

## Práctica 1: Selección del conjunto de datos

**Nombre y apellidos:** Paula Muñoz Lago

**Usuario UOC:** paulamlago

**Estudios que cursa (Máster o Grado):** Máster en Ciencia de Datos

Para abordar la práctica final se proponen datos sobre la evolución temporal del precio del Bitcoin, para así estudiar la popularización de las criptomonedas con el paso del tiempo en la sociedad. Se puede comparar dicha evolución con otras monedas populares como el Ethereum que, aunque no tienen ni la misma capitalización ni precio, han sufrido también un incremento en la popularidad, cuya tendencia podría ser comparada con la del Bitcoin.

El dataset conteniendo información de cada moneda se encuentra en el siguiente link de Kaggle: <https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory>, a éste se le han aplicado unas ciertas transformaciones para enriquecer el conjunto utilizando Python, cuyo código y resultado se encuentra en el siguiente repositorio de Github: <https://github.com/paulamlago/StockDataVisualisation>.

Ahí encontramos archivos de datos correspondientes a la historia de cada moneda (por ejemplo, el Bitcoin incluye información diaria desde 2013). Las columnas más relevantes para el estudio son; la fecha, que se ha convertido en el índice tras el procesamiento de datos previo a la visualización, el precio más alto, bajo, de apertura y cierre de cada día, así como el volumen de operaciones y la capitalización del mercado en dicha fecha. A estas variables se han unido nuevas tras dicho proceso de transformación, como la diferencia entre el precio de apertura y cierre (DiffDay), viendo así si ha fluctuado mucho el precio a lo largo del día, o si finalmente han ganado los bajistas o alcistas (Up/Down). También se han añadido columnas como el porcentaje de variación del precio con respecto al día anterior (Change), y variables técnicas como las medias móviles de 5 y 10 días y el MACD (Moving Average Convergence Divergence) y RSI (Índice de fuerza relativa).

Se ha escogido este dataset dada su relevancia en el contexto económico y social, puesto que desde hace unos años el uso de las criptomonedas se ha incrementado, siendo cada vez más los países en los que se acepta Bitcoin, es decir, existe una ley que lo regula. En este caso los datos no tratan información de colectivos sociales, por lo tanto no se tiene en cuenta la perspectiva de género. El número de registros son 2862 en el caso del Bitcoin y 2031 en el caso del Ethereum, y 10 variables cada conjunto. La mayoría de datos son cuantitativos continuos, dado que tratan precios, el volumen de operaciones, la capitalización de la moneda o valores técnicos. Sin embargo también dispone de variables Categóricas, el símbolo de la moneda, que no se ha eliminado tras el proceso de limpieza de los datos para poder comparar los datos de varias monedas, disponiendo así de su etiqueta.

Este tipo de datos se visualizan constantemente en brokers o páginas de información de criptomonedas, como [Binance](https://www.binance.com/), en la que se aprecia una visualización en la que se muestra una

evolución del precio del Bitcoin de las últimas horas, aunque al ser interactiva puede ampliarse el rango, también incluyendo información del volumen. Sin embargo, esta visualización no dispone de información de posibles diferencias en la tendencia con otras criptomonedas o capitalización de la moneda. Además, la vista está sobrecargada con información sobre ordenes de compra y venta de los usuarios. La visualización que se propone en esta práctica es un dashboard claro y directo, también con visualizaciones interactivas, que constituirá una evolución de las visualizaciones ya existentes.

Así, concluyo con que este dataset ofrece grandes posibilidades de visualización a través de las variables originalmente presentes y otras que han sido computadas con Python. De esta forma, podemos responder a las preguntas siguientes: ¿Cómo ha fluctuado el precio de estos activos a lo largo de su historia? ¿Ha habido un incremento en el volumen de las operaciones diario? (En caso de que sí, nos indicaría que más personas están comprando y vendiendo, siendo este un indicativo del aumento de su popularidad). ¿Existen similitudes entre la tendencia creciente del uso del Bitcoin y el resto de criptomonedas como el Ethereum? ¿Y con el resto? ¿Ha habido algún día en el que el precio se haya disparado o desplomado? ¿En qué se ha basado?