

**Unidad 1. Evidencia de Aprendizaje 1 - Creación de una base de datos analítica.**

Realizado por:

Eduard Andrés Páez Bohórquez

Profesor:

Andres Felipe Callejas Jaramillo

Facultad de Ingenierías, Institución Universitaria Digital de Antioquia Proyecto Integrado

Big Data - PREICA2501B020109

11 de mayo de 2025

## Contenido

<b>Unidad 1. Evidencia de Aprendizaje 1 - Creación de una base de datos analítica. ....</b>	<b>1</b>
<b>Introducción. ....</b>	<b>3</b>
<b>Descripción del problema. ....</b>	<b>4</b>
<b>Objetivos. ....</b>	<b>5</b>
<b>objetivo General: ....</b>	<b>5</b>
<b>Objetivos Específicos: ....</b>	<b>5</b>
<b>Descripción de los datos disponibles. ....</b>	<b>6</b>
<b>Solución propuesta. ....</b>	<b>7</b>
<b>Elección del SGBD: ....</b>	<b>7</b>
<b>Esquema diseñado: ....</b>	<b>7</b>
<b>Estructura del esquema: ....</b>	<b>7</b>
<b>Relación entre las tablas: ....</b>	<b>8</b>
<b>Metodología empleada. ....</b>	<b>9</b>
<b>Resultados. ....</b>	<b>10</b>
<b>Conclusiones. ....</b>	<b>11</b>
<b>Bibliografía ....</b>	<b>12</b>

## **Introducción.**

El objetivo del presente documento es describir la implementación y el diseño de un sistema de gestión de base de datos (SGBD), destinado para el almacenamiento y análisis de los datos históricos de la Tasa Representativa del Mercado (TRM) de la divisa del dólar estadounidense (USD) frente al peso colombiano (COP). El objetivo principal que se persigue es el de proporcionar una herramienta que brinde la posibilidad de realizar el seguimiento y la comprensión de la Tasa Representativa del Mercado (TRM), para una empresa que importa repuestos automotrices, radicada en Valledupar, Cesar, de tal manera que reduzca los riesgos financieros que conlleva.

### **Descripción del problema.**

En la ciudad de Valledupar, situada en el departamento del Cesar, se encuentra una empresa muy importante dedicada a la importación de repuestos para automóviles que constantemente debe enfrentarse a problemas financieros provocados por la variación que pueda tener el dólar estadounidense (USD) con el peso colombiano (COP). Dado que la empresa realiza todas las compras a su proveedor en dólares, la TRM puede influir en los costos de adquisición de los repuestos, en el margen de utilidad y en los precios entregados al cliente final, que al mismo tiempo hace que estos puedan verse afectados.

En este momento, la empresa no tiene un software que lleve control ni análisis del comportamiento histórico del dólar, ya que las consultas de la TRM son realizadas manualmente, esporádicamente y sin una fuente que les permita ver una tendencia, un promedio mensual o un periodo crítico. Esto implica que hay una limitación en la planificación de precios.

Como respuesta a tal necesidad, se propone crear una base de datos analítica que recoja y organice de manera automatizada los datos históricos del comportamiento del dólar con respecto al peso colombiano; la obtención de esta información se podrá extraer mediante el uso de técnicas de web scraping de una dependencia de Python llamada yfinance. Con dicha base de datos la empresa podrá extraer consultas, generar informes, visualizar tendencias y tomar decisiones con información respecto a la importación de productos y la previsión de precios.

## **Objetivos.**

### **objetivo General:**

Desarrollar un sistema de gestión de base de datos que facilite el almacenamiento, la consulta y el análisis exhaustivo de datos históricos de la TRM, con el propósito de apoyar la toma de decisiones financieras fundamentadas en una empresa importadora de repuestos automotrices en Valledupar.

### **Objetivos Específicos:**

1. Definir un modelo de datos estructurado y eficiente, basado en la información histórica de la TRM proporcionada.
2. Seleccionar un Sistema de Gestión de Bases de Datos (SGBD) apropiado para los requerimientos del proyecto.
3. Implementar el esquema de la base de datos en el SGBD seleccionado.
4. Desarrollar consultas fundamentales para la visualización y el análisis de los datos de la TRM.

### **Descripción de los datos disponibles.**

Los datos existentes se encuentran organizados en forma de tabla HTML, y contienen la siguiente información histórica de interés relacionada con la TRM:

Fecha: el día que se recoge el valor de la TRM.

Cerrar: valor de cierre de la TRM en la fecha indicada.

Alto: valor alto que ha llegado a alcanzar la TRM en esa fecha.

Bajo: valor bajo que ha llegado a registrar la TRM en dicha fecha.

Abrir: valor de apertura de la TRM en la fecha indicada.

Volumen: número de acciones transaccionadas (no especialmente relevante para el análisis inicial de la TRM).

El modelo de datos se considera únicamente la Fecha, más los valores directos de la TRM que nos interesa; Apertura, Alto, Bajo, Cerrar y Cierre ajustado. El Volumen no se considerará métrica intrínseca de la TRM, por lo que se le dará un enfoque secundario en este primer enfoque.

## Solución propuesta.

### Elección del SGBD:

Se propone la utilización de SQLite. Este sistema se distingue por su naturaleza autónoma, su reducido tamaño, su velocidad operativa y su alta fiabilidad, ofreciendo todas las funcionalidades esenciales de un motor de base de datos SQL. Al ser un SGBD embebido, se ejecuta dentro de la aplicación, evitando la complejidad de un servidor de base de datos independiente, lo que lo hace adecuado para las necesidades de este proyecto.

### Esquema diseñado:

Para almacenar y analizar el comportamiento histórico del dólar frente al peso colombiano, se diseñó una base de datos en **SQLite** denominada histórico\_dólar. Este esquema permite organizar eficientemente los datos capturados mediante web scraping desde la librería de Python llamada yfinance y facilitar su posterior análisis en herramientas de BI.

### Estructura del esquema:

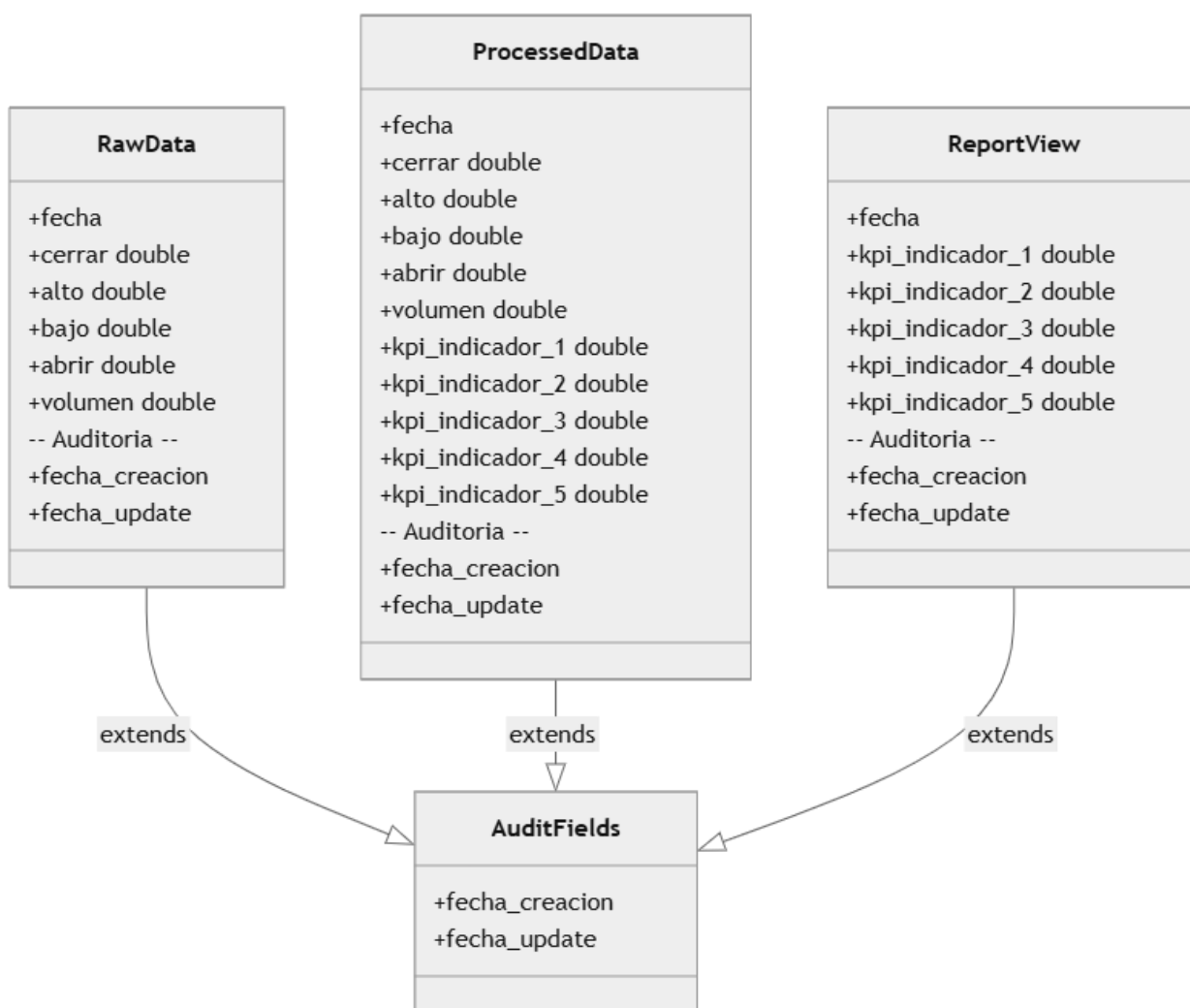
El esquema contiene las siguientes tablas:

- **RawData**: almacena los datos crudos directamente extraídos desde la fuente. Contiene campos como fecha, abrir, cerrar, alto, bajo, volumen, entre otros.
- **ProcessedData**: esta tabla contiene los datos transformados y enriquecidos, incluyendo KPIs calculados como variaciones porcentuales, rangos y promedios.
- **ReportView**: vista optimizada para visualización en herramientas de inteligencia de negocios. Incluye solo los indicadores clave y métricas relevantes.
- **AuditFields**: tabla abstracta reutilizada en las otras tres para registrar los campos de auditoría (fecha\_creacion, fecha\_update).

### Relación entre las tablas:

Las tablas RawData, ProcessedData y ReportView heredan los campos de auditoría de AuditFields, lo cual permite mantener trazabilidad de los registros.

A continuación, se presenta el diagrama del esquema diseñado:





### **Metodología empleada.**

Definición del modelo de datos: Se realizó un análisis de la estructura de la tabla HTML para discernir sobre las entidades y atributos pertinentes para el seguimiento del histórico del dólar frente al peso colombiano. El resultado fue la concepción de una tabla única, histórico\_dólar, con los campos necesarios para consignar la información histórica.

Selección del SGBD: La elección de SQLite se fundamentó en su simplicidad, su carácter embebido y su facilidad de implementación, atributos que lo hacen apropiado para este contexto donde no se requiere una infraestructura de base de datos compleja.

Implementación del esquema de la base de datos: Se proporcionó el código alojado en un repositorio de GitHub el cual facilita el trabajo en equipo y la actualización de la información de una manera más ágil y segura.

Desarrollo de consultas básicas: Se incluirán ejemplos de consultas SQL diseñadas para facilitar la recuperación y el análisis inicial de los datos.

## **Resultados.**

Se diseñó una base de datos utilizando SQLite para almacenar los datos históricos del comportamiento del dólar (USD) frente al peso colombiano (COP). A través de web scraping con yfinance, los datos de las tasas de cambio fueron obtenidos y almacenados de forma eficiente para su posterior análisis.

Estructura de la Base de Datos:

### **1. Tabla tasas\_cambio:**

- **id** (INTEGER, PK, AUTOINCREMENT)
- **fecha** (DATE)
- **cerrar** (DECIMAL)
- **alto** (DECIMAL)
- **bajo** (DECIMAL)
- **abrir** (DECIMAL)
- **volumen** (DECIMAL)

### **Conclusiones.**

El sistema propuesto para la gestión de base de datos está basado en SQLite, y en el esquema diseñado permite establecer una base de datos con los servicios de un gestor muy robusto para el almacenamiento y especial para el almacenamiento y consulta de los históricos de la TRM. Este es un valioso recurso para la empresa de Valledupar que le permite analizar las tendencias del tipo de cambio, y potencialmente tomar decisiones financieras más informadas que ayuden a mitigar riesgos derivados de la volatilidad del dólar. Las siguientes etapas pueden ir orientadas en la integración con herramientas de visualización y análisis avanzado.

## Bibliografía

foundation, P. s. (2025). *Python.org*. Obtenido de Python : <https://www.python.org/>

GitHub · Build and ship software on a single, collaborative platform. (2025). GitHub.  
<https://github.com/>