



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

Documentación general de proyecto

Identificación de campos clave del dataset:

En este proyecto se identificó el campo "code", también conocido como código EAN, como el campo clave del dataset. Esto se debe a que dicho campo es único para cada producto registrado y está compuesto únicamente por caracteres numéricos, lo cual permite utilizarlo de manera eficiente en funciones de dispersión o hash. Al ser un identificador inequívoco, resulta ideal para establecer relaciones uno a uno entre los productos y sus datos correspondientes.

Definición de criterios de búsqueda relevantes para el tipo de datos:

Se definió como principal criterio de búsqueda el uso del código de barras EAN. La decisión parte del supuesto de uso realista en el cual un usuario tiene a la mano un artículo del supermercado y desea consultar la información nutricional, comercial o ambiental del mismo. De este modo, el código de barras se convierte en el punto de entrada natural para el sistema, permitiendo búsquedas rápidas y específicas.

Sobre el código EAN:

El código EAN (*European Article Number*) es un estándar internacional para la identificación de productos a través de códigos de barras. Fue desarrollado originalmente por la organización europea EAN International y actualmente es gestionado por GS1. Está compuesto típicamente por 13 dígitos que identifican de manera única un producto comercial en el mercado global.

Justificación del criterio de búsqueda implementado:

La búsqueda basada en el campo `code` se considera directa, ya que se accede a los datos del producto sin necesidad de recorrer el conjunto completo de registros. Esta estrategia se fundamenta en el uso de una tabla hash, la cual ofrece un tiempo de acceso promedio constante. Esta aproximación no solo es eficiente, sino que permite introducir conceptos importantes en la programación de estructuras de datos, como el acceso por clave única, la resolución de colisiones y la transformación de cadenas en claves numéricas. Además, por su simplicidad, es una excelente base para extender el sistema hacia búsquedas más avanzadas en el futuro.

Rangos de valores válidos para la entrada:

En cuanto a los valores válidos para la entrada del usuario, se estableció que únicamente se permitirán cadenas compuestas exclusivamente por dígitos, con una longitud mínima de

un carácter y una máxima de trece. Esto garantiza la compatibilidad con el formato estándar del código EAN y permite evitar entradas erróneas o mal formateadas que puedan comprometer la estabilidad del sistema.

General Project Documentation

Identification of Key Dataset Fields:

In this project, the field "code", also known as the EAN code, was identified as the key field of the dataset. This is because the field is unique for each registered product and consists only of numeric characters, which makes it efficient to use in hash or dispersion functions. As an unambiguous identifier, it is ideal for establishing one-to-one relationships between products and their corresponding data.

Definition of Relevant Search Criteria for the Data Type:

The primary search criterion was defined as the use of the EAN barcode. This decision is based on a realistic usage scenario where a user has a supermarket item in hand and wants to consult its nutritional, commercial, or environmental information. In this way, the barcode becomes the natural entry point for the system, enabling fast and specific searches.

About the EAN Code:

The EAN code (European Article Number) is an international standard for identifying products using barcodes. It was originally developed by the European organization EAN International and is now managed by GS1. It typically consists of 13 digits that uniquely identify a commercial product in the global market.

Justification for the Implemented Search Criterion:

The search based on the "code" field is considered direct, as it allows access to product data without scanning the entire set of records. This strategy is based on the use of a hash table, which provides an average constant access time. This approach is not only efficient but also introduces important concepts in data structure programming, such as access via unique keys, collision resolution, and the transformation of strings into numerical keys. Additionally, due to its simplicity, it serves as an excellent foundation for extending the system to more advanced search capabilities in the future.

Valid Input Value Ranges:

As for valid user input values, it was established that only strings composed exclusively of digits will be allowed, with a minimum length of one character and a maximum of thirteen. This ensures compatibility with the standard EAN code format and helps avoid erroneous or poorly formatted entries that could compromise the stability of the system.