

Proyecto Integrado V - Evidencia de aprendizaje 1

Mateo Valencia Minota

Carlos Andrés Cardona Quintero

Proyecto Integrado V

PREICA2501B020128

Andrés Felipe Callejas Jaramillo

Medellín - Antioquia

11/05/20245

Resumen

Este trabajo presenta el desarrollo de una solución automatizada para la recolección, almacenamiento y análisis del precio del petróleo crudo (Crude Oil Futures - CL=F), considerado un indicador clave de la economía global. La solución implementada permite obtener y mantener actualizada una base histórica de datos con trazabilidad y control de versiones mediante GitHub Actions. El enfoque orientado a datos tiene como finalidad preparar una base sólida para el análisis exploratorio, la generación de KPIs (indicadores clave de rendimiento) y el posterior entrenamiento de modelos predictivos. Este documento introduce el marco conceptual, los objetivos y la relevancia económica del indicador seleccionado.

Introducción

El precio del petróleo crudo, representado en los mercados financieros mediante contratos de futuros como **CL=F (Crude Oil Futures)**, es uno de los indicadores económicos más influyentes a nivel global. Este valor refleja el precio al que se negocia un barril de petróleo con entrega futura, y es utilizado ampliamente en decisiones macroeconómicas, políticas energéticas, estrategias industriales y proyecciones financieras.

Además de su uso tradicional, el precio del petróleo ofrece **valor predictivo** sobre otros sectores: inflación, transporte, industria manufacturera, entre otros. En este contexto, construir una base sólida de datos históricos, automatizada y trazable, permite no solo su seguimiento, sino también su análisis con herramientas estadísticas y modelos de machine learning para comprender y anticipar su comportamiento, como lo haremos en entregas posteriores.

Objetivo general

Diseñar e implementar una solución automatizada mediante GitHub Actions, usando un scrapper para recolectar y respaldar la información histórica del precio del petróleo crudo, desde el año 200 hasta la actualidad, con el fin de entender su comportamiento y facilitar la toma de decisiones económicas y financieras basadas en datos y, adicionalmente, como ejercicio de práctica para reforzar los conocimientos recibidos en diferentes materias.

Objetivos específicos

1. **Recolectar datos históricos diarios** del precio del petróleo crudo (Open, High, Low, Close, Adj Close y Volume), desde la fuente Yahoo Finance, utilizando técnicas de web scraping, automatización mediante workflows en GitHub Actions y control de versiones (GitHub).
2. **Almacenar los datos recolectados**, asegurando su integridad, homogeneidad temporal y correcta persistencia en formato .csv
3. **Preprocesar** los datos almacenados mediante validación y estandarización, preparándolos para el cálculo de indicadores clave para posteriores entregas.
4. **Documentar** la solución con base en el alcance de la entrega 1.

Bibliografía

GitHub. (s.f.). *Documentation*. <https://docs.github.com/>

GitHub Actions. (s.f.). *Documentation*. <https://docs.github.com/actions>

Python Software Foundation. (s.f.). *Python 3 Documentation*. <https://docs.python.org/3>

scikit-learn developers. (s.f.). *scikit-learn: Machine Learning in Python*. <https://scikit-learn.org/>

Yahoo Finance. (s.f.). *Crude Oil WTI (CL=F)*. <https://finance.yahoo.com/quote/CL%3DF/>

Yahoo Finance. (s.f.). *Yahoo Finanzas*. <https://finance.yahoo.com/>

Seabold, S., & Perktold, J. (2010). *Statsmodels: Econometric and Statistical Modeling with Python*. <https://www.statsmodels.org/stable/>