Acerca del conjunto de datos

El set de datos que voy a utilizar en principio es un csv que se compone de 6 columnas, 2860 filas que baje de la página Kaggle, el nombre original del csv es "Precios de Bitcoin (2014-2024)" del colaborador Alexander Kapturov (Propietario) la licencia es CCO de dominio público, la última fecha de actualización del dataset fue hace en mayo 2024.

					Date	0pen	High	Low	Close
lunes,	21	de	marzo	de	2016	413417999	413417999	410381012	413307007
martes,	22	de	marzo	de	2016	413131989	418375000	412531006	418088989
miércoles,	23	de	marzo	de	2016	418161011	419268005	417364014	418040985
jueves,	24	de	marzo	de	2016	418424011	418679993	415485992	416394012
viernes,	25	de	marzo	de	2016	416507996	418079987	415558014	417177002

Primer columna: Fecha incluye día mes y año

segunda columna: Precios de Bitcoin en dólares estadounidenses de apertura(open)

tercer columna:Precios de Bitcoin en dólares estadounidenses de subida(high) cuarta columna: Precio de bitcoin en dólares estadounidense de bajada (low) quinta columna: Precio de bitcoin en dólares estadounidense de cierre (close) sexta columna: Volumen movimientos Bitcoin (volume)

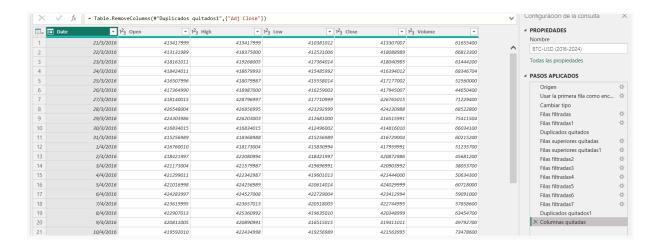
En principio el dataset tenía 3375 filas y 7 columnas, realice el proceso de ETL en Power Bi, donde elimine los primeros 2 años ya que no era información que necesitaba, elimine una columna duplicada, elimine los datos nulos. trabaje el tipo de datos unificando todos los valores numéricos en entero y la fecha en día mes y año. Lo guarde en un archivo esv listo para trabajar en mi proyecto. La ruta de mi archivo es

"C:/Users/SILCA/OneDrive/Escritorio/dataset/BTC-USD.csv"

Foto1

Consultas [1]	×	√ fx -	Csv.Document(File.Content	Configuración de la consulta	>					
BTC-USD (2016-2024)		A ^B C Column1	▼ A ^B _C Column2	▼ A ^B _C Column3	▼ A ^B _C Column4	▼ A ^B _C Column5	▼ A ^B _C Column6	▼ A ^B _C Column7	4 PROPIEDADES	
	1	Date	Open	High	Low	Close	Adj Close	Volume	Nombre	
	2	2014-09-18	456.859985	456.859985	413.104004	424.440002	424.440002	34483200	BTC-USD (2016-2024)	
	3	2014-09-19	424.102997	427.834991	384.532013	394.795990	394.795990	37919700	Todas las propiedades	
	4	2014-09-20	394.673004	423.295990	389.882996	408.903992	408.903992	36863600	■ PASOS APLICADOS	
	5	2014-09-21	408.084991	412.425995	393.181000	398.821014	398.821014	26580100		
	6	2014-09-22	399.100006	406.915985	397.130005	402.152008	402.152008	24127600	Origen	
	7	2014-09-23	402.092010	441.557007	396.196991	435.790985	435.790985	45099500	Usar la primera fila como enc	0
	8	2014-09-24	435.751007	436.112000	421.131989	423.204987	423.204987	30627700	Cambiar tipo	-85
	9	2014-09-25	423.156006	423.519989	409.467987	411.574005	411.574005	26814400	Filas filtradas Filas filtradas1	8
	10	2014-09-26	411.428986	414.937988	400.009003	404.424988	404.424988	21460800	Duplicados quitados	747
	11	2014-09-27	403.556000	406.622986	397.372009	399.519989	399.519989	15029300	Filas superiores guitadas	ö
	12	2014-09-28	399.471008	401.016998	374.332001	377.181000	377.181000	23613300	Filas superiores quitadas1	8
	13	2014-09-29	376.928009	385.210999	372.239990	375.467010	375.467010	32497700	Filas filtradas2	0
	14	2014-09-30	376.088013	390.976990	373.442993	386.944000	386.944000	34707300	Filas filtradas3	0
	15	2014-10-01	387.427002	391.378998	380.779999	383.614990	383.614990	26229400	Filas filtradas4	-0
	16	2014-10-02	383.988007	385.497009	372.946014	375.071991	375.071991	21777700	Filas filtradas5	-0
	17	2014-10-03	375.181000	377.695007	357.859009	359.511993	359.511993	30901200	Filas filtradas6	45
	18	2014-10-04	359.891998	364.487000	325.885986	328.865997	328.865997	47236500	Filas filtradas7	0
	19	2014-10-05	328.915985	341.800995	289.295990	320.510010	320.510010	83308096	Duplicados quitados1	
	20	2014-10-06	320.389008	345.134003	302.559998	330.079010	330.079010	79011800	Columnas quitadas	

Foto2



En la foto 1 se puede observar el set de datos original, en la foto 2 se puede observar el proceso de ETL realizado con los pasos aplicados.

Diccionario de datos del dataset

- Date: Fecha en la que se registraron los datos. Tipo de dato: fecha
- Open: Precio de apertura del mercado para el activo en la fecha especificada. Tipo de dato: entero
- High: Precio más alto alcanzado por el activo durante la jornada de mercado. Tipo de dato: entero
- Low: Precio más bajo alcanzado por el activo durante la jornada de mercado. Tipo de dato: entero
- Close: Precio de cierre del mercado para el activo en la fecha especificada. Tipo de dato: entero
- Volume: Volumen de acciones negociadas durante la jornada de mercado.
 Tipo de dato: entero

Fecha de entrega segunda entrega; 17/06/2024 Equipo de trabajo de Aprendizaje Automático Bolaña silvia(alumna) Mirabete Martin(profesor)