



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Universidad Nacional de Colombia - sede Bogotá
Facultad de Ingeniería
Departamento de Ingeniería de Sistemas e Industrial
Curso: Sistemas Operativos.

Estudiante: Juan Mateo Rozo Torres

Descripción detallada de los campos del dataset seleccionado

Este conjunto de datos proviene de la plataforma colaborativa Open Food Facts, y contiene información detallada sobre millones de productos alimenticios comercializados en todo el mundo. Está organizado en secciones temáticas que permiten analizar cada producto desde distintas perspectivas: identificación, ingredientes, valores nutricionales, etiquetado, impacto ambiental y otros aspectos relevantes.

Descripción general

En su conjunto, el dataset ofrece una visión integral de cada producto alimenticio, combinando información técnica, comercial, nutricional y ambiental. Su estructura facilita tanto el análisis estadístico como el desarrollo de herramientas para la toma de decisiones alimentarias, estudios de salud pública o aplicaciones orientadas al consumo consciente.

- **Total de columnas (campos): 209**
- **Total de filas (productos): 3.881.332**

Información general

Incluye los datos básicos que identifican cada producto, como su código de barras, nombre comercial, cantidad, fechas de creación y modificación, el usuario que aportó la información, y el enlace al producto en la plataforma.

- **Campos incluidos: 34**

Etiquetas

Esta sección contiene metadatos asociados al producto, como marcas, categorías, etiquetas de empaque o nutrición, países donde se vende, lugares de fabricación, y otros elementos relacionados con su distribución y comercialización. Es especialmente útil para el análisis por criterios geográficos, logísticos o de mercado.

- **Campos incluidos: 38**

Ingredientes

Describe la lista de ingredientes del producto en texto libre, e incluye información adicional sobre posibles alérgenos, trazas o componentes secundarios. Es clave para personas con restricciones alimentarias o alergias específicas.

- **Campos incluidos: 18**

Datos misceláneos

Agrupar información complementaria como el tamaño de la porción, número de aditivos, presencia de aceite de palma, categoría principal del producto, y enlaces a imágenes del envase. Son datos diversos que no encajan estrictamente en las categorías nutricionales, pero enriquecen el análisis general del producto.

- **Campos incluidos: 28**

Información nutricional

Ofrece una descripción detallada de la composición del producto por cada 100 gramos. Incluye valores energéticos, macronutrientes (grasas, carbohidratos, proteínas), tipos de ácidos grasos, azúcares, fibra, alcohol, cafeína, vitaminas y minerales. Es la sección más extensa y técnica del conjunto de datos.

- **Campos incluidos: 77**

Huella de carbono

Registra, cuando está disponible, la huella de carbono del producto según la información en el empaque. Permite evaluar el impacto ambiental del producto en términos de emisiones.

- **Campos incluidos: 2**

Nutri-Score y otros indicadores

Incluye calificaciones nutricionales como el Nutri-Score (sistema francés), el Nutrition Score UK, el grupo NOVA (procesamiento de alimentos), y puntajes ambientales. Estas métricas permiten comparar de forma visual y rápida la calidad nutricional y sostenibilidad de los productos.

- **Campos incluidos: 12**

Detailed Description of the Dataset Fields

This dataset comes from the collaborative platform Open Food Facts and contains detailed information on millions of food products marketed worldwide. It is organized into thematic sections that allow products to be analyzed from different perspectives: identification, ingredients, nutritional values, labeling, environmental impact, and more.

General Overview

Overall, the dataset offers a comprehensive view of each food product, combining technical, commercial, nutritional, and environmental data. Its structure facilitates both statistical analysis and the development of tools for dietary decision-making, public health studies, or applications aimed at informed consumption.

- **Total columns (fields): 209**
- **Total rows (products): 3,881,332**

General Information

Includes the basic data that identifies each product, such as barcode, product name, quantity, creation and modification dates, the user who contributed the data, and the URL of the product on the platform.

- **Fields included: 34**

Tags

This section contains metadata associated with the product, such as brands, categories, packaging or nutrition labels, countries of sale, manufacturing locations, and other elements related to distribution and commercialization. It is especially useful for analysis by geographic, logistic, or market criteria.

- **Fields included: 38**

Ingredients

Describes the list of ingredients in free text format and includes additional information on potential allergens, traces, or secondary components. This section is crucial for people with dietary restrictions or specific allergies.

- **Fields included: 18**

Miscellaneous Data

Groups complementary information such as serving size, number of additives, presence of palm oil, main assigned category, and links to packaging images. These are diverse data fields that don't strictly fit into nutritional categories but enrich the overall product analysis.

- **Fields included: 28**

Nutritional Information

Provides a detailed breakdown of the product's composition per 100 grams. It includes energy values, macronutrients (fats, carbohydrates, proteins), types of fatty acids, sugars, fiber, alcohol, caffeine, vitamins, and minerals. It is the most extensive and technical section of the dataset.

- **Fields included: 77**

Carbon Footprint

Records, when available, the product's carbon footprint based on packaging information. This data allows for environmental impact assessments of certain foods.

- **Fields included: 2**

Nutri-Score and Other Indicators

Includes nutritional ratings such as the Nutri-Score (French system), the UK Nutrition Score, the NOVA group (degree of processing), and environmental scores. These metrics make it easier to compare the nutritional quality and sustainability of products visually and quickly.

- **Fields included: 12**