

## **Proyecto Integrado V - Linea de Énfasis (Entrega 1)**

### **Estudiante:**

Sergio Raúl Bustos

### **Docente:**

Andres Felipe Callejas Jaramillo

### **Curso:**

Proyecto Integrado V

Programa Ingeniería De Software Y

Datos Facultad Ingenierías y

Ciencias Agropecuarias Institución

Universitaria Digital de Antioquia

Mayo

2025

## **Resumen**

Este proyecto tiene como objetivo desarrollar un sistema automatizado para la recopilación y almacenamiento de datos históricos del valor de las acciones de Apple Inc. (AAPL), desde su debut en la bolsa el 12 de diciembre de 1980 hasta la fecha actual. La obtención de los datos se llevó a cabo a través del sitio web Yahoo Finanzas, utilizando para ello Python y técnicas de Web Scraping. El procesamiento de los datos se realizó con la biblioteca Pandas, y los resultados fueron guardados tanto en archivos CSV como en una base de datos SQLite. Como producto final, se construyó una aplicación funcional que permite consultar y almacenar esta información de forma automatizada, organizada en torno a un flujo de ejecución principal (Main), con capacidad de integrarse en entornos productivos mediante herramientas de integración continua como GitHub Actions.

## **Introducción**

En este proyecto se presenta un sistema automatizado para la recopilación, almacenamiento y análisis de datos históricos del precio de las acciones de Apple Inc. (AAPL) en el mercado bursátil. El objetivo principal es facilitar el acceso y tratamiento de esta información con fines de análisis financiero, permitiendo comprender de forma más detallada la evolución y comportamiento de la empresa en el mercado a lo largo del tiempo. La motivación radica en el creciente interés por el análisis de datos históricos como base para la toma de decisiones estratégicas, estudios de tendencias y desarrollo de modelos predictivos en el ámbito bursátil.

## Metodología

La metodología empleada en esta etapa inicial del proyecto consistió en la creación de un script automatizado enfocado en la extracción, transformación y almacenamiento de datos históricos de las acciones de Apple Inc. (AAPL). Este proceso se dividió en las siguientes fases:

- **Obtención de datos:** Se utilizó la librería requests para realizar peticiones HTTP al portal Yahoo Finanzas, incorporando un encabezado User-Agent que imita la actividad de un navegador web. El contenido HTML recuperado fue interpretado con la ayuda de BeautifulSoup, permitiendo ubicar y extraer la tabla correspondiente a los registros históricos del precio de las acciones de Apple.
- **Tratamiento de datos:** Posteriormente, los datos fueron estructurados en un DataFrame mediante la biblioteca Pandas. Durante esta etapa se procedió al ajuste de nombres de columnas, verificación de integridad de los datos y eliminación de registros vacíos o incorrectos, asegurando así un formato limpio y apto para análisis posteriores.
- **Monitoreo de ejecución:** Se integró un sistema de registro personalizado mediante la librería logging, con el fin de documentar la actividad del programa. Este sistema captura eventos clave como fallos de conexión, inicialización de clases, y la correcta creación de los archivos procesados.
- **Persistencia de la información:** Los datos ya transformados se almacenaron en dos formatos: como archivos .csv y dentro de una base de datos SQLite. Ambos tipos de archivos se guardan automáticamente en directorios generados por el sistema, específicamente en las carpetas data y static.
- **Diseño modular del sistema:** La aplicación fue construida con una estructura modular, distribuyendo sus funcionalidades en clases independientes: Logger, Collector y main. Esta arquitectura mejora la legibilidad del código, facilita las pruebas unitarias y permite futuras ampliaciones. La ejecución general se controla desde el módulo main, el cual también gestiona la creación de carpetas y la generación de registros de actividad.