

Web Scrapping: Valores históricos de precios de criptomonedas

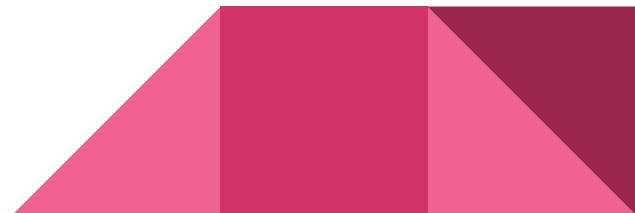
Rafael López García
Carlos Luis Gento de Celis

Noviembre 2021



Índice

1. Introducción
2. Contexto
3. Descripción y contenido del Dataset
4. Código Web Scraping
5. Conclusión



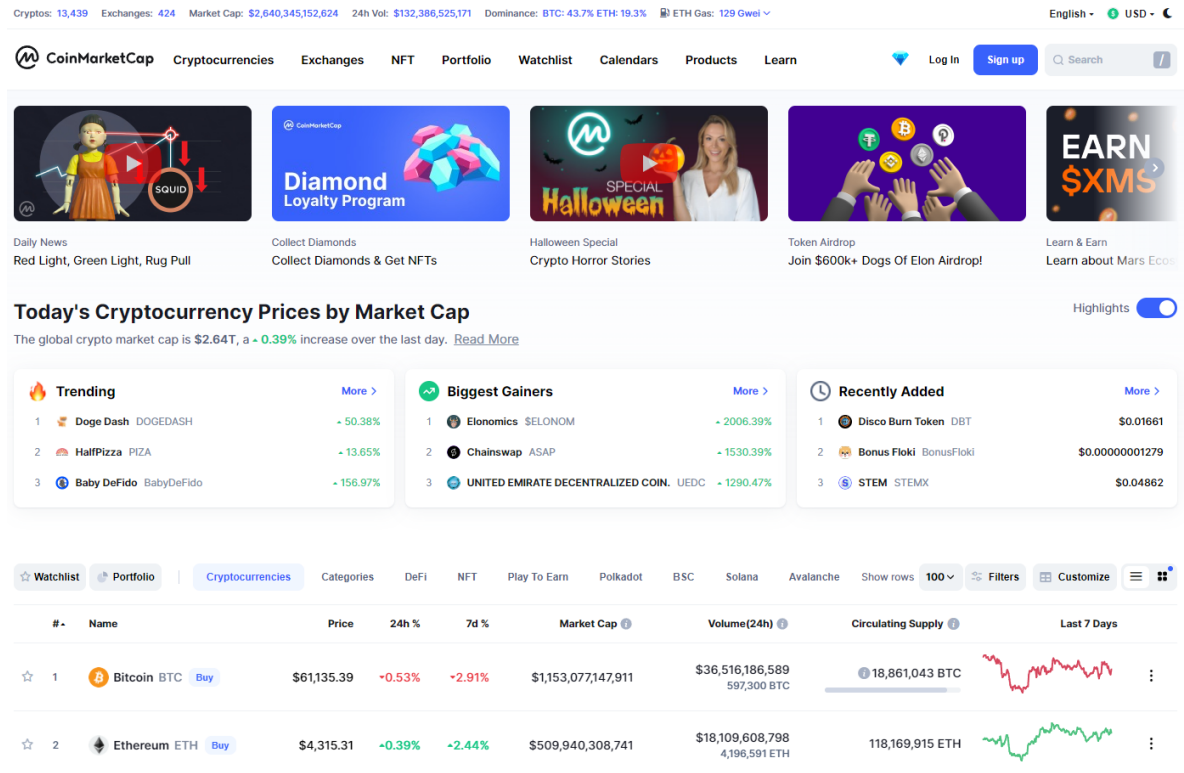
Introducción

- Las criptomonedas han generado gran interés desde que el Bitcoin fue presentado en 2008 (S. Nakamoto, 2008)
 - Actualmente existen miles de criptomonedas diferentes (Bitcoin, Ethereum, Litecoin, XRP...)
 - Gran debate sobre si las criptomonedas pueden ser “almacenes de valor”
- Presentamos un dataset sobre precios de criptomonedas generado con técnicas de Web Scraping



Contexto (I)

- Información extraída de coinmarketcap.com



Contexto (II)

- Información histórica de precios de criptomonedas (2013 – actualidad)

Bitcoin BTC

Price: \$61,200.45 -0.18%

Add to Main Watchlist

Overview

Market

Historical Data

Holders

Project Info

Wallets

News

Socials

Ratings

Analysis

Price Estimates

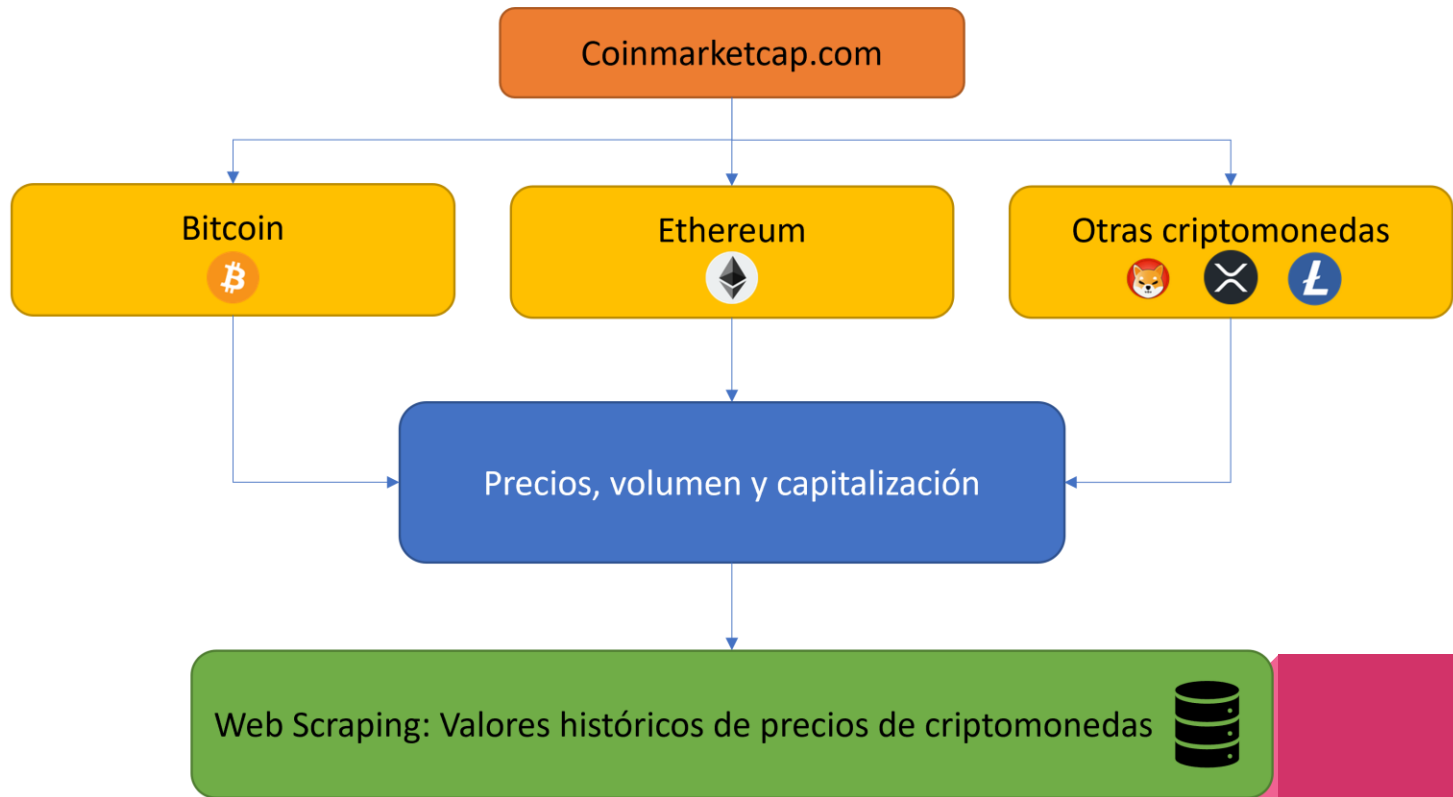
Share

Historical Data for Bitcoin

Date Range

Date	Open*	High	Low	Close**	Volume	Market Cap
Oct 31, 2021	\$61,850.49	\$62,406.17	\$60,074.33	\$61,318.96	\$32,241,199,927	\$1156,485,845,063
Oct 30, 2021	\$62,239.36	\$62,330.15	\$60,918.39	\$61,888.83	\$32,157,938,616	\$1,167,178,054,167
Oct 29, 2021	\$60,624.87	\$62,927.61	\$60,329.97	\$62,227.96	\$36,856,881,767	\$1,173,512,442,368
Oct 28, 2021	\$58,470.73	\$62,128.63	\$58,206.92	\$60,622.14	\$45,257,083,247	\$1,143,165,999,365
Oct 27, 2021	\$60,352.00	\$61,435.18	\$58,208.19	\$58,482.39	\$43,657,076,893	\$1,102,757,767,323
Oct 26, 2021	\$63,032.76	\$63,229.03	\$59,991.16	\$60,363.79	\$34,878,965,587	\$1,138,177,770,430
Oct 25, 2021	\$60,893.93	\$63,729.33	\$60,691.80	\$63,039.83	\$31,064,911,614	\$1,188,575,312,422
Oct 24, 2021	\$61,368.34	\$61,505.80	\$59,643.35	\$60,930.84	\$27,316,183,882	\$1,148,743,134,468
Oct 23, 2021	\$60,694.63	\$61,743.88	\$59,826.52	\$61,393.62	\$26,882,546,034	\$1,157,410,091,263
Oct 22, 2021	\$62,237.89	\$63,715.02	\$60,122.80	\$60,692.26	\$38,434,082,775	\$1,144,131,483,274
Oct 21, 2021	\$66,002.23	\$66,600.55	\$62,117.41	\$62,210.17	\$45,908,121,370	\$1,172,684,282,558
Oct 20, 2021	\$64,284.59	\$66,930.39	\$63,610.67	\$65,992.84	\$40,788,955,582	\$1,243,927,428,207
Oct 19, 2021	\$62,043.17	\$64,434.53	\$61,622.93	\$64,261.99	\$40,471,196,346	\$1,211,242,526,571

Contexto (III)



Descripción y contenido del Dataset (I)

- El dataset contiene información diaria de más de 6000 criptomonedas.
- Período de referencia: 31 de octubre 2020 – 29 octubre 2021
- El dataset tiene 3 partes:
 - Lista de criptomonedas
 - Información histórica de las criptomonedas
 - Logos de criptomonedas
- Lista de criptomonedas
 - Contiene una lista con las criptomonedas del dataset
 - Campos:
 - **name:** Nombre de la criptomoneda
 - **symbol:** Código identificador de criptomoneda
 - **logo:** Nombre del fichero que contiene el logo de la criptomoneda
 - **historicalFile:** Fichero con la información histórica de la criptomoneda
 - **url:** url de la criptomoneda dentro de coinmarketcap.com

Descripción y contenido del Dataset (II)

- Información histórica de las criptomonedas:
 - Contiene la información diaria sobre los precios de las criptomonedas y su capitalización
 - Campos para cada criptomoneda:
 - **symbol:** Código identificador de la criptomoneda
 - **date:** Fecha de la sesión
 - **open:** precio de la criptomoneda en el momento de la apertura de la sesión diaria
 - **high:** precio máximo alcanzado durante la sesión diaria
 - **low:** precio mínimo alcanzado durante la sesión diaria
 - **close:** precio de la criptomoneda en el momento de cierre de la sesión diaria
 - **volumen:** Valor de las criptomonedas intercambiadas durante la sesión diaria
 - **marketcap:** Valor de capitalización de la criptomoneda.
- Logos de las criptomonedas



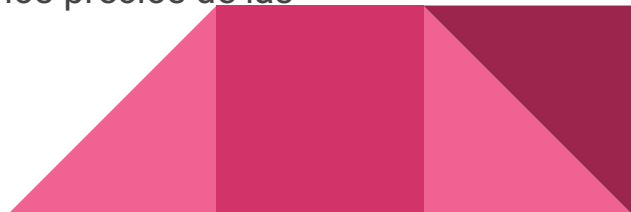
Código Web Scraping

- Se extrae información alojada en páginas HTML con técnicas de Web Scraping
- Lenguaje utilizado: Python
- Librerías utilizadas para Web Scraping: requests, BeautifulSoup y selenium
- El código se divide en 2 partes:
 1. Creación inicial de los ficheros del dataset
 - Costosa en tiempo
 2. Actualización de la información histórica diaria
 - Obtener la información más actualizada posible



Conclusión

- Dataset con información diaria de más de 6000 criptomonedas
- Información extraída de coinmarketcap.com a través de técnicas de Web Scraping y lenguaje Python
- El código desarrollado crea un dataset inicial (costoso en tiempo), pero también permite actualizar la información diaria
- Este dataset podría ayudar a responder, entre otras, las siguientes preguntas:
 - ¿Cuánto ha variado el precio del Bitcoin en las últimas semanas, meses o años? ¿Y cuánto lo ha hecho el del resto de criptomonedas?
 - ¿Hay patrones de comportamiento comunes y periódicos en los precios de las criptomonedas, como por ejemplo efectos estacionales?





¡Muchas gracias!