

• Pour info, sur la slide 11, il y a une animation qui cache le texte en dessous donc si possible plutôt regarder le ppt en mode diaporama pour voir le texte sur cette slide.

Appel à projets : idées innovantes d'application en lien avec l'alimentation





- 1. Rappel de l'appel à projets
- 2. Explication de l'idée d'application
- 3. Démarche méthodologique de nettoyage des données
- 4. Démarche méthodologique d'exploration des données
- 5. Conclusion



Rappel de l'appel à projets

Lancé par Santé publique France

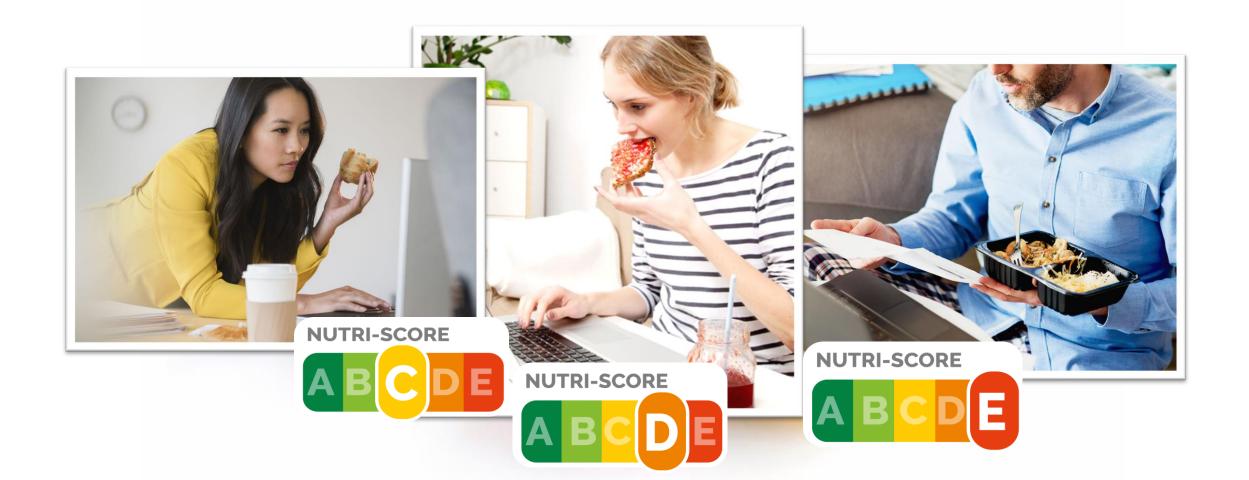


• Trouver des idées innovantes d'applications en lien avec l'alimentation

Utilisation du jeu de données Open Food Facts



Télétravail et alimentation





Impact du télétravail sur l'alimentation

• Le stress et la proximité avec le réfrigérateur amène les télétravailleurs à grignoter et à casser le rythme de leur repas

• Conséquences : déséquilibre alimentaire, prise de poids, fatigue, etc.

Maladies liées à l'alimentation : l'alerte

Cancers: 1 re cause de mortalité en France.

1/3 des cas seraie

1/3 des cas seraient évitables par une alimentation plus équilibrée.

Maladies cardio-vasculaires:



2^e cause de mortalité en France.

Les liens de causalité en cas de surconsommation



Recommandation/consommation: le grand écart



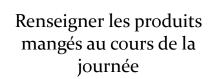






AMELIONER EQUILIBRE ALIMENTAISE

Principe de l'application :



Soit en scannant directement le code barre du produit Soit en cherchant le produit en utilisant la barre de recherche Indiquer la quantité consommée

Par défaut : portion simple mais possibilité de la modifier si besoin. Calcul automatique des indicateurs de la journée et comparaison aux recommandations

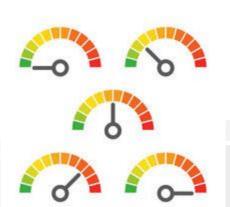
Calories:

1 800 kcal/jour femme 2 400 kcal/jour homme Acides gras saturés : 27 g/jour

Sucre : 50 g/jour

Sel : 5 g/jour

Fibres : 25 g/jour





Pour chaque indicateur en dehors des limites :

Liste des 3 produits les plus contributeurs dans l'indicateur

Pour chacun des 3 produits :

→ Proposition des 5 produits de la **même catégorie** avec le **meilleur nutriscore** et la meilleure valeur pour l'indicateur concerné

Choix des produits de "remplacement" puis recalcul automatique des indicateurs de la journée et comparaison aux recommandations





 $1 \rightarrow 18,7 g$







2 → 11,7 g





 $3 \rightarrow 7.9 \mathrm{g}$





 $1 \rightarrow 2,1 g$





 $2 \rightarrow 14,4 g$







 $3 \rightarrow 16,2 g$





4 > 15,3 g







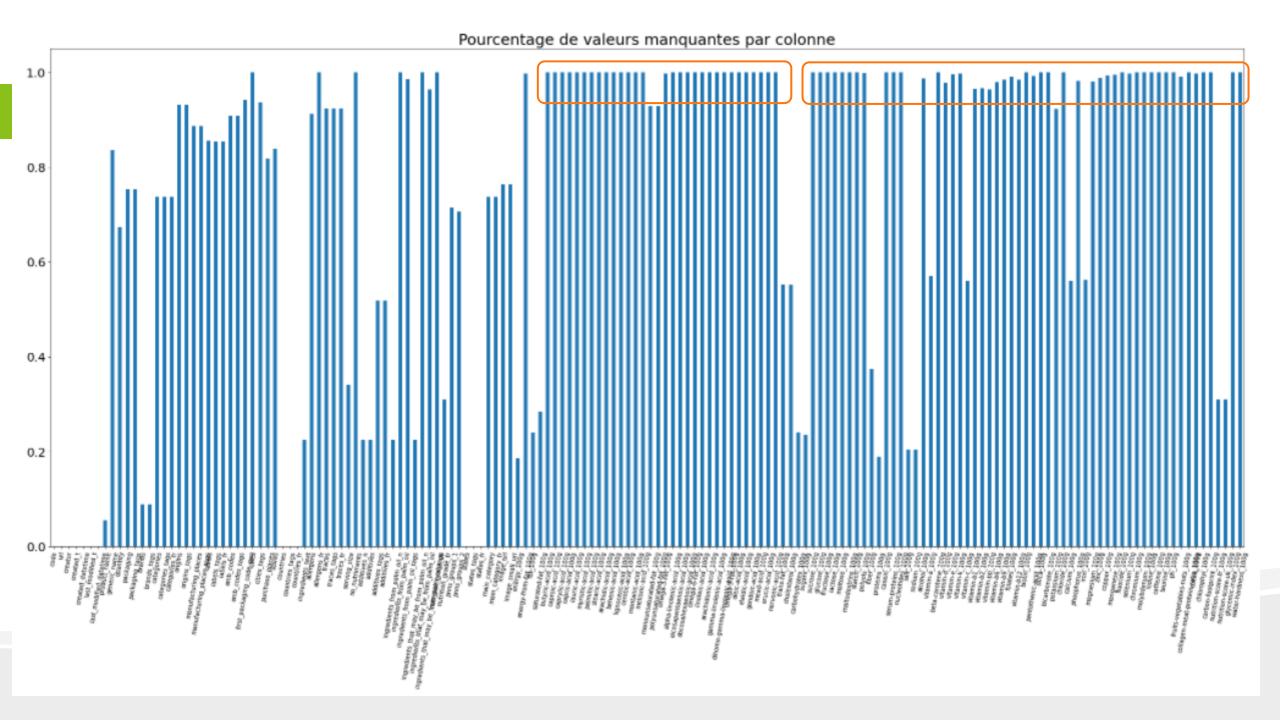
 $5 \rightarrow 18 g$



Donner la possibilité de sauvegarder les choix de remplacement pour garder une liste de meilleurs produits en mémoire







Nettoyage du jeu de données (2/4)

- Filtre sur les produits commercialisés en France
- Sélection des colonnes nécessaires à l'application :

Nom du produit

Sel pour 100g

Marque

Acides gras saturés pour 100g

Groupe PNNS 1 et 2

Fibres pour 100g

Calories pour 100g

Nutriscore

- Sucre pour 100g

Portion pour 1 personne



Nettoyage du jeu de données (3/4)

• Traitement des valeurs aberrantes :

- Somme de nutriments pour 100 g inférieure à 105 g (marge d'erreur possible)
- Limite du nombre de calories maximum à 3 770 kJ (valeur énergétique maximum d'un aliment)
- Limite entre o et 100 de la valeur d'un nutriment pour 100g
- Homogénéisation des catégories des groupes PNNS 1 et 2



Nettoyage du jeu de données (4/4)

• Traitement des valeurs manquantes :

- Suppression des lignes ayant au moins 2 valeurs manquantes sur l'ensemble des colonnes suivantes : calories, sucre, sel, acides gras saturés et fibres.
- Pour la colonne groupe PNNS 1 : utilisation d'un k-NN Classifier.
- Pour les autres colonnes : utilisation d'un k-NN Imputer.



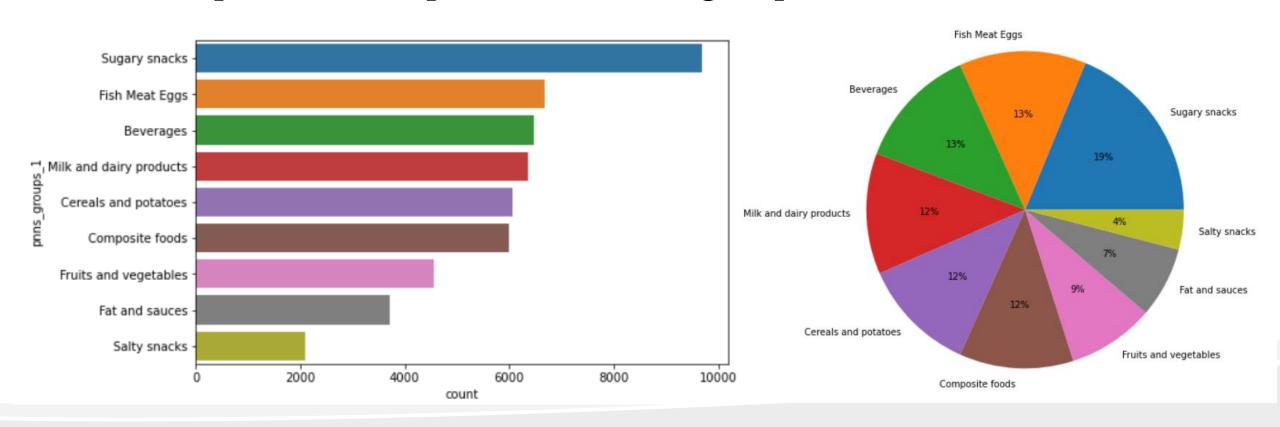
Analyse exploratoire du jeu de données (1/9)

• Après avoir effectué ces traitements on obtient un jeu de données de 61 394 lignes et 12 colonnes rempli à 100%.

• Cependant pour l'analyse exploratoire nous allons utiliser le jeu de données sans le traitement des valeurs manquantes.

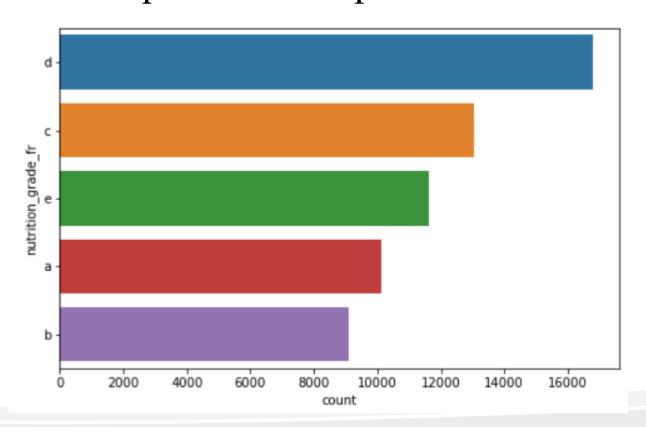
Analyse exploratoire du jeu de données (2/9)

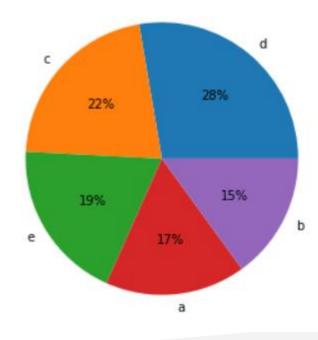
• Répartition des produits selon le groupe PNNS 1 :



Analyse exploratoire du jeu de données (3/9)

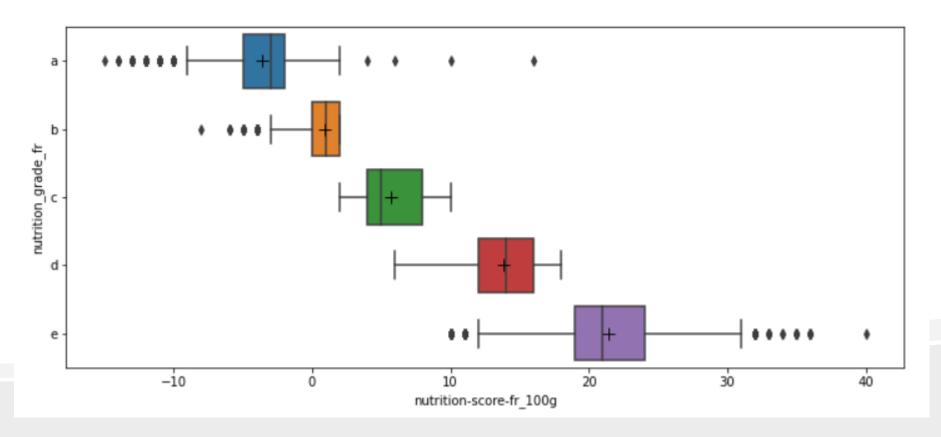
• Répartition des produits selon le nutriscore :





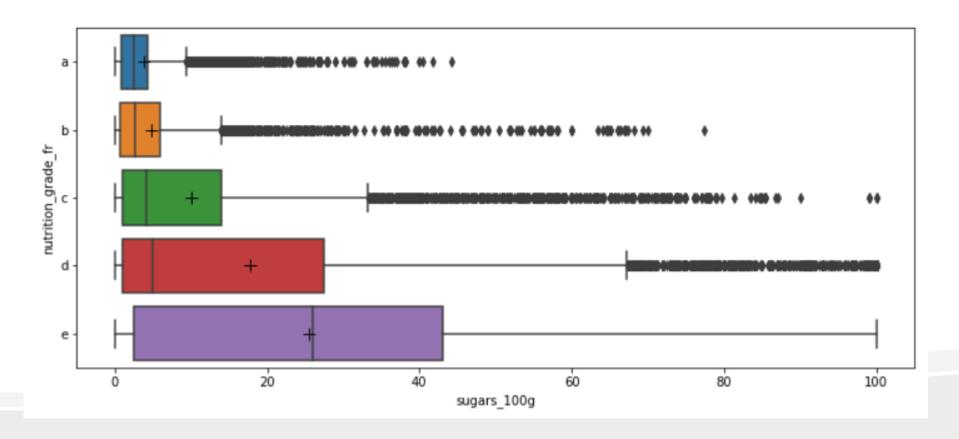
Analyse exploratoire du jeu de données (4/9)

• Lien entre la valeur numérique du nutriscore et la catégorie associée :



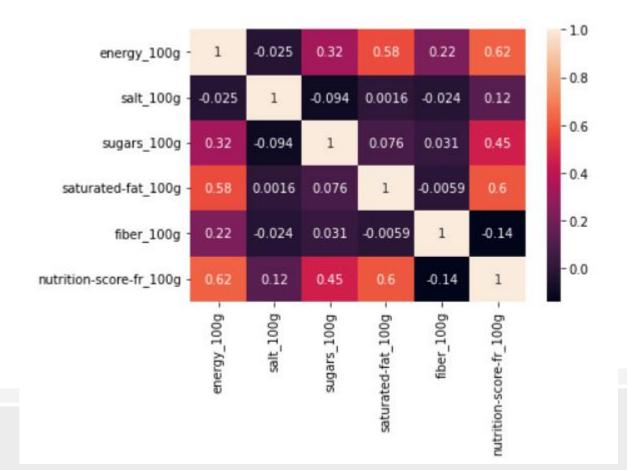
Analyse exploratoire du jeu de données (5/9)

• Analyse de l'impact du sucre sur le nutriscore :



Analyse exploratoire du jeu de données (6/9)

• Matrice de corrélation :



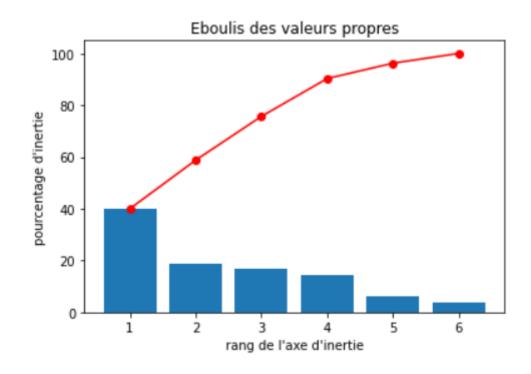


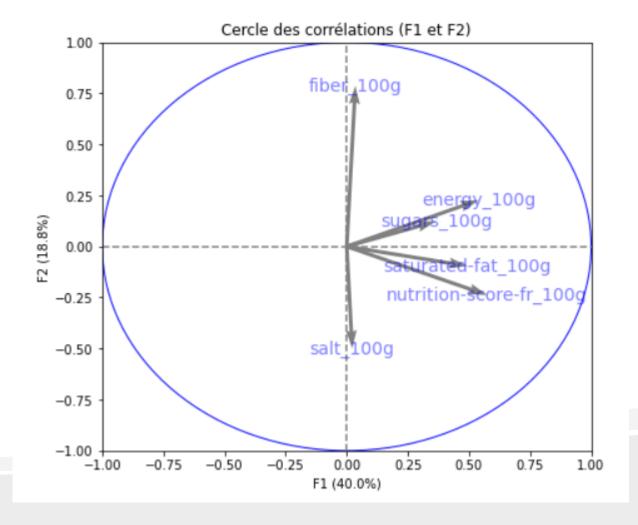
Analyse exploratoire du jeu de données (7/9)

- Réalisation d'une Analyse en Composantes Principales (ACP) sur les variables suivantes :
 - Calories pour 100 g
 - Sucre pour 100 g
 - Sel pour 100 g
 - Acides gras saturés pour 100 g
 - Fibres pour 100 g
 - Valeur numérique du nutriscore

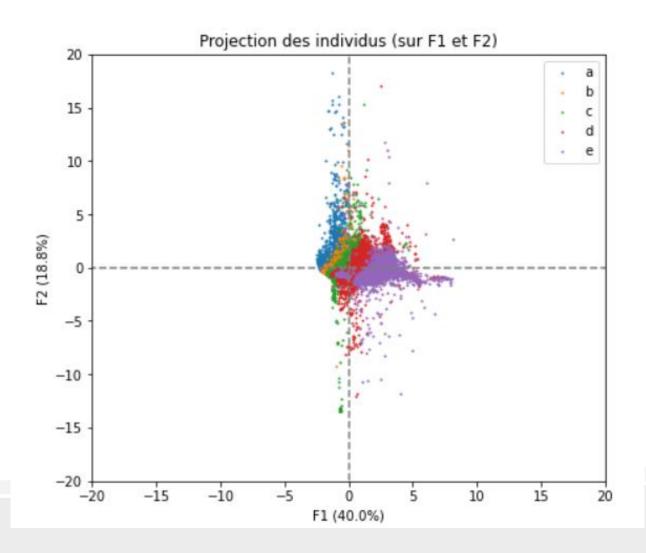
Analyse exploratoire du jeu de données (8/9)

• Résultats :









- On retrouve bien des groupes d'individus par rapport à F1 :
 - Plus on augmente sur F1 et plus la catégorie du nutriscore est mauvaise



Conclusion

- Il est pertinent pour l'application de proposer des produits avec un meilleur nutriscore car il y a une corrélation entre le nutriscore et le sucre, les calories et les acides gras saturés.
- Plus de 25 000 produits utilisables directement dans l'application.
- Plus de 61 000 après un traitement préalable.