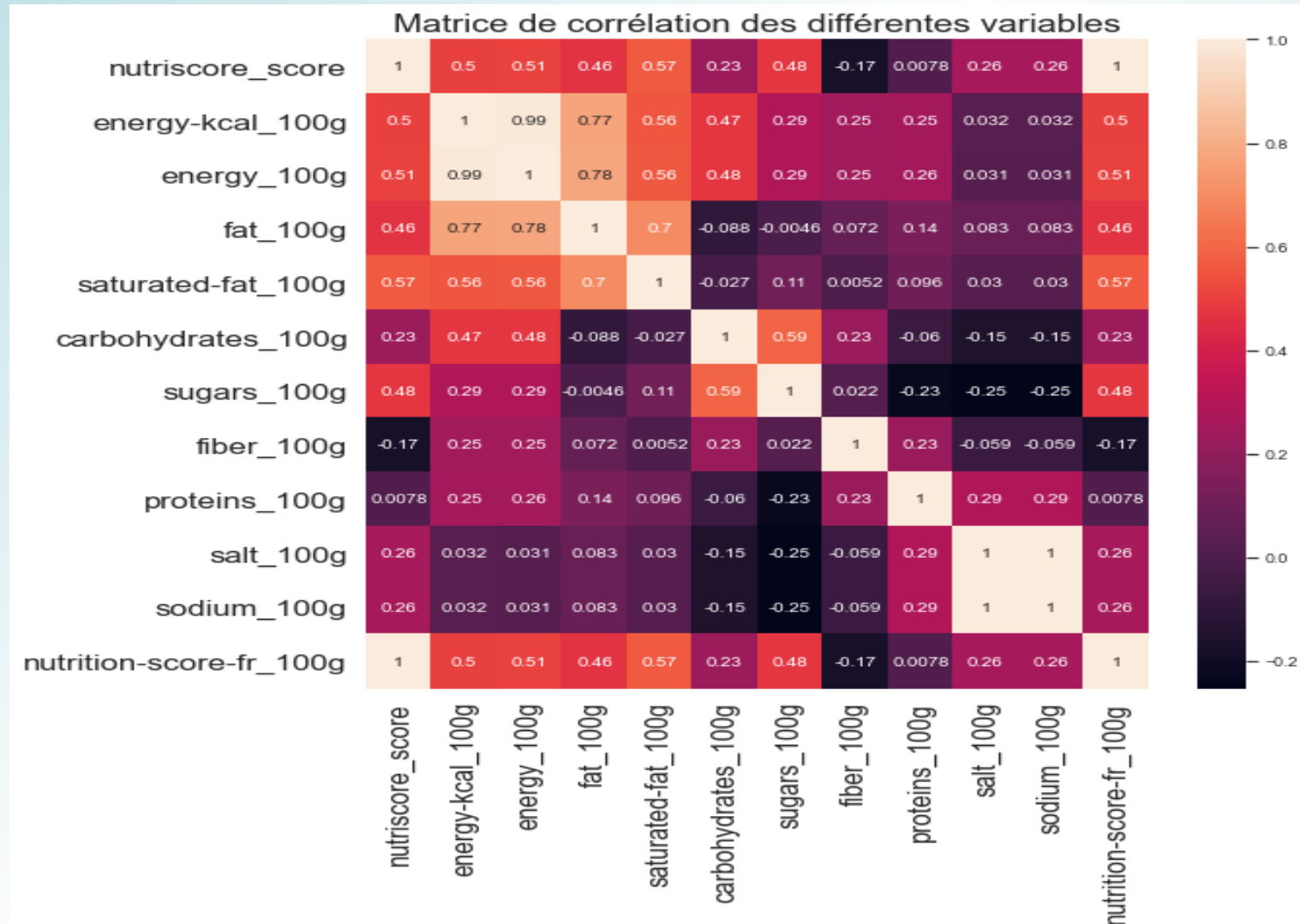
Répartition des Nutrigrades



Présentation de l'analyse exploratoire:

Analyse multivariée – Matrice des corrélations

- La matrice des corrélations permet de visualiser le niveau de corrélation entre variables.



Présentation de l'analyse exploratoire:

Analyse multivariée – Matrice des corrélations

- Forte corrélation ($\text{corr} = 1$) entre les variables « `nutriscore_score` » et « `nutrition-score-fr_100g` », entre « `salt_100g` » et « `sodium_100g` » et entre « `energy_100g` » et « `energy-kcal_100g` ».
- => ces variables représentent le même phénomène.**
- => suppression des variables « `salt_100g` », « `nutrition-score-fr_100g` » et « `energy-kcal_100g` ». On garde 8 nutriments et le `nutri-score`.**
- La variable « `energy_100g` » est corrélée ($\text{corr} = 0,78$) étroitement avec les graisses.
 - Les variables « `fat_100g` » et « `saturated-fat_100g` » sont corrélées ($\text{corr} = 0,7$).
 - La variable « `sugars_100g` » est corrélée ($\text{corr} = 0,59$) avec les glucides.

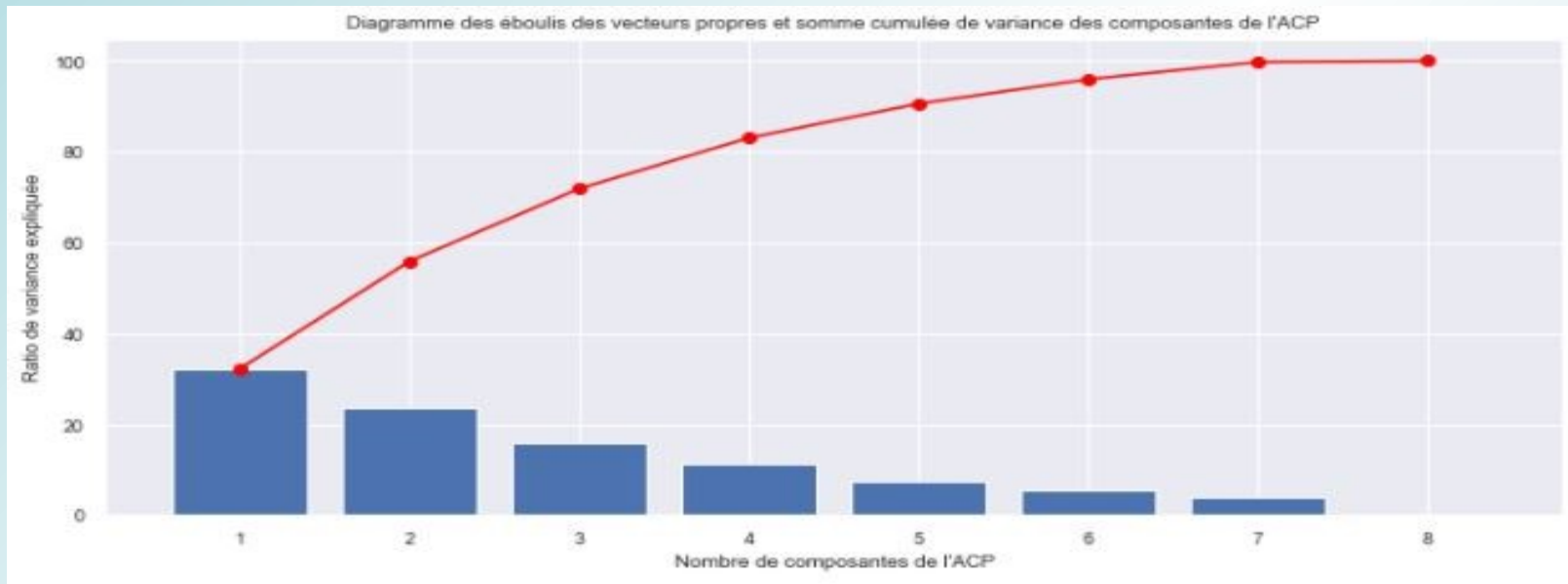
Présentation de l'analyse exploratoire:

Analyse multivariée – Réalisation d'une ACP

- ACP pour l'étude et la visualisation des corrélations entre variables, afin éventuellement de limiter le nombre de variables par une réduction de dimensions.
- ACP pour la visualisation des observations dans un espace à 2 dimensions, afin d'identifier des groupes homogènes d'observations, ou au contraire des valeurs atypiques.
- Validation de l'interêt de faire une réduction de dimensions pour notre application.
- Traçage du cercle des corrélations, calcul de la qualité de représentation des variables (COS^2) pour vérifier la représentativité des variables aux axes du plan factoriel.
- Calcul de la contribution des variables aux axes (CTR) sur les 2 premiers plans factoriels.

Présentation de l'analyse exploratoire:

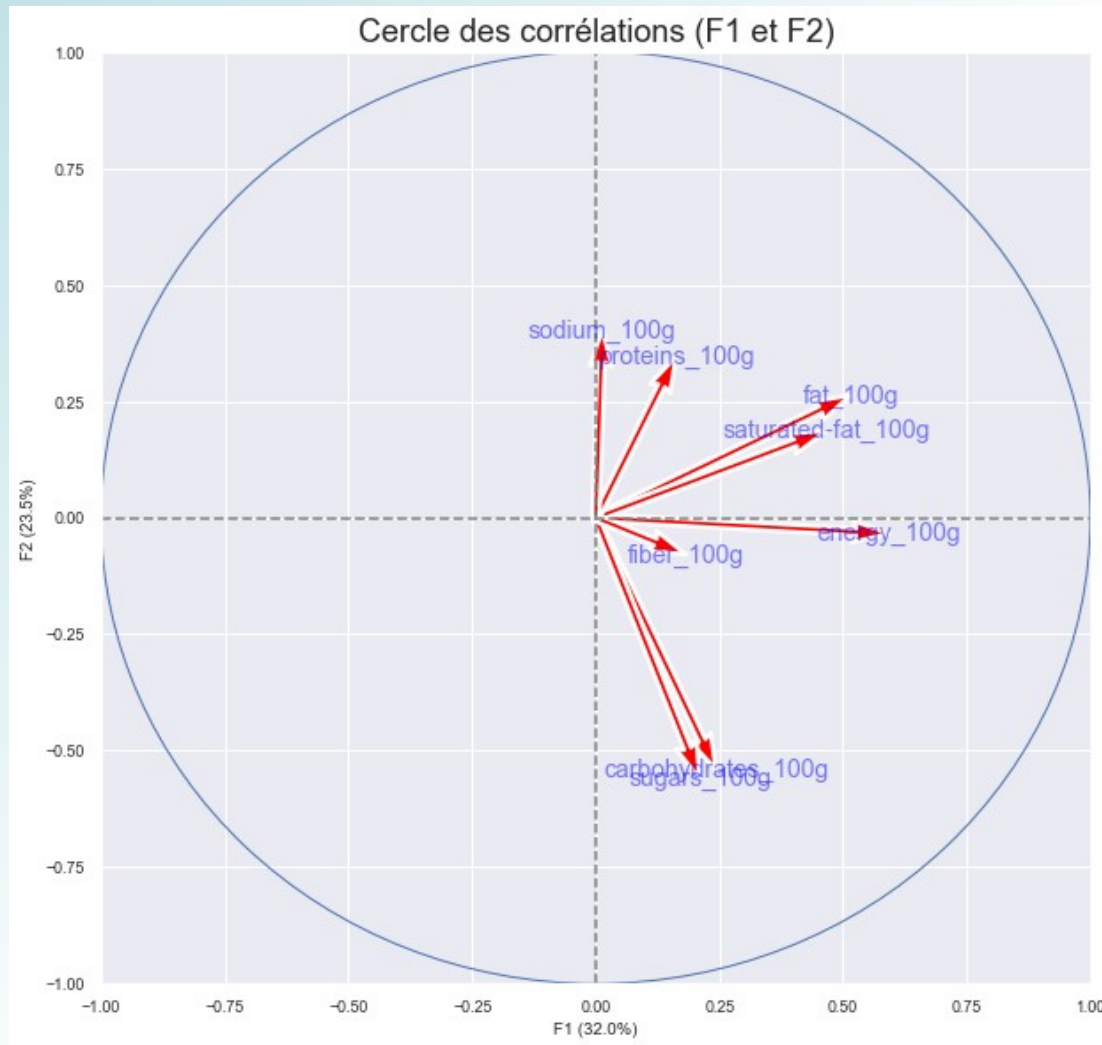
Analyse multivariée – ACP : réduction de dimensions



- Calcul d'un ACP sur les 8 variables nutriments et traçage du diagramme des éboulis.
- Il apparaît qu'il serait possible de réduire le dataset à 4 composantes principales (le 2 premiers plans factoriels représentent 83 % de la variance cumulée).
- La réduction de dimensions n'est pas pertinente car le nettoyage de données a permis de réduire fortement le nombre de variables (8 nutriments).

Présentation de l'analyse exploratoire:

Analyse multivariée – ACP : cercle des corrélations



Cercle des corrélations sur le 1^{er} plan factoriel (F1/F2).

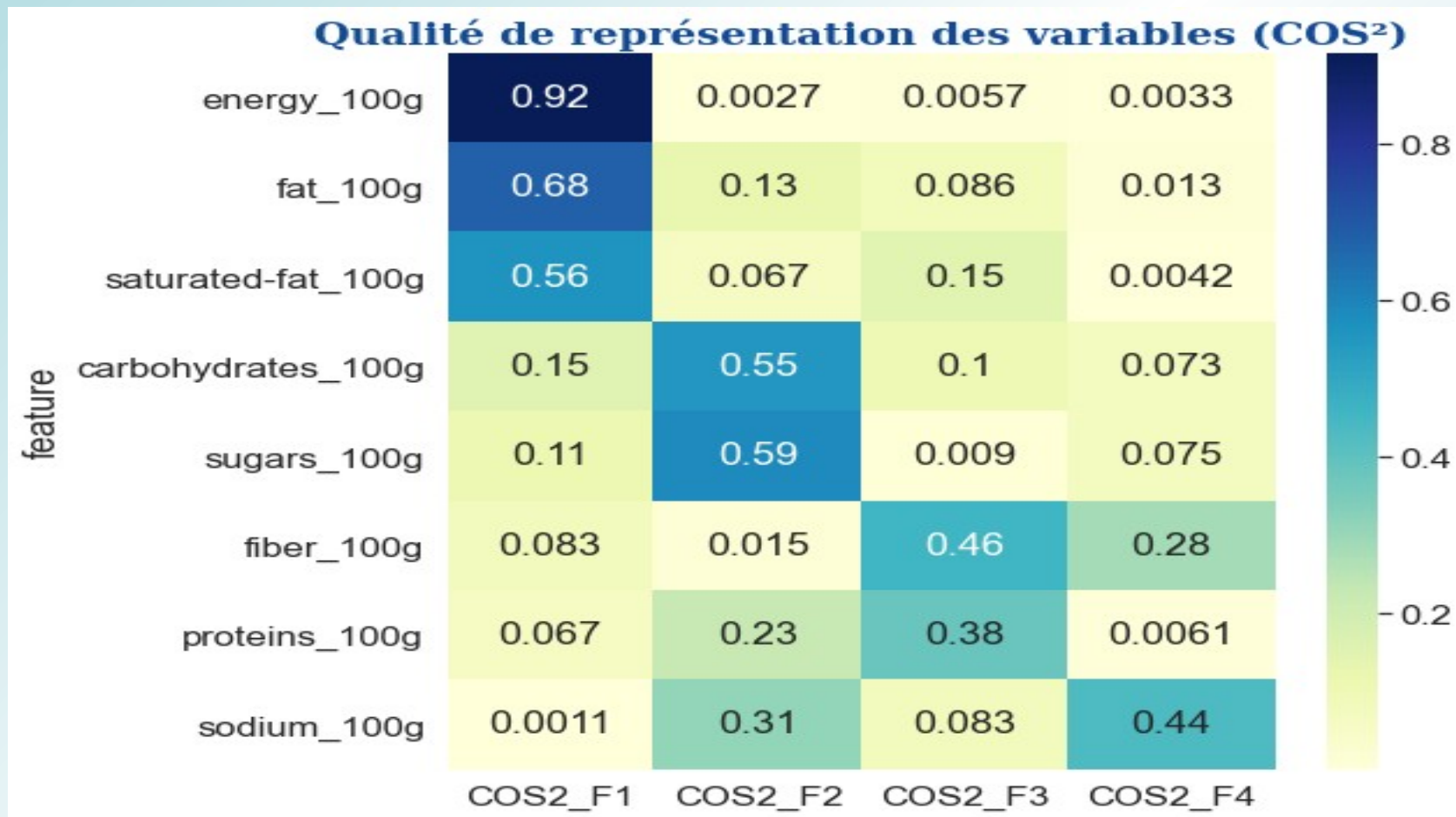
Présentation de l'analyse exploratoire:

Analyse multivariée – ACP : interprétation cercle corrélations

- Les variables « energy_100g », « fat_100g » et « saturated-fat_100g » sont corrélées positivement à l'axe F1.
- Les variables « fat_100g » et « saturated-fat_100g » sont fortement corrélées, les graisses saturées font partie de la famille des graisses.
- Sur l'axe F2, Les variables « sugars_100g », « carbohydrates_100g » sont fortement corrélées négativement, les sucres sont le maillon le plus simple des glucides .
- L'axe F1 pourrait représenter les sources des apports énergétiques, notamment les lipides.
- Difficulté pour interpréter l'axe F2, à ce stade de l'analyse.

Présentation de l'analyse exploratoire:

Analyse multivariée – Qualité représentation variables COS²



Analyse de la qualité de représentation des variables sur les axes des 2 premiers plans factoriels.

Présentation de l'analyse exploratoire:

Analyse multivariée – Interprétation COS²

- Les COS² confirment que l'axe F1 représente le facteur « apport en énergie », les variables « energy_100g » et les variables « fat_100g » et « saturated-fat_100g » permettent d'interpréter l'axe F1.
- L'axe F2 correspond à des produits sucrés (avec une forte teneur en sucres et glucides), les variables « sugars_100g » et « carbohydrates_100g » sont importantes pour interpréter l'axe F2. Ce sont des produits plutôt néfastes pour la santé.
- L'axe F3 correspond à des produits comportant des fibres et protéines, les variables « fiber_100g » et « proteins_100g » sont importantes pour interpréter l'axe F3 . Ce sont plutôt des produits bons pour la santé.
- L'axe F4 correspond à des produits comportant du sodium, la variable « sodium_100g » est importante pour interpréter l'axe F4. Ce sont plutôt des produits néfastes pour la santé.

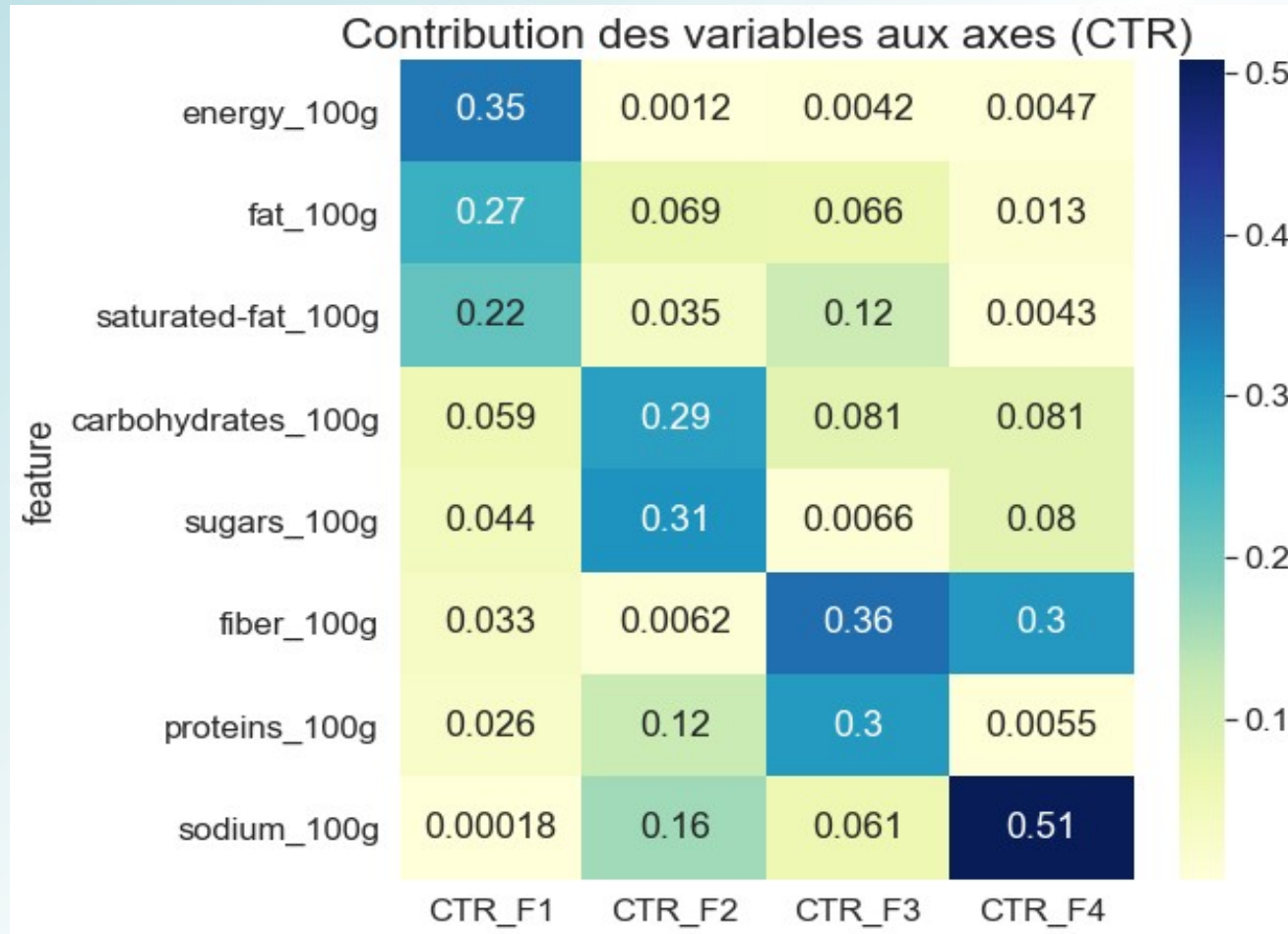
Présentation des faits pertinents pour le Nutri-rein:

1ère observation - Pertinence des variables nutriments retenues

- Les variables sur l'axe F1 et F2, à savoir l'énergie, les graisses et les graisses saturées (axe F1) et les glucides et les sucres (axe F2) sont les variables les plus importantes pour expliquer la variabilité dans le jeu de données.
 - Sur les axes F3 et F4, les fibres et les protéines (axe F3) et le sodium (axe F4) sont corrélées avec les axes F3 et F4
- => 1ère observation : les 8 variables nutriments retenues pour le calcul du Nutri-rein sont corrélées à un des axes des 2 premiers plans factoriels, donc sont des variables principales.**
- => Ces variables contribuent aux 4 composantes principales (qui représentent 83 % de la variance cumulée) et sont donc représentatives pour le calcul du Nutri-rein.**

Présentation des faits pertinents pour le Nutri-rein:

1ère observation - CTR contribution des variables aux axes



Contribution des variables aux axes (CTR)

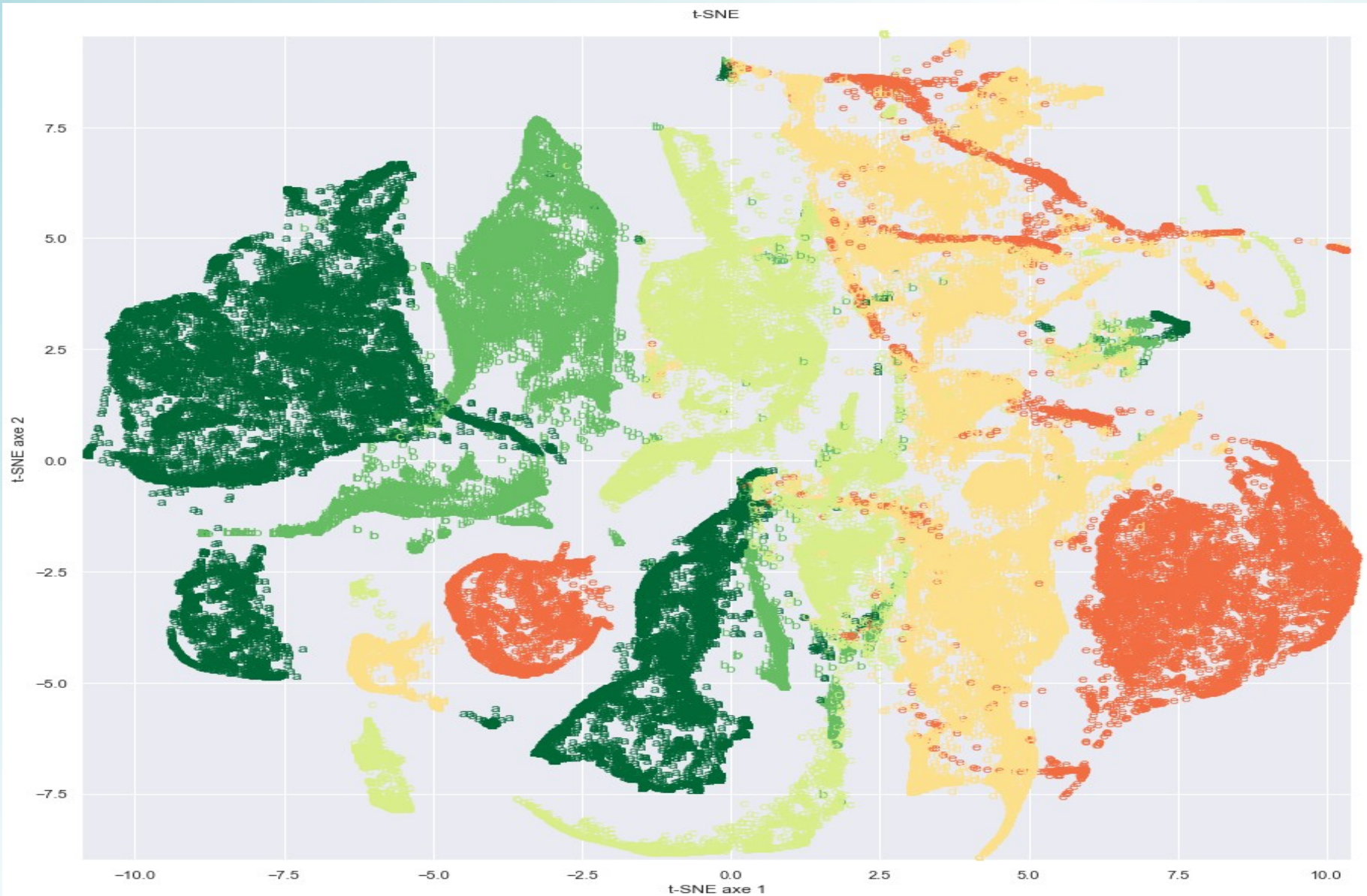
Présentation des faits pertinents pour le Nutri-rein:

2ème observation - Identification de groupes de produits

- Méthode retenue : réduction de dimensions par t-SNE et visualisation du nuage de points.
- Méthodes non retenues car inadéquates pour l'application :
 - Projection des produits dans l'espace ACP, suite ACP
=> cette méthode de visualisation n'a pas permis de visualiser « clairement » des groupes homogènes (voir détail dans les slides en annexes).
 - Algorithme de clustering K-means
=> cette méthode n'a pas permis d'identifier des groupes de produits homogènes selon leur qualité nutritionnelle.
=> la distribution des grades par cluster mélange des produits de qualité nutritionnelle disparates sur tous les grades (voir détail dans les slides en annexes).

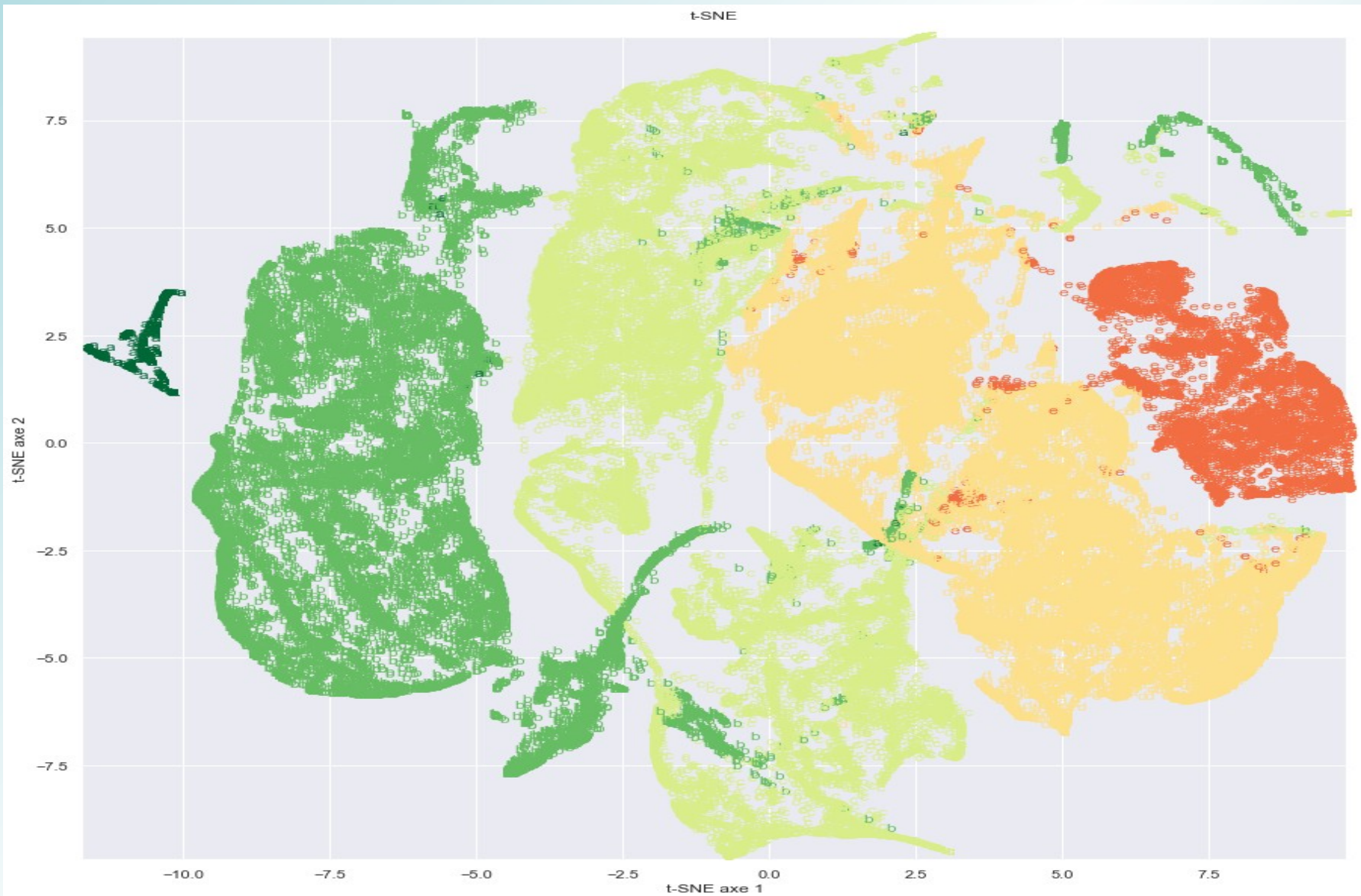
Présentation des faits pertinents pour le Nutri-rein:

2ème observation - Visualisation nuage points t-SNE (Nutri-score)



Présentation des faits pertinents pour le Nutri-rein:

2ème observation - Visualisation nuage points t-SNE (Nutri-rein)



Présentation des faits pertinents pour le Nutri-rein:

2ème observation - Comparaison t-SNE Nutri-score / Nutri-rein

- Pour le Nutri-rein, le t-SNE montre que les produits sont regroupés en groupes bien séparés correspondant aux différents grades du Nutri-rein.
- Par contre, pour le Nutri-score, le regroupement en groupes est moins évident
 - notamment pour les produits avec un grade 'a' ou 'e' qui sont disséminés dans plusieurs sous-groupes.

=> 2ème observation : le nutri-rein est plus pertinent que le nutri-score pour représenter une distribution cohérente des produits selon leur qualité nutritionnelle.

Présentation des faits pertinents pour le Nutri-rein:

3ème observation - *Distribution normale des grades Nutri-rein*

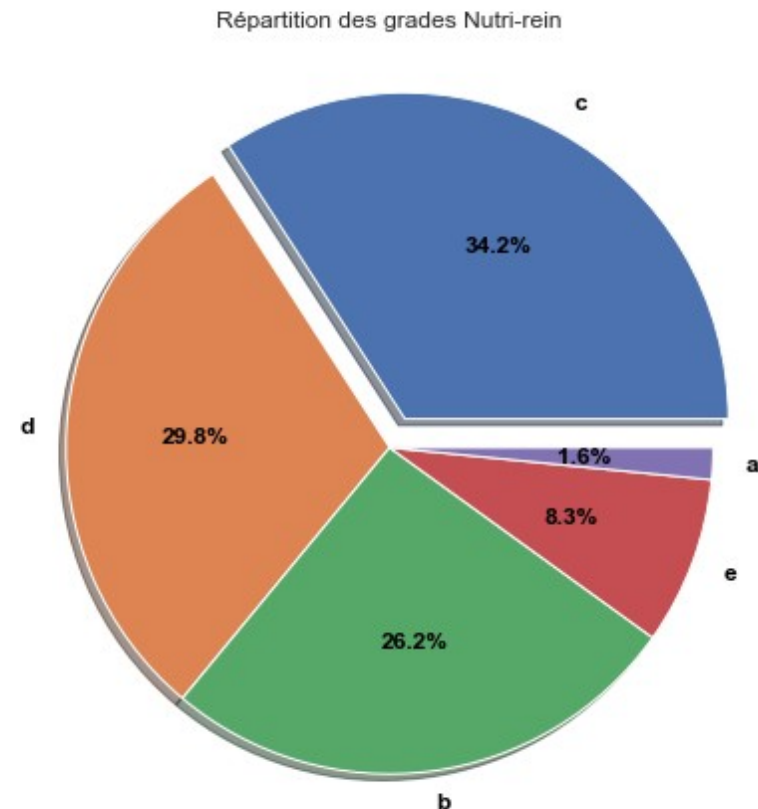
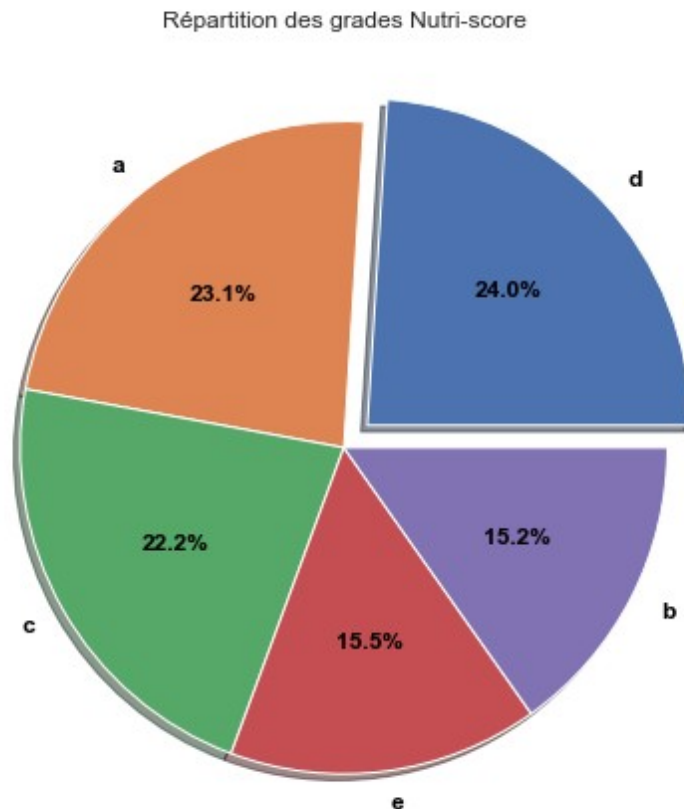
- La répartition des grades et scores Nutri-rein ressemble à une distribution normale
 - avec une prépondérance des produits avec de qualité nutritionnelle moyenne (grade 'c') et peu de produits de très bonne qualité (grade 'a') et de très mauvaise qualité nutritionnelle (grade 'e') .
- La répartition des grades et scores Nutri-rein tient compte des apports nutritionnels attendus dans le cadre d'un produit entrant dans la composition d'un repas.
- La distribution des scores Nutri-rein par catégorie de produit suit aussi une distribution ressemblant à une distribution normale.

=> 3ème observation : la répartition des grades et scores Nutri-rein est plus réaliste et « naturelle » que celle du Nutri-score pour évaluer la qualité nutritionnelle des produits.

Présentation des faits pertinents pour le Nutri-rein:

3ème observation - Comparaison répartition produits par grade

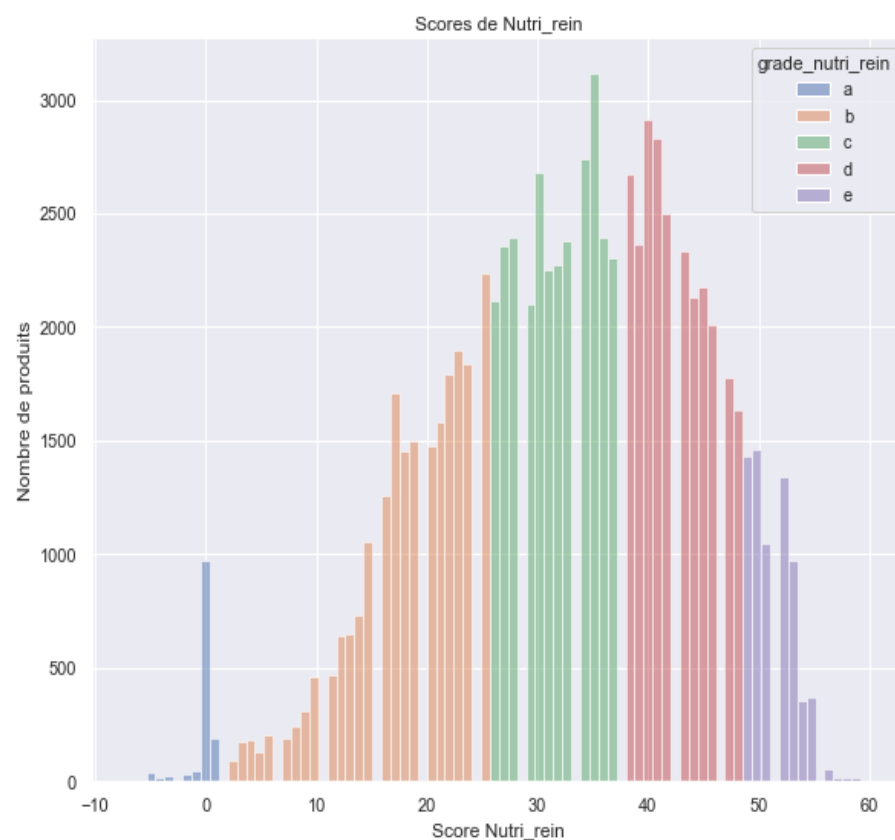
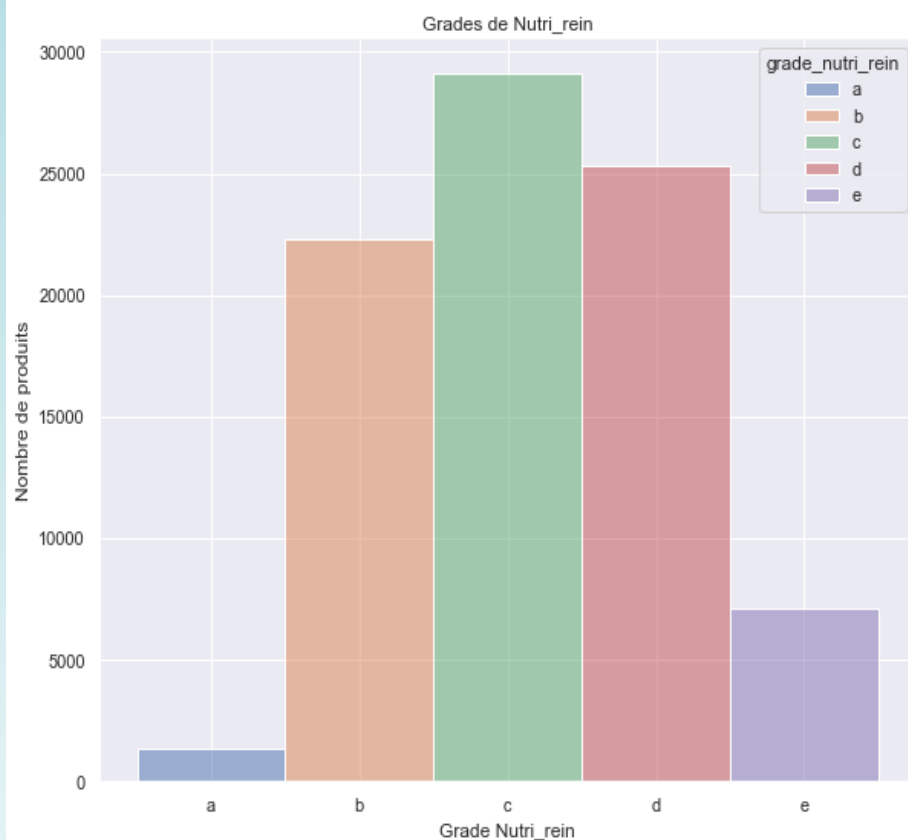
Répartition des produits par grade



Présentation des faits pertinents pour le Nutri-rein:

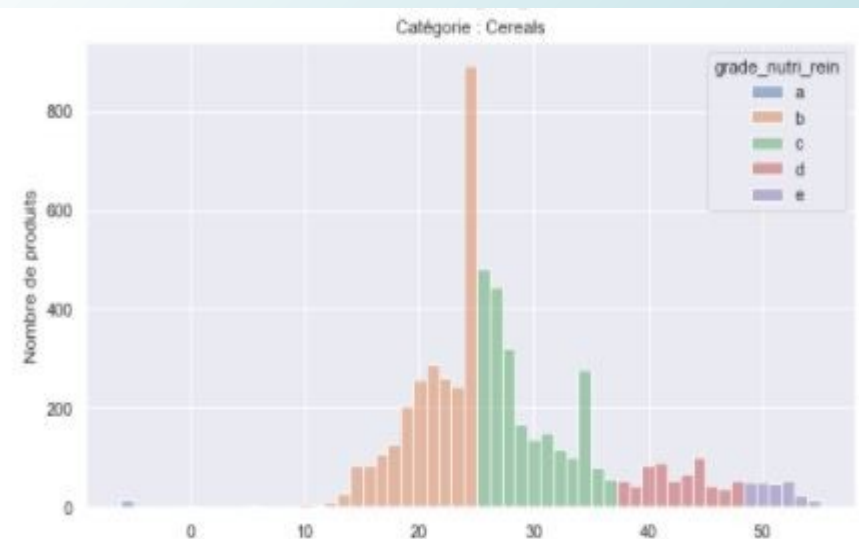
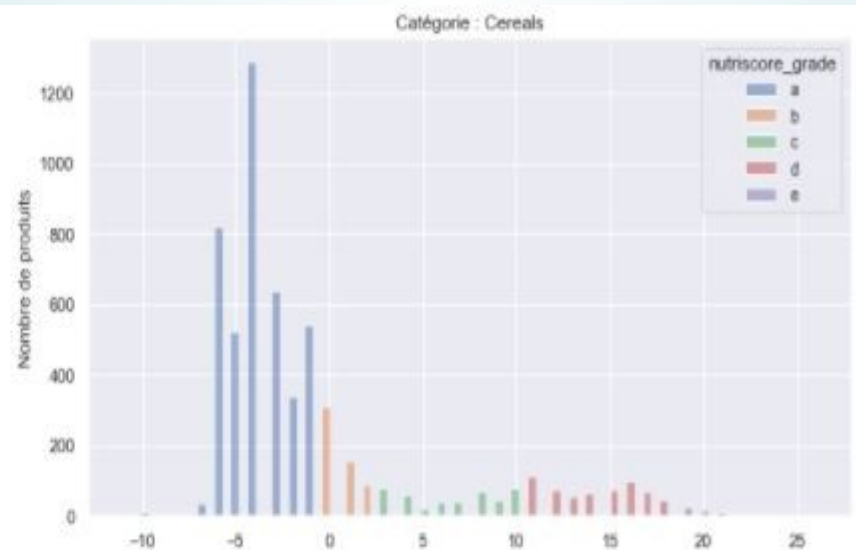
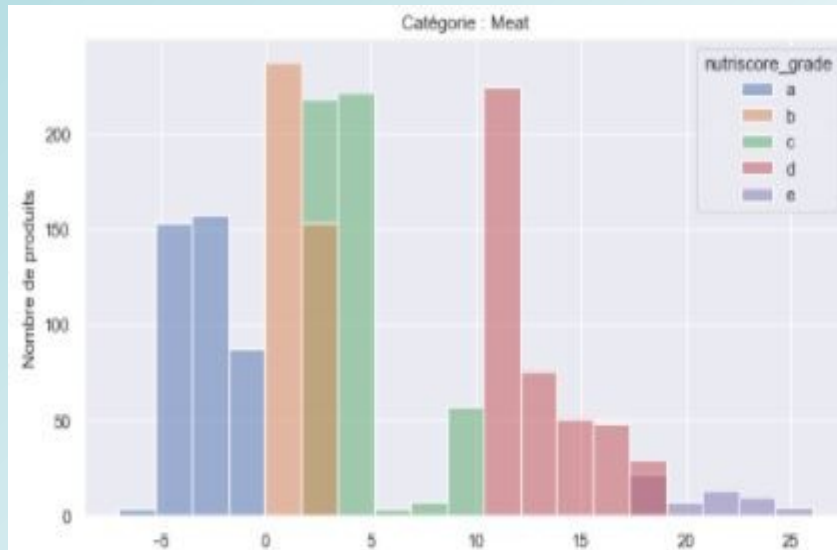
3ème observation - Répartition des grades/scores Nutri-rein

Répartition des scores Nutri-rein et de leurs grades



Présentation des faits pertinents pour le Nutri-rein:

3ème observation - Comparaison distribution scores par catégorie



Conclusion sur la faisabilité de l'application Nutri-rein

- Le Nutri-rein a pu être calculé sur la base uniquement de 8 nutriments (sodium, glucides, fibres,...) et les apports en énergie, représentatifs des besoins pour le patient en insuffisance rénale. Le calcul mobilise moins de variables que le Nutri-score.
- Le Nutri-rein est plus précis que le Nutri-score car il est calculé par rapport aux apports nutritionnels et énergétiques attendu d'un produit intégré comme élément d'un repas ou du petit-déjeuner.
- Les grades/scores Nutri-rein suivent une distribution plus « naturelle » que le Nutri-score, avec une majorité de produits de moyenne qualité nutritionnelle .

Axes d'amélioration de l'application Nutri-rein

- Prise en compte de variables sous-représentées dans la jeu de données : calcium, potassium, phosphore.
- Ajout de variables afin de rendre l'application plus précise : quantité consommée, poids, sexe, âge de la personne, niveau d'IRC
- Transformation du Nutri-rein en assistant numérique culinaire pour confectionner et évaluer la qualité nutritionnelle des repas de la journée (petit-déjeuner, repas midi, repas soir) d'un IRC
 - Ajout d'un classement des produits comme élément d'un repas (entrée, plat principal, dessert,...)
 - Pondération (à la baisse) du Nutri-rein dans la cas d'un dépassement de l'apport énergétique attendu

Annexes

Annexe 1 : Insuffisance rénale en chiffres

Annexe 2 : Besoins en nutriments / énergie d'un IRC

Annexe 3 : ACP – cercle corrélations axes F3 / F4

Annexe 4 : Projection des produits dans l'espace ACP

Annexe 5 : Classification K-means

Annexe 1 : L'insuffisance rénale en chiffres

- Un patient en insuffisance rénale chronique (IRC) est un patient dont le débit de filtration glomérulaire (DFG) est inférieur à 60 mL/min/1.73m². Il représente la quantité de liquide filtrée par les reins. Il se mesure grâce à la créatinine.
- Un patient est considéré en insuffisance modérée quand son DFG est en 30 et 40 mL/min/1.73m².
- Un patient est en insuffisance rénale sévère lorsque son DFG est entre 15 et 30 mL/min/1.73m².
- Une dialyse est envisagée lorsque le DFG est inférieur à 15 mL/min/1.73m².
- Le patient en insuffisance rénale doit limiter ses apports en sodium, en calcium, phosphores, potassium et en protéines essentiellement.

Annexe 2 : Besoins en nutriments / énergie d'un IRC

Apport énergétique : environ 146 kJ/kg/jour pour une personne exerçant une activité moyenne, soit environ un apport énergétique de $80 \times 146 = 11680$ kJ par jour pour une personne de 80kg. Ce chiffre est à majorer ou à minorer selon l'âge et l'activité de l'individu.

Pour info, l'Anses préconise un apport énergétique de 8400 kJ pour un femme adulte avec une activité moyenne, et de 11300 kJ pour un homme adulte avec une activité moyenne, sans pathologie rénale.

Apport en sodium (en g/jour): pas plus de 3g de sodium par jour

Pour info, l'OMS recommande un apport de moins de 5 g par jour pour une personne sans pathologie rénale.

Apport en protéines (en g/kg de poids corporel/jour): pas plus de 0,8g/kg/jour, soit environ un apport en protéines de $80 \times 0,8 = 64$ g par jour pour une personne de 80 kg.

Pour info, les hommes et les femmes sédentaires sans pathologie rénale ont besoin d'environ 0,8 g / kg. Un entraînement en endurance fait grimper cette quantité à 1,2 – 1,4 g / kg / jour. Un entraînement en force la fait passer à 1,2 – 1,8 g / kg / jour.

Apport en fibres (en g/jour): au moins 30g/jour avec plus de 50% d'aliments végétaux.

Pour info, l'OMS recommande un apport en fibres journalier entre 25 g et 29 g jour.

Apport en glucides (exprimé en % de l'apport énergétique total journalier): pas plus de 50% de l'apport énergétique total journalier).

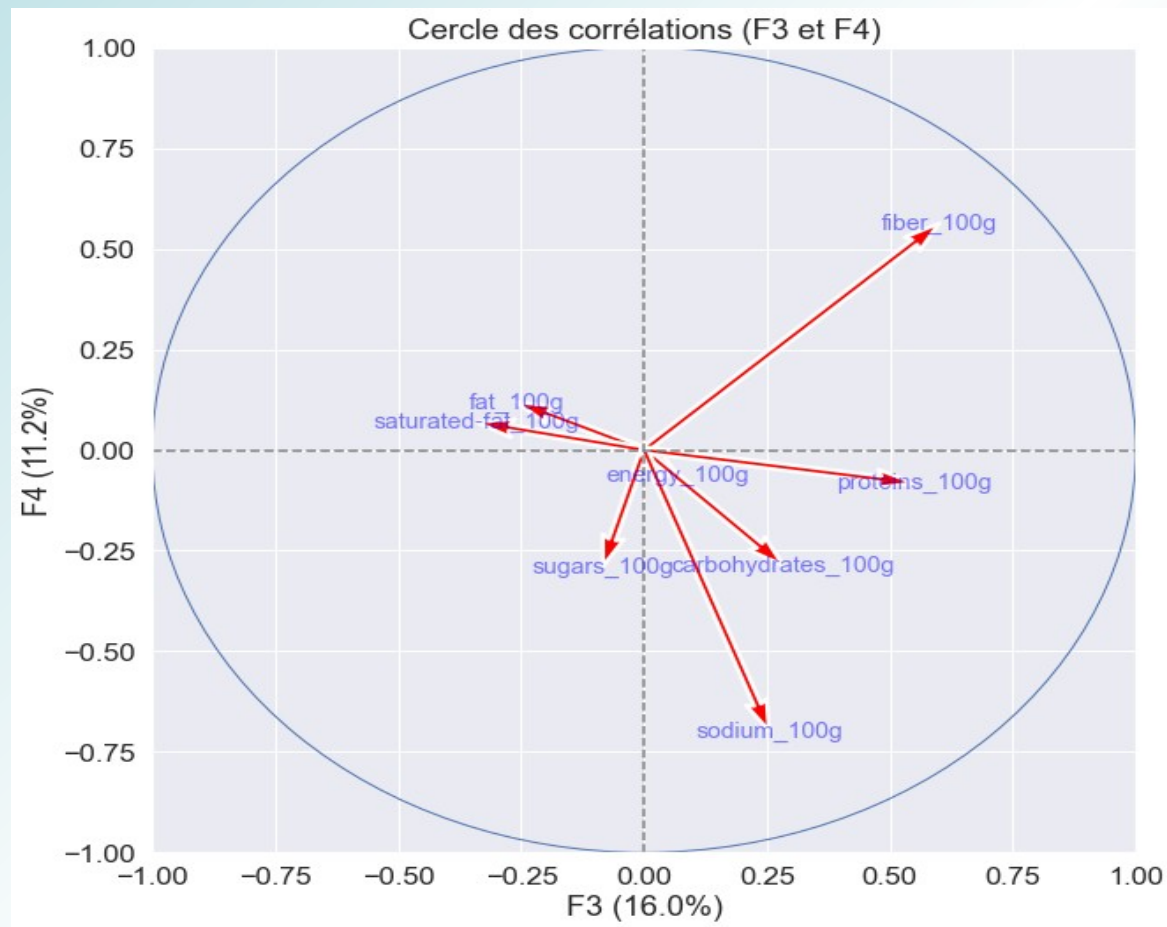
Par exemple, pour une personne de 80kg, on ne dépassera pas 5840 kJ par jour, soit 308 g de glucides par jour. Pour info, 1g de glucides = 17 kJ.

Apport en lipides (graisses dont graisses saturées) (exprimé en % de l'apport énergétique total journalier) pas plus de 40% de l'apport énergétique total journalier, dont au moins 50 % de graisses insaturées.

Par exemple, pour une personne de 80 kg, on ne dépassera pas 4672 kJ par jour, soit 90g de lipides par jour. Pour info, 1 g de lipides = 37 kJ.

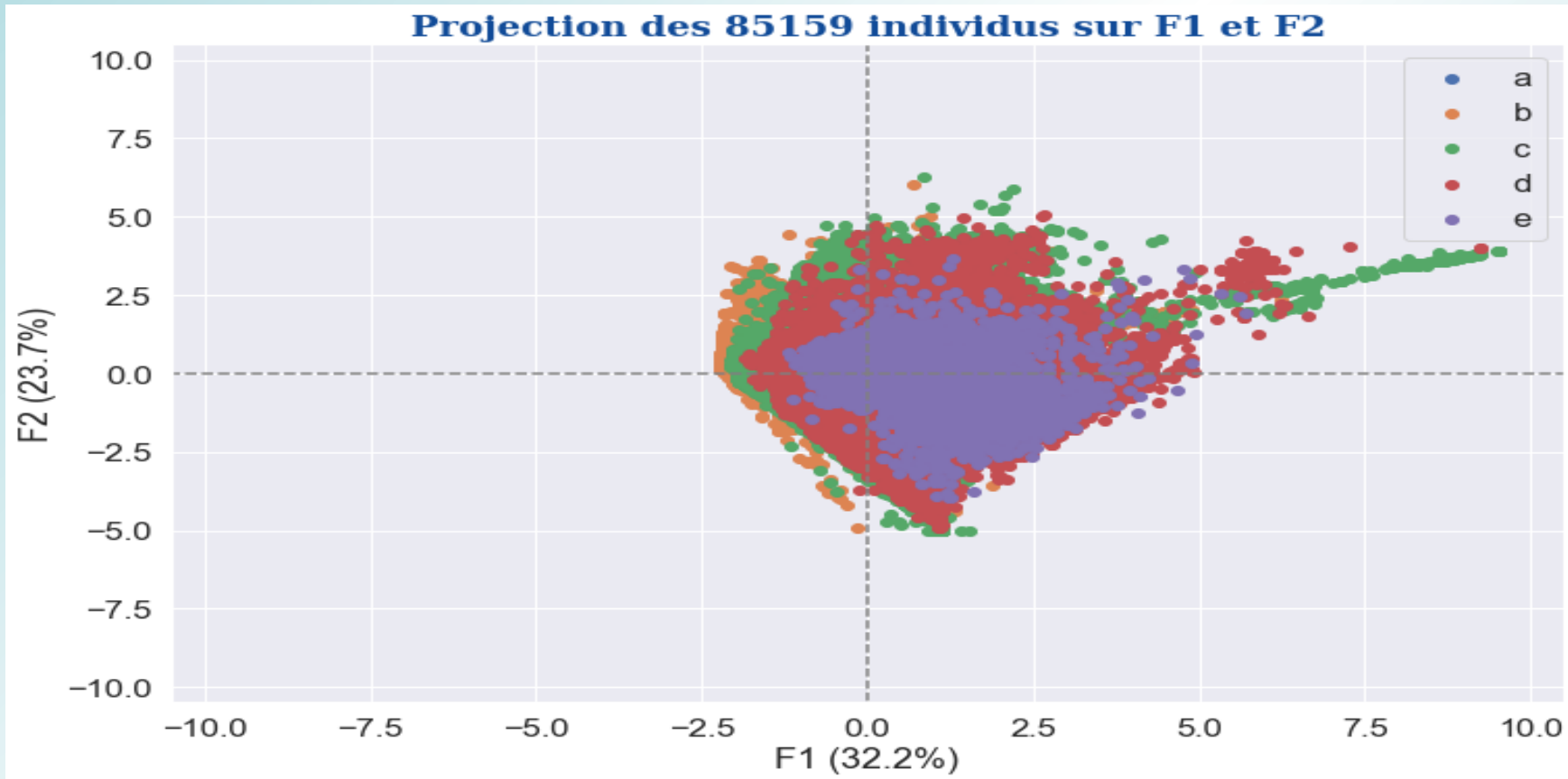
Apport en sucres (en g/jour): pas plus de 25 g par jour.

Annexe 3 : ACP – Cercle des corrélations axes F3 / F4



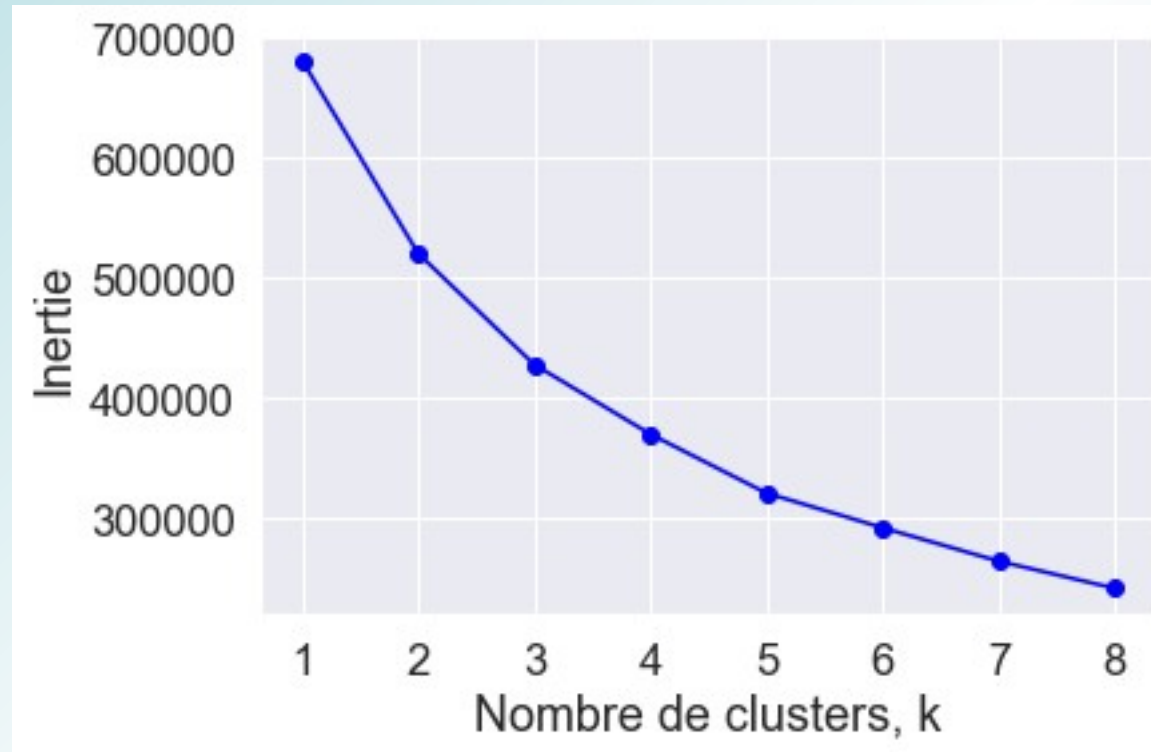
L'interprétation de la corrélation des variables sur les axes F3 / F4 est difficile. Analyse de la qualité de représentation des variables aux axes (COS^2 et CTR) pour affiner l'analyse.

Annexe 4 : Projection des produits dans espace ACP



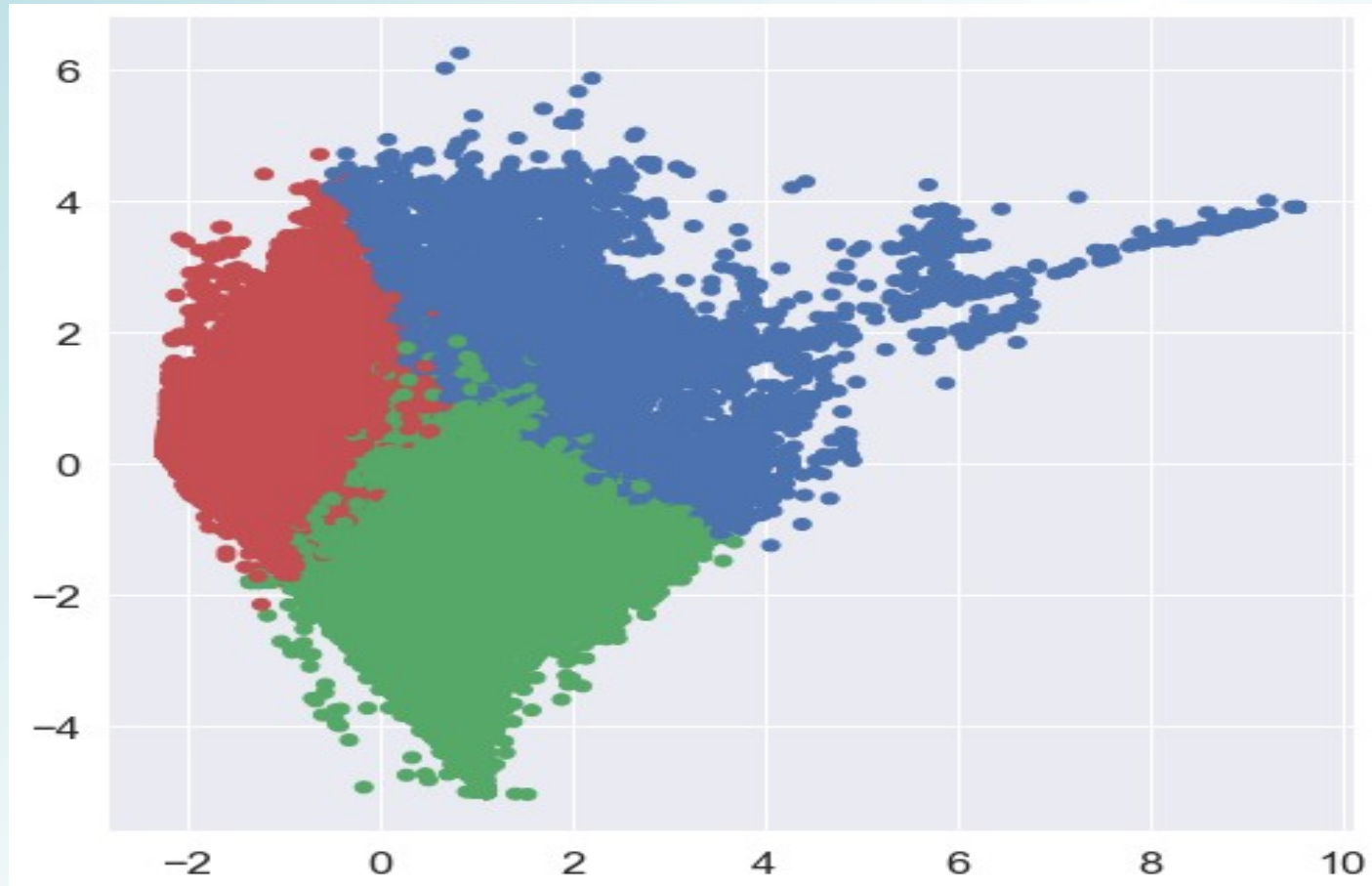
La projection des produits dans l'espace ACP (PCA) ne montre pas de classification "naturelle" des produits selon leur qualité nutritionnelle en fonction du grade nutri-rein. Idem pour la projection sur axe F3/F4.

Annexe 5 : K-means – Détermination du « coude »



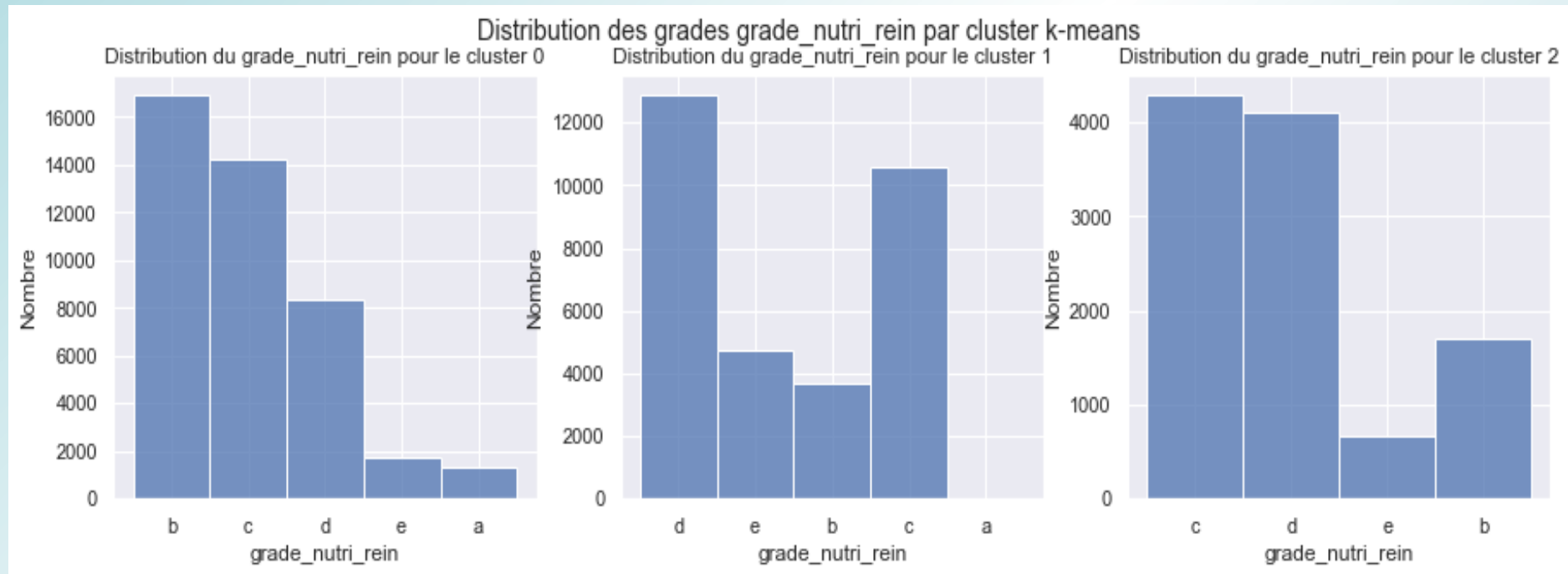
Sur le graphique ci-dessus, la détermination du nombre de clusters n'est pas évidente, il n'y a pas de "coude" visible. La valeur 3 semble la plus pertinente.

Annexe 5 : K-means – Classification avec $k = 3$



L'affichage des points sur les 3 clusters ne permet pas en l'état de valider la classification proposée par k-means. Il n'y a pas de relation évidente avec les grades Nutri-rein.

Annexe 5 : K-means – Distribution Nutri-rein par cluster



Pour les grades du nutri-rein, le grade 'a' n'est pas présent dans le cluster 1, et les clusters ne permettent pas de classer les produits selon leur qualité nutritionnelle, puisque tous les grades ou presque sont présents dans les 3 clusters.