

1. Introducción

El presente proyecto tiene como objetivo el diseño e implementación de una base de datos relacional centrada en datos de café, utilizando como fuente principal un conjunto de archivos CSV extraídos de Kaggle. Estos datos reflejan características físicas, químicas y de calidad de distintas variedades de café como *Arabica* y *Robusta*.

El proyecto busca aplicar prácticas de normalización, integridad referencial y diseño orientado a consultas analíticas, logísticas y operativas que podrían utilizarse en un entorno real, como un marketplace o laboratorio de calidad cafetera.

2. Objetivo

El objetivo del proyecto es crear una base de datos robusta, bien estructurada y normalizada que permita:

- Gestionar y consultar información sobre cafés producidos y evaluados.
 - Proveer soporte para operaciones analíticas, informes de calidad y seguimiento de variedades por país, tipo y finca.
 - Integrar vistas, funciones y procedimientos almacenados que permitan automatizar análisis y facilitar reportes.
-

3. Situación Problemática

Los productores de café, laboratorios y distribuidores manejan grandes volúmenes de información de manera desestructurada, lo que dificulta los análisis de calidad, trazabilidad, rendimiento por origen y comparativas entre variedades.

Con la implementación de una base de datos estructurada, se cubren brechas relacionadas con la eficiencia en la toma de decisiones, reducción de errores humanos en evaluaciones y mejora del acceso a la información histórica de evaluaciones.

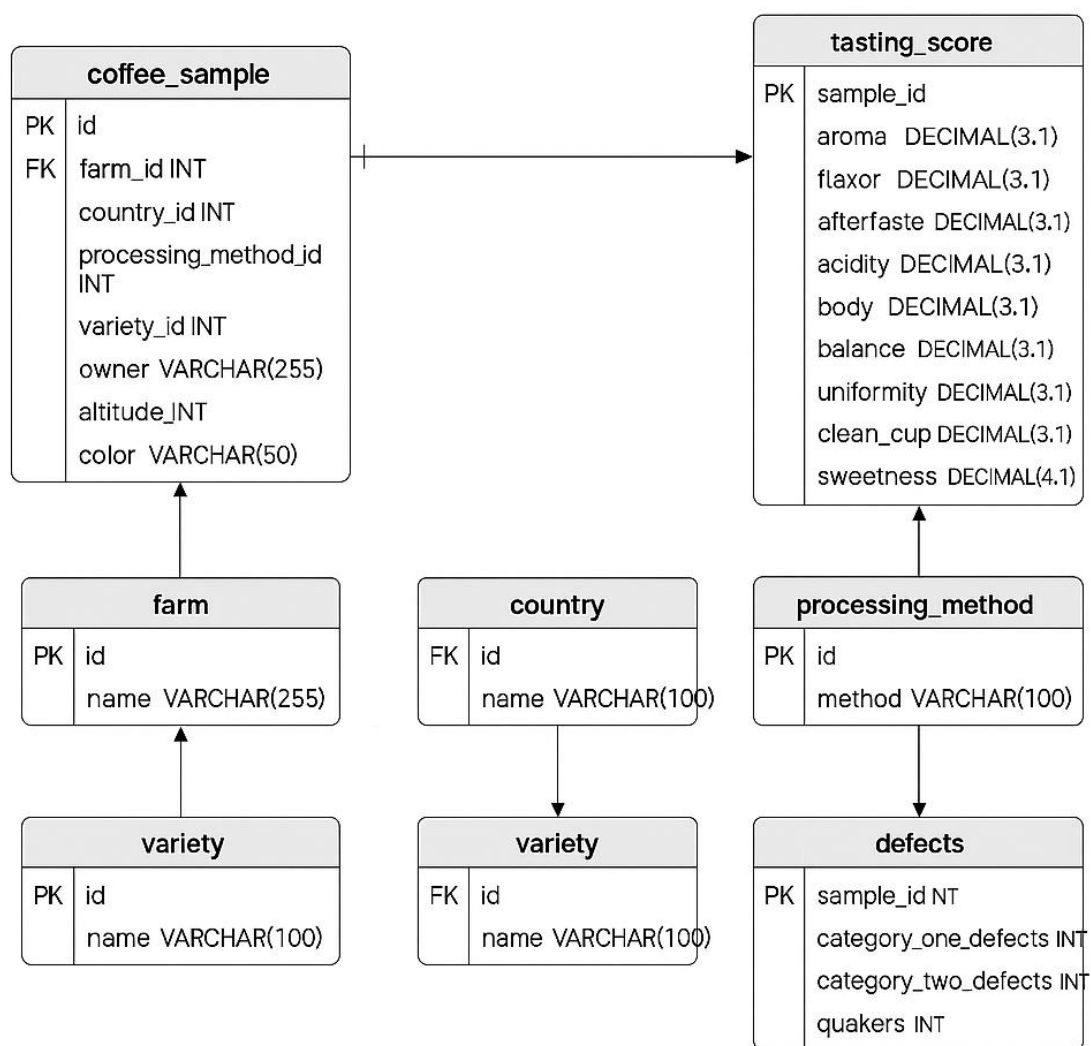
4. Modelo de Negocio

La solución se presenta como parte del sistema de una empresa dedicada al comercio y control de calidad de café a nivel internacional. Esta organización recibe

muestras de café, realiza evaluaciones con catadores certificados y conserva registros para cada envío/productor.

Este modelo podría servir a:

- **Laboratorios de calidad cafetera**
- **Marketplaces de cafés especiales**
- **Organizaciones exportadoras de café**



(Diagrama Entidad-Relación generado para la base de datos)

5. Listado de Tablas

Tabla	Descripción
-------	-------------

Coffee	Información general del café, origen, variedad, finca, etc.
--------	---

Evaluation	Resultados y notas de cata.
------------	-----------------------------

Roaster	Información sobre el tostador o exportador.
---------	---

Country	País de origen.
---------	-----------------

Species	Tipo de café (Arabica / Robusta).
---------	-----------------------------------

Processing	Método de procesamiento.
------------	--------------------------

Grade	Clasificación y estándar de calidad.
-------	--------------------------------------

Incluye descripciones completas, tipos de datos y claves en el archivo
sql/01_create_tables_mysql.sql.

6. Vistas

Vista 1: vw_top_quality_countries

- **Objetivo:** Mostrar los países con cafés de mayor puntuación promedio.
- **Tablas:** Evaluation, Country, Coffee
- **SQL:** Ver archivo 02_views_functions_storedprocedures_triggers.sql

Vista 2: vw_avg_flavor_by_species

- **Objetivo:** Promedio de sabor por tipo de especie.
- **Tablas:** Evaluation, Species, Coffee

7. Funciones Personalizadas

Función 1: fn_get_coffee_quality_level(score FLOAT)

- **Objetivo:** Clasificar calidad del café (Excelente, Bueno, Regular, etc.).

- **Tablas usadas:** Evaluaciones

Función 2: fn_country_coffee_count(country_name TEXT)

- **Objetivo:** Retorna el total de cafés registrados por país.
-

8. Stored Procedures

SP 1: sp_insert_new_coffee(...)

- **Objetivo:** Insertar un nuevo café y asociarlo a tablas relacionadas.
- **Tablas usadas:** Coffee, Species, Country, Roaster

SP 2: sp_get_country_report(country_name TEXT)

- **Objetivo:** Informe detallado de cafés por país.
 - **Tablas:** Coffee, Evaluation, Grade
-

9. Triggers

Trigger 1: trg_update_score_category

- **Objetivo:** Actualiza automáticamente la categoría de calidad al insertar una evaluación.
- **Tabla:** Evaluation

Trigger 2: trg_prevent_null_scores

- **Objetivo:** Previene inserciones con valores nulos críticos.
 - **Tabla:** Evaluation
-

10. Inserción de Datos

Archivos CSV usados:

- data/arabica_data_cleaned.csv
- data/robusta_data_cleaned.csv
- data/merged_data_cleaned.csv

Instrucciones para importar:

1. Colocar los archivos .csv en la carpeta /data del repositorio.
2. Asegurarse de tener habilitada la opción LOAD DATA INFILE en MySQL.
3. Ejecutar el script sql/03_insert_data_from_csv.sql para insertar todos los datos.

11. Estructura de Repositorio

/data

arabica_data_cleaned.csv

robusta_data_cleaned.csv

merged_data_cleaned.csv

/sql

01_create_tables_mysql.sql

02_views_functions_storedprocedures_triggers.sql

03_insert_data_from_csv.sql

/diagrams

Diagrama E-R.png

README.md

Entrega2_Martin.pdf

12. Conclusión

Este proyecto ofrece una solución escalable y funcional para la gestión de datos cafetaleros, utilizando las mejores prácticas en SQL y diseño relacional. La estructura

modular permite mantener, consultar y extender el sistema fácilmente para nuevos escenarios de negocio.
