

Informe Ejecutivo

Análisis Exploratorio de Datos sobre Precios de Bitcoin (BTC/USD)

Introducción

El análisis exploratorio realizado se centró en los precios históricos de **Bitcoin por minuto (BTC/USD)**, comprendiendo más de **7,1 millones de registros**. El objetivo principal fue comprender la estructura de los datos, su calidad, las correlaciones existentes y posibles anomalías que pudieran influir en un modelo predictivo o una toma de decisiones de negocio.

Características del Dataset

- **Tamaño total:** 7.188.157 registros con 6 columnas.
 - **Variables incluidas:**
 - Timestamp
 - Open
 - High
 - Low
 - Close
 - Volume
 - **Consistencia:** No se encontraron valores nulos en ninguna de las variables.
 - **Uso de memoria:** Aproximadamente 329 MB.
-

Calidad de los Datos

- El dataset es **completo y sin valores faltantes**, lo que permite trabajar de manera inmediata en análisis descriptivos y modelos.
 - Las variables numéricas presentan formato correcto en `float64`.
-

Correlaciones Observadas

Mediante análisis de correlación:

- Existe una **fuerte correlación positiva** entre las variables **Open, High, Low y Close**, lo cual es coherente dado que representan diferentes puntos del mismo precio en la ventana temporal.
 - **Volume** muestra una correlación más baja respecto al comportamiento de precios, sugiriendo que no siempre aumenta en sincronía con las variaciones de precio.
-

Identificación de Anomalías

Se aplicó un método basado en **dos desviaciones estándar respecto a la media** sobre la variable *Low* indexada por *High*.

- Se detectaron valores que caen fuera del rango esperado.
- Estos puntos atípicos pueden deberse a errores en la captura de datos, gaps de mercado, volatilidad extrema o actividad inusual en transacciones.
- La identificación de anomalías es crítica ya que podrían distorsionar análisis estadísticos y modelos predictivos si no se tratan adecuadamente.

Principales Conclusiones

1. **Calidad de los datos:** El dataset es robusto, sin valores vacíos, y está listo para análisis avanzados y modelado.
2. **Estructura del mercado:** Los precios de apertura, cierre, máximo y mínimo mantienen una correlación natural muy fuerte entre ellos, lo que confirma la consistencia del dataset.
3. **Volumen transaccional:** Su menor correlación frente a precios señala que las variaciones de volumen no siempre predicen movimientos directos en el precio.
4. **Anomalías detectadas:** La existencia de outliers debe revisarse con mayor detalle, ya que pueden representar tanto oportunidades (picos de volatilidad) como ruido en los datos.

Recomendaciones

- **Preprocesamiento adicional:** Estandarizar fechas (Timestamp) y convertirlas a formatos temporales útiles (horas, días, semanas).
- **Tratamiento de outliers:** Aplicar filtros o técnicas de suavizado para minimizar el impacto de anomalías.
- **Exploración temporal:** Analizar la evolución del precio y volumen en distintas escalas (horaria, diaria, semanal) para detectar patrones o estacionalidades.
- **Uso de técnicas avanzadas:** Considerar modelos de series temporales (ARIMA, Prophet, LSTM) que incorporen tanto precio como volumen.