

Orange is the new black

Analyse sur les jus d'orange vendus en France



Sommaire

01 Data

Open Food Facts et leurs données open-data

02 Cleaning

Nettoyage des données

03 Analyse

Description avec visualisation

04 Idée d'app

ou non ?



01

Data



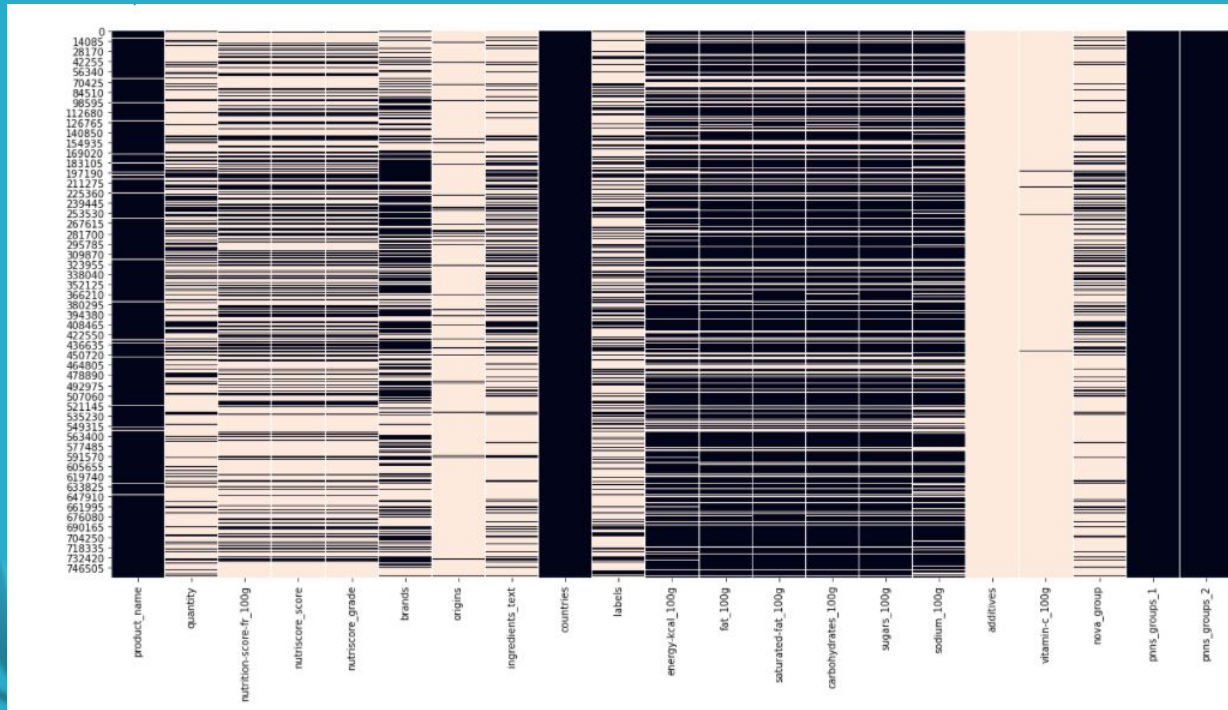
Whoa!

Open Food Facts est une base de données sur les produits alimentaires par tout le monde et pour tout le monde.

Les données sont open-data, elles sont ainsi utilisables pour tout usage.



En résumé, Open Food Facts





02

Cleaning

Les étapes du cleaning :

01



JUS D'ORANGE

Isoler notre cible, le jus d'orange, dans 'product_name'

02



FRANCE

Convertir et conserver l'information FRANCE dans 'countries'

03



21 COLONNES

Conserver au moins 60% ou plus de données valides

04



NaN

Remplacer les NaN par des zéros



Nous avons désormais ainsi un total de 516
lignes et 21 colonnes exploitables pour
notre analyse.

03

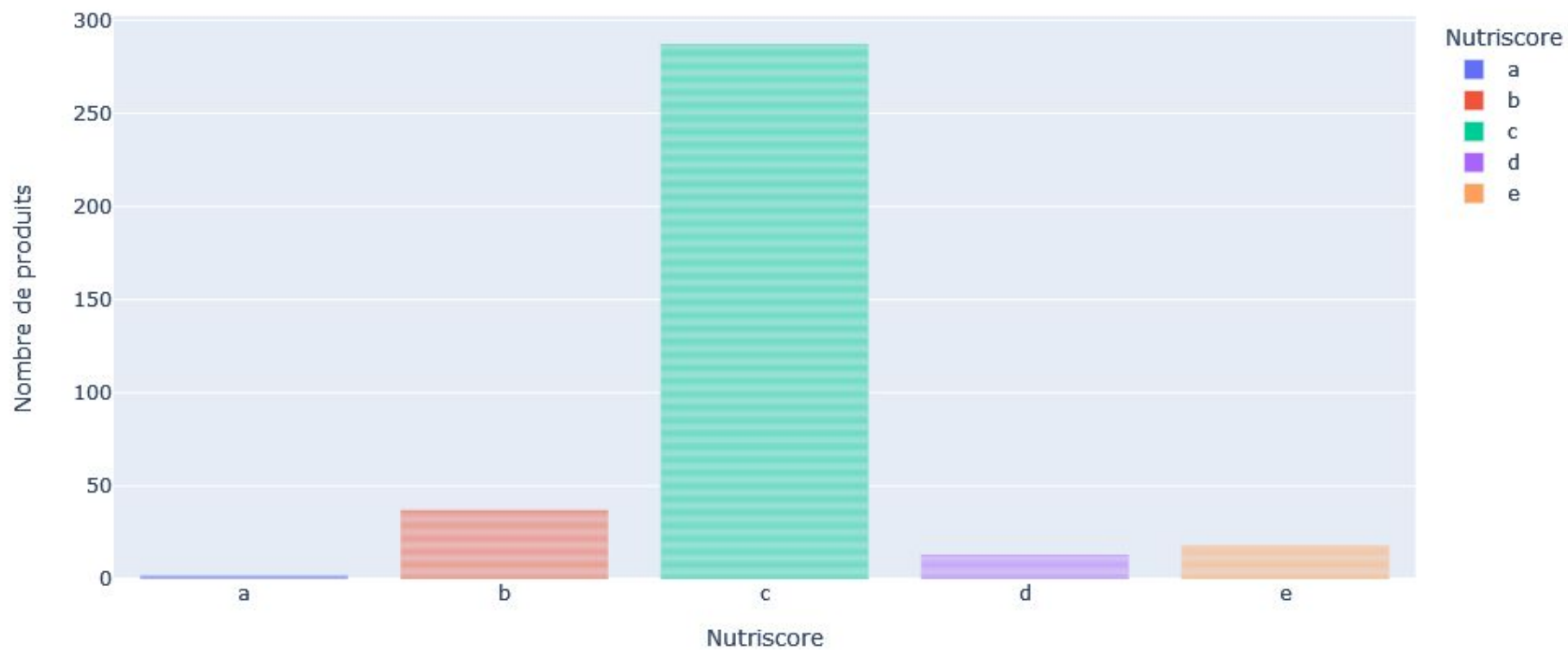
Analyse



Répartition des produits avec taux de sucre entre 0,0001 et 50 grammes



Répartition des jus d'orange par Nutriscore



Les marques les plus représentées



Note : toutes les marques les plus représentées sont des marques distributeurs

Les jus sans conservateurs, colorants et sucre ajouté



Pur jus orange bio de Les Fées Bio, 100% pur jus d'oranges pressées d'Andros, Pur jus d'orange sans pulpe de U Bio, Pressade pur jus orange de Pressade, Le pur jus orange du matin de Joker

Les jus contenant le plus de vitamine C



Échantillon de 5 jus d'orange possédant entre 0,0420 et 0,0450 grammes de vitamine C pour 100 gr : 100 % pur jus orange de Franprix, Premium 100% jus d'orange sanguine de Monoprix, Le Pur jus - Sans pulpe Jus d'orange de Joker, Pur jus de fruit pressé 100% orange de Innocent, Pur jus d'oranges sanguines pressées de Super U

Les jus avec un nutriscore E



Les jus d'orange qui possèdent le nutriscore E sont majoritairement fabriqués à partir de jus concentré : Nectar d'orange de Casino, Jus d'orange de Kas, Jus d'orange de Fruima, Nectar d'orange de Tous les jours



04

Une app ?



Non, désolée

Il n'y aura pas d'applications, même si
le jus d'orange c'est super bon.

Trop de valeurs erronées dans les données open-data d'Open Food Facts

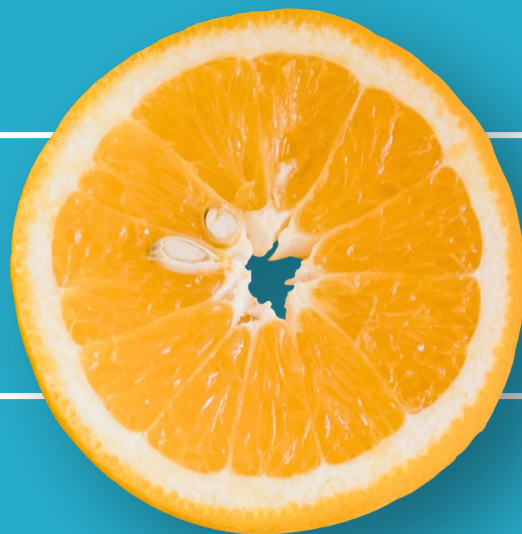
Exemple : ce jus d'orange annoncé avec un nutriscore A dans les données mais qui possède en réalité un nutriscore B après vérification sur le site web du distributeur.



Et trop de valeurs manquantes

1 969 352 lignes

186 colonnes



292 300 424
valeurs
manquantes

soit **80%**



Merci !