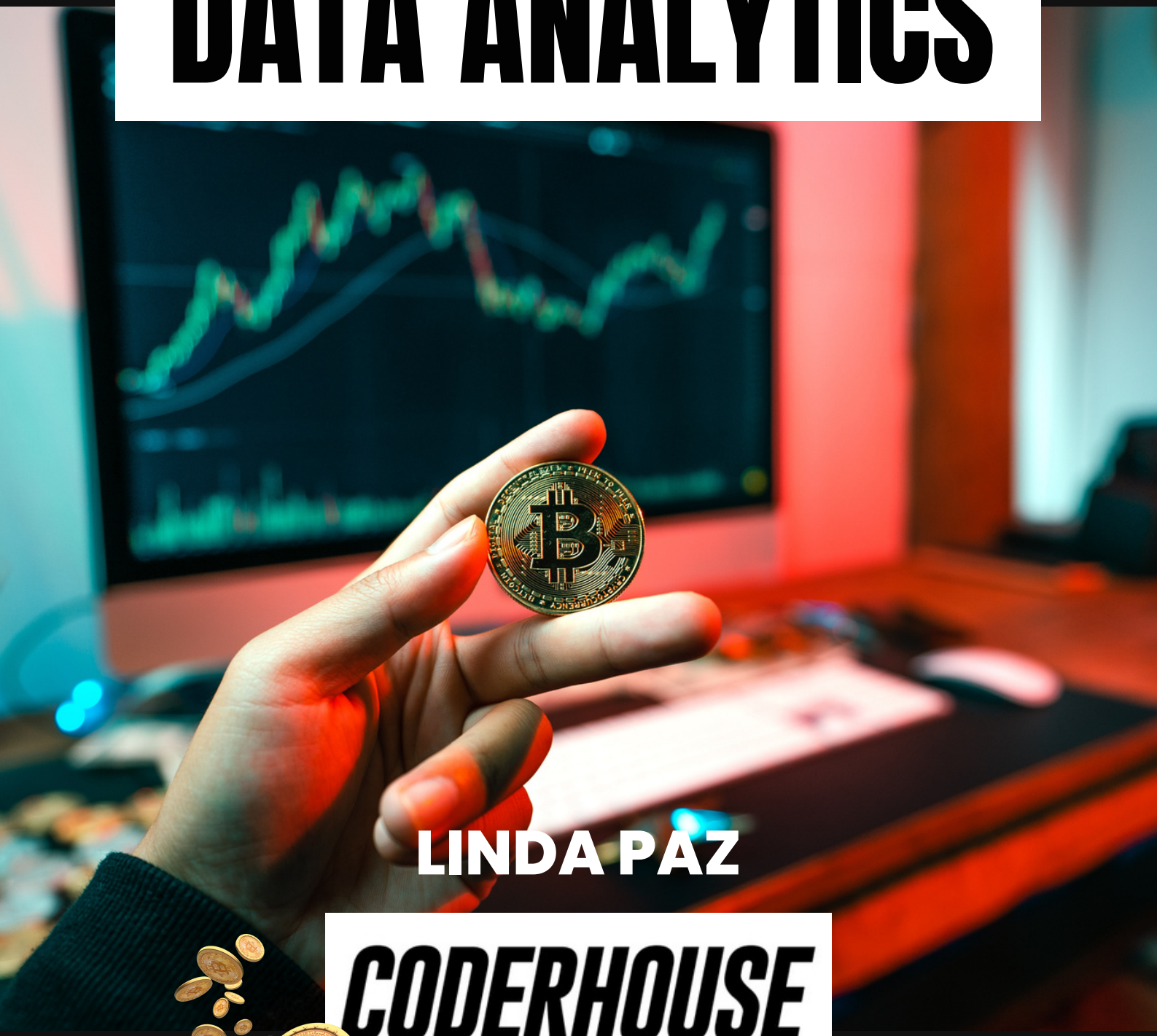
PROYECTO:
" BITCOIN HISTORICAL DATASET 2017-2022"

DATA ANALYTICS



LINDA PAZ

CODERHOUSE

2022

ÍNDICE:

Primera parte:

1. Descripción de la temática de los datos ..... Pág 2
2. Hipótesis ..... Pág 2
3. Datasets ..... Pág 2

Segunda parte:

4. Objetivos del proyecto ..... Pág 3
5. Alcance ..... Pág 3
6. Usuario final ..... Pág 3
7. Diagrama entiedad-relación ..... Pág 4
8. Listado de tablas ..... Pág 5
9. Listado de columnas ..... Pág 5

Tercera parte:

10. Transformaciones realizadas ..... Pág 7
11. Medidas calculadas generadas y sus fórmulas..... Pág 12

Cuarta parte:

12. Análisis funcional del tablero. ..... Pág 17

Quinta parte:

13. Futuras líneas ..... Pág 27

PRIMERA PARTE

1. Descripción de la temática de los datos

Siguiendo la metodología SMART, el siguiendo proyecto se definió como temática a tratar, el análisis de los diversos cambios históricos que tuvo el BITCOIN.
Estos datos fueron extraídos durante el periodo 2017-2022 y los mismos se centraron en registrar las caídas y las subas de precio que tuvo la moneda virtual. Con lo cual se puede analizar el volumen de transacciones realizadas en un período y la conversión en dólares que tuvo esta a lo largo de los años.

Siendo el tema mas importante de este proyecto, la gran volatibilidad que tuvo esta criptomoneda, pudiendo hacer que esta caiga un 50% en unos días o aumente un 12.000% en solo pocos meses.

2. Hipótesis

Por medio de las diversas operaciones realizadas en el periodo definido, se busca comprender el nivel de magnitud que tuvo esta moneda virtual (BITCOIN) y ver si las bajadas y subidas en los precios son coherentes con los sucesos externos.
Pudiendo medir por medio de la convertibilidad a dólar, las caídas, al igual que su mayor precio histórico.
Esto servirá para tener una visión estadística acerca del alcance que tuvo, su precio y además de saber por medio de un análisis cuales fueron las razones que ayudaron a que esta criptomoneda se vuelva conocida a nivel masivo en un corto periodo de tiempo.

3. Datasets

En el presente apartado se adjuntará el enlace del conjunto de datos o Dataset que se empleara para realizar el análisis del corriente trabajo:

Proyecto_Bitcoin.xls

SEGUNDA PARTE

4. Objetivos del proyecto

El presente proyecto tiene como objetivo resaltar de forma comparativa, la dimensión de la fluctuación del Bitcoin en el periodo definido 2017 - 2022.
Clasificando las mismas a fin de que dicho análisis sirva (o sea aplicable) en casos de medición de impactos de las transacciones para futuras correcciones u observaciones.

5. Alcance

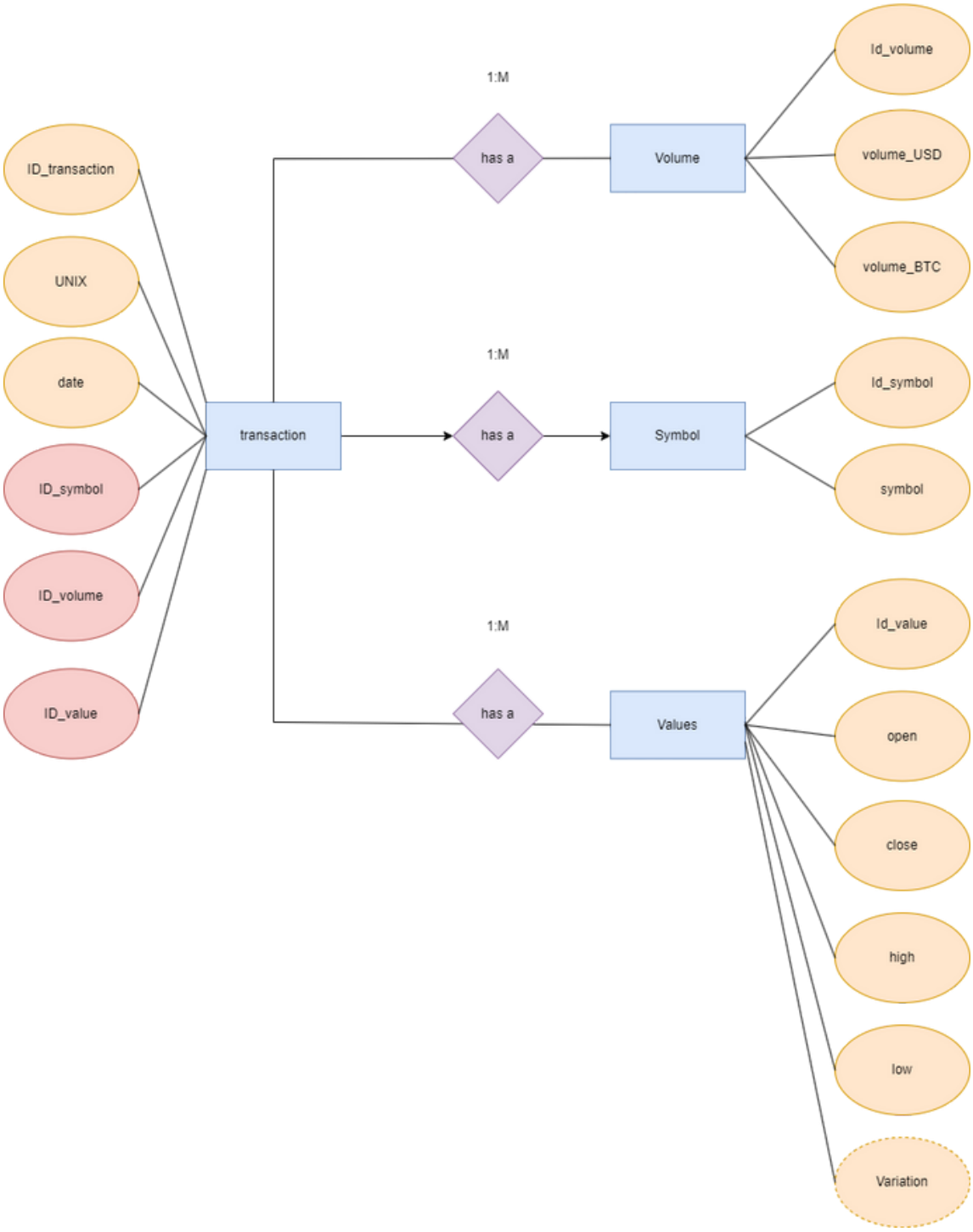
El corriente proyecto describe el desarrollo de un informe desarrollado con la herramienta de Microsoft Power BI, el cual estará conectado con un base de datos trabajada en SQL Server que cuenta con un total aproximado de 500 mil registros de transacciones confirmados por mes, dentro de un período de 5 meses, contando con un total aproximado de más de 1.500.000 operaciones.
Estas serán tratadas y desarrolladas con el fin de obtener gráficas que transmitan la información necesaria para alcanzar el objetivo formal del proyecto.

6. Usuario final

Se tiene como usuario final a: inversores, traiders, estadistas, corredores de bolsa, empresas, economistas o cualquier tercero que tenga interés en el asunto ya sea para fines economicos o simplemente curiosidad.
Los resultados del análisis son a un nivel mundial, ya que cualquier persona con internet y una billetera virtual (**wallets**), podía comprar y vender esta criptomoneda, por medio de un **broker** (es una entidad financiera autorizada y regulada que ejecuta las órdenes de los traders en los mercados. También es responsable de la seguridad de los fondos que pertenecen a sus clientes).
Se aplicará un análisis tanto descriptivo como diagnostico exponiendo información por medio de gráficos relacionados.

7. Diagrama de entiedad-relación

(*) En Bitcoin, cada bloque contiene una marca de tiempo tipo Unix. Esta marca sirve como fuente de variación en el hash del bloque y evitar su manipulación.



8. Listado de tablas

En esta sección, se brinda una descripción de cada tabla disponible en la base de datos a trabajar, incluyendo cuales son sus claves primarias y secundarias:

- **Transaction:** Contiene la información de la transaccion realizada, como fecha, volumen, valor, etc.
  PK: Id_transaction
  FK: id_symbol
  FK: id_volume
  FK: id_value

- **Volume:** Contiene la información del volumen de la transaccion, diferenciada en volumen de dolares y volumen de bitcoin.
  PK: Id_volume

- **Symbol:** Contiene la información acerca del simbolo al cual vamos a hacer la equivalencia de bitcoin.
  PK: Id_symbol

- **Values:** Contiene la información del valor de la transferencia, como por ejemplo: su precio más alto y más bajo, al igual que su precio de apertura y de cierre.
  PK: Id_value

9. Listado de columnas

A continuación se presentará, por tabla, los diversos campos con su tipo de datos correspondiente que las componen:

Tabla: SYMBOL

Nombre del campo	Primary Key	Candidate Key	Foreing Key	Tipo de dato
Id_symbol	X			Int
Symbol				varchar(45)

Tabla: TRANSACTION

Nombre del campo	Primary Key	Candidate Key	Foreing Key	Tipo de dato
Id_transaction	X			Int
Unix				Int
Date				date
Id_symbol			X	Int
Id_volume			X	Int
Id_value			X	Int

Tabla: Volume

Nombre del campo	Primary Key	Candidate Key	Foreing Key	Tipo de dato
Id_volume	X			Int
Volume_USD				Float
Volume_BTC				Float

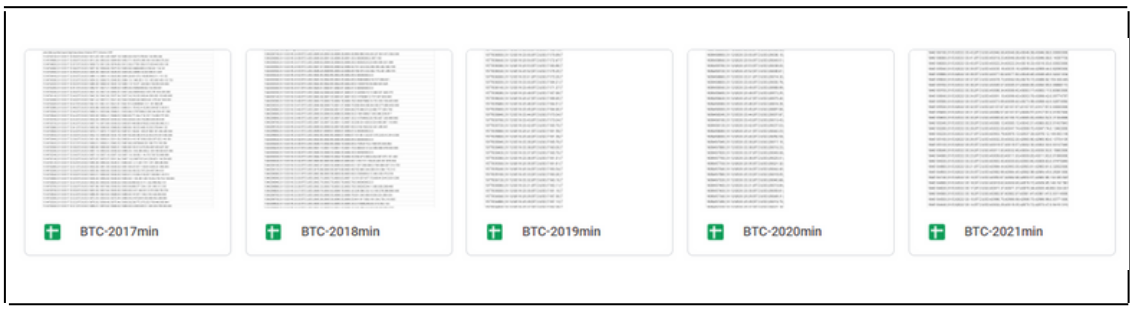
Tabla: VALUES

Nombre del campo	Primary Key	Candidate Key	Foreing Key	Tipo de dato
Id_value	X			Int
Open				Float
Close				Float
High				Float
Low				Float
Variation				Float

TERCERA PARTE

10. Transformaciones realizadas

Para el diseño de la base de datos, la cual originalmente se encontraba con toda la información en una sola tabla extensa en formato Excel y separada en 5 partes iguales, dividida por los años (2017, 2018, 2019, 2020, 2021), se veía de la siguiente manera:



Para el diseño de la base de datos que se vinculó con el proyecto de power BI (.pbix), lo primero que hice fue modificar la tabla original, para de esa forma dividirla según el modelo de "Entidad - relación", que se hizo en puntos anteriores del proyecto, con la finalidad de normalizar los datos.

Se tuvo que realizar cambios en la tabla, por ejemplo: "BTC-2017min", la cual contenía:

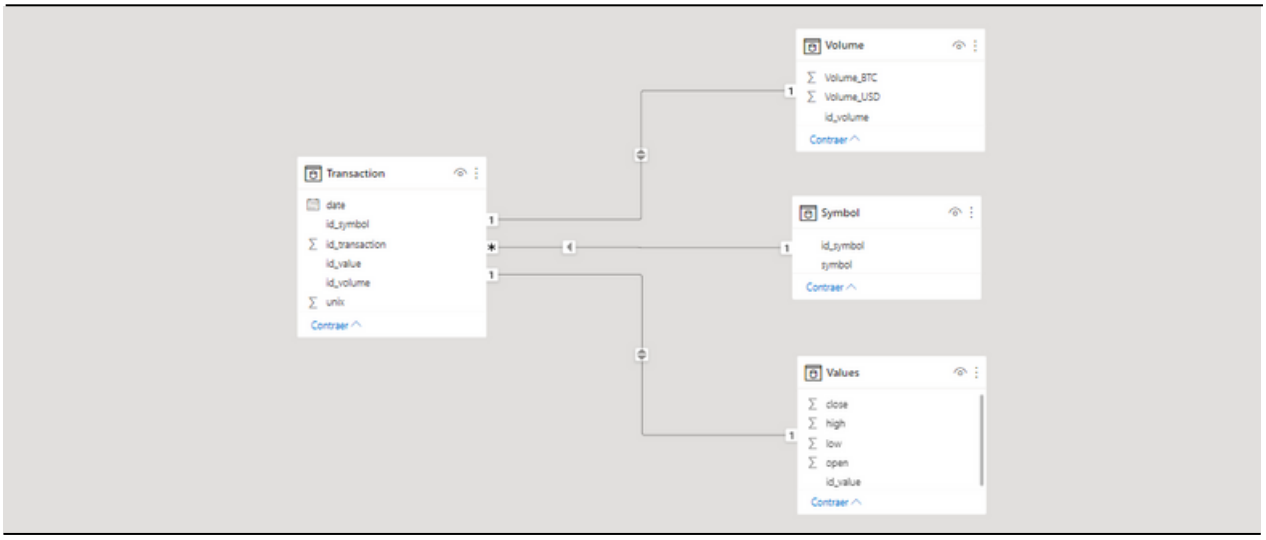
- unix;
- date;
- symbol;
- open;
- high;
- low;
- close;
- Volume BTC;
- Volume USD

Y crear con estos atributos, las tablas:

- transaction
- Symbol
- Volume
- Values

A cada nueva tabla le asigne un identificador único, para que cada transacción este vinculada correctamente con su símbolo, su volumen y su valor correspondiente.

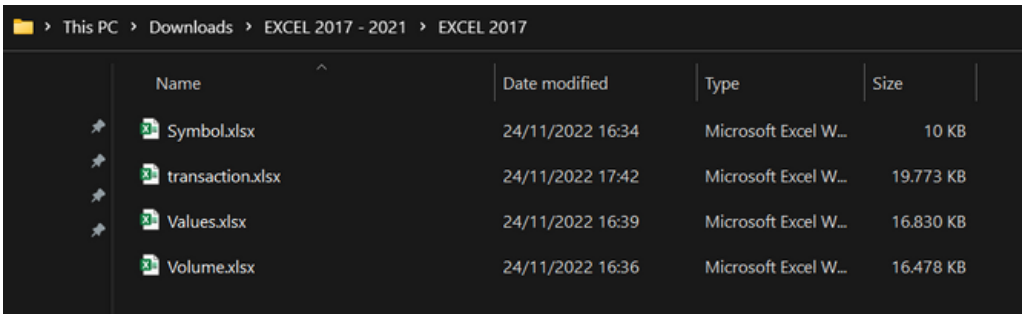
Quedando las relaciones de esta manera:



El mismo procedimiento lo realice en el resto de las tablas.

Para dichas modificaciones se empleó **Excel** por medio de la herramienta de datos y analisis de datos, realizando en primera instancia la segmentacion de los datos y dandole formato de tabla, para que de esa manera sea más fácil la manipulación de los mismos.

Dejando (por ejemplo) a "BTC-2017min" segmentada de la siguiente forma:



Y (por ejemplo) la tabla "transaction" se veia de la siguiente forma:

id_transactio	unix	date	id_symbol	id_volume	id_value
1	1514764740	31/12/2017 23:59	1	1	1
2	1514764680	31/12/2017 23:58	1	2	2
3	1514764620	31/12/2017 23:57	1	3	3
4	1514764560	31/12/2017 23:56	1	4	4
5	1514764500	31/12/2017 23:55	1	5	5
6	1514764440	31/12/2017 23:54	1	6	6
7	1514764380	31/12/2017 23:53	1	7	7
8	1514764320	31/12/2017 23:52	1	8	8
9	1514764260	31/12/2017 23:51	1	9	9
10	1514764200	31/12/2017 23:50	1	10	10
11	1514764140	31/12/2017 23:49	1	11	11
12	1514764080	31/12/2017 23:48	1	12	12
13	1514764020	31/12/2017 23:47	1	13	13
14	1514763960	31/12/2017 23:46	1	14	14
15	1514763900	31/12/2017 23:45	1	15	15

Para realizar las relaciones de las tablas utilicé la herramienta de **Power Bi**, en la cual primero tuve que cargar los datos, luego realizar una limpieza de los mismos y por último fijarme que los valores estén cargados con su "tipo de dato" correspondiente, para de esa manera evitar cualquier tipo de error a la hora de realizar los gráficos.

Otro paso muy importante que tuve que hacer es unir las tablas, ya que al estar dividas en años, tenía muchas tablas con la misma información, pero en tablas diferentes, esta unión la realice en Power Bi, de la siguiente forma:

Carga y transformación de datos:

-2017-

```
= Table.TransformColumnTypes(Symbol_Table,{{"id_symbol", Int64.Type}, {"symbol", type text}})
```

	id_symbol	symbol
1	1	BTC/USD

```
= Table.TransformColumnTypes(Transaction_Table,{{"id_transaction", Int64.Type}, {"unix", Int64.Type}, {"date", type datetime},
```

	id_transaction	unix	date	id_symbol	id_volume	id_value
1	1	1514764740	31/12/2017 23:59:00	1	1	1

```
= Table.TransformColumnTypes(Values_Table,{{"id_value", Int64.Type}, {" open ", type number}, {" high ",
```

	id_value	open	high	low	close
1	1	13913,28	13913,28	13867,18	13880

	id_volume	Volume_BTC	Volume_USD
1	1	0,59	82134565492

La tabla "volume" me dio error en la columna de "volume_usd" y "volume_btc", ya que el volumen de dólares por transacción superaba el límite que se podía poner en power bi, entonces tuve que dividir el volumen, por un billón (1000000000000), con el fin de que el número se reduzca.

	id_volume	Volume_BTC	Volume_USD
1	1	0,59	82.134.565.492,00
2	2	139878396	195.183.096.575.292,00
3	3	0,78	107.829.442.939.156,00
4	4	0,67	9.236.841.134.241,00
5	5	0,07	90.556.013.007,00
6	6	481070190	66.648.811.119.132,00
7	7	1046953115	Error
8	8	715371462	Error
9	9	0,6	82.744.638.901,00
10	10	0,57	Error
11	11	316353009	Error
12	12	605440980	Error
13	13	0	121.181.868.281,00

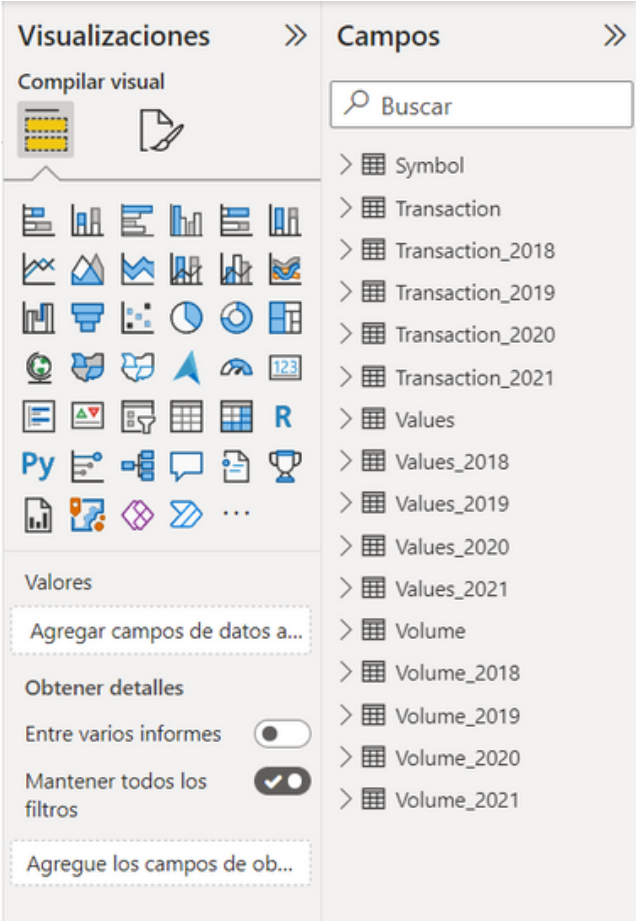
Expression.Error: El número está fuera del rango del valor de moneda. Los valores válidos son de -922337203685477.5808 a 922337203685477.5807.
Detalles:
[Function]

```
=[@[ Volume_USD ]]/ 1000000000000
```

B	C	D
Volume_BTC	Volume_USD	Volume_USD en millones
$ 139.878.396,00	$ 195.183.096.575.292	$ 195
$ 0,78	$ 107.829.442.939.156	$ 108
$ 0,67	$ 9.236.841.134.241	$ 9

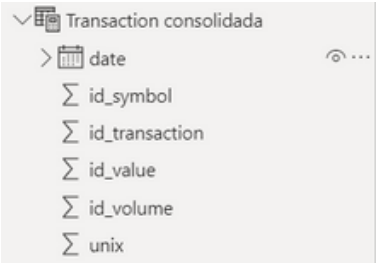
El mismo procedimiento lo realice con el resto de tablas, hasta terminar de cargar todos los datos de la muestra en Power BI.

Antes de unir las tablas se veían de esta forma:



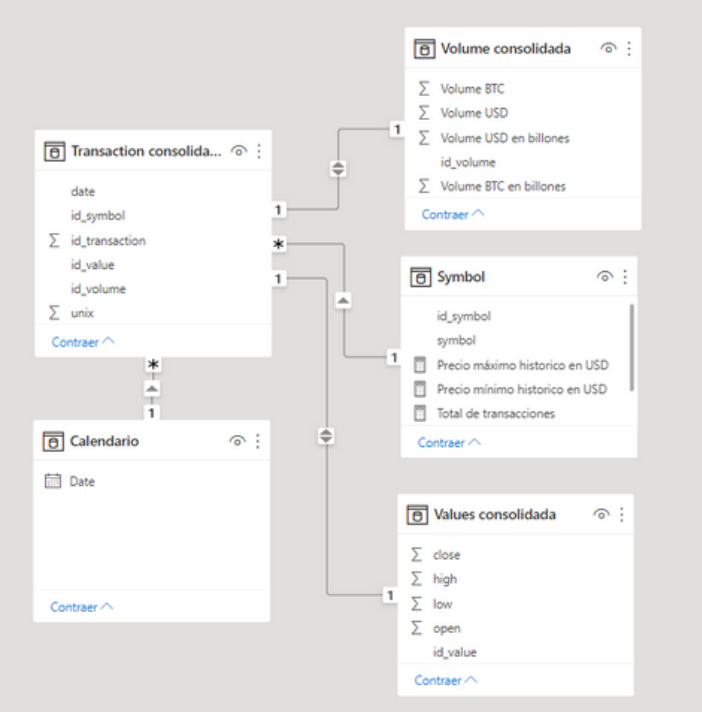
La union de (por ejemplo), la tabla "transaction", la hice en Power BI, de la siguiente forma:

```
1 Transaction consolidada = UNION('Transaction', 'Transaction_2018', 'Transaction_2019', 'Transaction_2020', 'Transaction_2021')
```



*) De esta manera uní las transacciones de todos los años, en una sola tabla de transacciones consolidadas.
Este mismo procedimiento lo realice con el resto de las tablas.
Con la finalidad de tener toda la información consolidada.

De esta manera nos quedarian las tablas consolidadas, con todos los datos de los años correspondientes al proyecto.



11. Medidas generadas y sus fórmulas.

A continuación se remarcaran las diversas medidas desarrolladas y los campos empleados para desarrollarlas:

Medida	Detalle	Descripción
Union de tabla "Transaction"	Transaction consolidada = **UNION**('Transaction', 'Transaction_2018', 'Transaction_2019', 'Transaction_2020', 'Transaction_2021')	Campo creado para unir las transacciones de todos los años en una misma tabla.

Medida	Detalle	Descripción
Union de tabla "Values"	Values consolidada = **UNION**('Values', 'Values_2018', 'Values_2019', 'Values_2020', 'Values_2021')	Tabla creada para unir los valores de todos los años en una misma tabla.
Union de tabla "Volume"	Volume consolidada = **UNION**('Volume', 'Volume_2018', 'Volume_2019', 'Volume_2020', 'Volume_2021')	Tabla creada para unir los volumenes de todos los años en una misma tabla.
Total de transacciones (Todas)	Total de transacciones = **DISTINCTCOUNT**('Transaction consolidada'[id_transaction])	(Dax) creado para expresar la cantidad total de transacciones captados en la muestra
Volumen BTC (en billones)	Volumen de BTC en billones = **DISTINCTCOUNT**('Volume consolidada'[Volume BTC en billones])	(Dax) creado para expresar el volumen DE bitcoin total de transacciones captadas en la muestra

Medida	Detalle	Descripción
Volumen USD (en billones)	Volumen de USD en billones = **DISTINCTCOUNT**('Volume consolidada'[Volume USD en billones])	(Dax) creado para expresar el volumen total dólares de las transacciones captadas en la muestra.
Precio máximo historico (en USD)	Precio máximo historico en USD = **MAX**('Values consolidada'[high])	(Dax) creado para expresar el precio máximo historico en dólares.
Precio mínimo historico (en USD)	Precio mínimo historico en USD = **MIN**('Values consolidada'[low])	(Dax) creado para expresar el precio mínimo historico en dólares.
Año actual (en USD)	Año actual = **SUM**('Values consolidada'[high])	(Dax) creado para expresar la sumatoria del precio de bitcoin en corriente año en dólares.

Medida	Detalle	Descripción
Año pasado (en USD)	Año pasado = **CALCULATE**([Año actual], **DATEADD**(Calendario[Date],-1,YEAR))	(Dax) creado para expresar la sumatoria del precio de bitcoin del período anterior en dólares.
fecha maxima	fecha maxima = **MAX**('Transaction consolidada'[date])	Este calculo me sirvio para la creación tabla calendario.
fecha mínima	fecha minima = **MIN**('Transaction consolidada'[date])	Este calculo me sirvio para la creación tabla calendario.
Calendario	Calendario = **CALENDAR**(DATE(2017,01,01), DATE(2022,03,01))	Creación de tabla calendario tomando la fecha de inicio de operaciones y la ultima fecha en la que se registró una operación.

Medida	Detalle	Descripción
Precio alto	Precio alto = **SUM**('Values consolidada'[high])	(Dax) creado para expresar la sumatoria del precio "alto"de bitcoin del corriente período en dólares.
Precio bajo	Precio bajo = **SUM**('Values consolidada'[low])	(Dax) creado para expresar la sumatoria del precio "bajo"de bitcoin del corriente período en dólares.
Precio de apertura	Precio de apertura = **SUM**('Values consolidada'[open])	(Dax) creado para expresar la sumatoria del precio "apertura" de bitcoin del corriente período en dólares.
Precio de cierre	Precio de cierre = **SUM**('Values consolidada'[close])	(Dax) creado para expresar la sumatoria del precio "cierre"de bitcoin del corriente período en dólares.

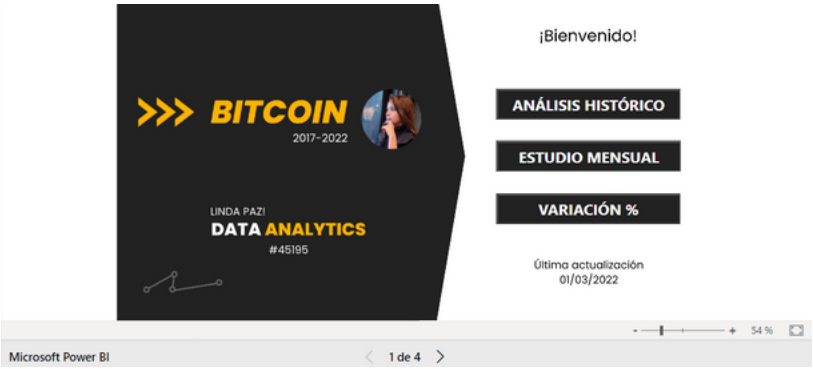
Medida	Detalle	Descripción
Variacion %	Variacion % = **DIVIDE**([Año actual]-[Año pasado],[Año pasado])	(Dax) creado para expresar la variación porcentual de bitcoin del corriente período en comparación al periodo anterior en dólares.

CUARTA PARTE

12. Análisis funcional del tablero:

En este apartado se desarrollará una explicación para cada solapa presente en el proyecto.

Para poder ver el proyecto en la web, pueden usar el siguiente link:



Proyecto_bitcoin2017

En este apartado se desarrollará una explicación para cada solapa presente en el proyecto.

El tablero está dividido en tres solapas:

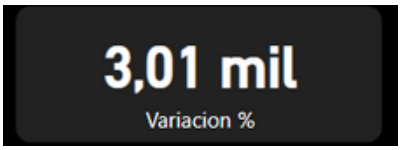
1. **Análisis histórico**
2. **Estudio mensual**
3. **Variación porcentual**

Sección 1: Análisis histórico

En este apartado es un análisis a nivel macro, lo que intente mostrar en él es la fluctuación de la criptomoneda a lo largo del periodo de tiempo que se está analizando en la muestra.

1. El **primer indicador** nos muestra el total de transacciones analizadas.

Esta tarjeta cambia los datos cuando se utiliza el filtro. Mostrando el total de transacciones del período que seleccionemos..



2. El **segundo indicador** el porcentaje total de variación del precio del bitcoin. En esta primera parte no me expando mucho sobre la variación. ya que le dedique una solapa entera a este análisis.

Esta tarjeta cambia los datos cuando se utiliza el filtro. Mostrando el total de transacciones del período que seleccionemos..

3. El **tercer indicador** nos muestra el precio máximo que tuvo el bitcoin, entre el período 2017 - 2022.

Este cálculo concuerda con el de Google.

66,99 mil
Precio máximo historico en USD

751,34
Precio mínimo historico en USD

4. El **cuarto indicador** nos muestra el precio mínimo que tuvo el bitcoin, entre el período 2017 - 2022.

Este cálculo concuerda con el de Google.

¿Cuánto costaba 1 Bitcoin en 2017?
Anexo:Historial de precios de bitcoin

Fecha	USD : 1 BTC
29 de diciembre de 2016	$979
Enero de **2017**	$800 — $1150
5-12 de enero de **2017**	$750 — $920
2-3 de marzo de **2017**	$1290+

5. El **quinto indicador** nos muestra el volumen de bitcoin en billones.

Por definición, el volumen mide la intensidad y el grado de importancia de las tendencias en el mercado.

Su correcto análisis permite entender el origen de los cambios de precio e indica cuándo los inversores deben vender o comprar criptodivisas.

1,51 mill.
Volumen de BTC en billones

6. El **sexto indicador** nos muestra el volumen de dólares en billones.

El volumen de intercambio de criptodivisas en dólares es una herramienta que utilizan los operadores más experimentados.

La bolsa de Bitcoin presenta las últimas cotizaciones, ofreciéndole datos precisos sobre las criptodivisas y los movimientos de esta en los distintos períodos de tiempo.

7. Se emplea un **segmentador de datos** para poder filtrar la información del proyecto según el periódo de tiempo que se desee analizar.

8. La última parte es el **gráfico**, en el que se analiza el total de las transacciones según su monto en dólares por el total del período de tiempo de la muestra.

El período de tiempo esta segmentado por año.

Este gráfico concuerda con el que se muestra en Google, mostrando las mismas curvas, en los mismos periodos de tiempo, y con los mismos montos en dólares.

La unica diferencia es que la que muestra el grafico del tablero se corta en 2022, que es la última fecha que se tiene en la base de datos.

Sección 2: Estudio mensual

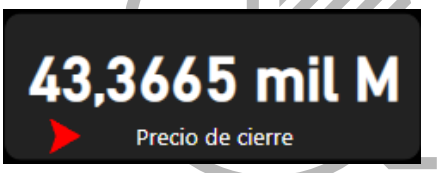
En este apartado se trata de mostrar un análisis más minucioso, apuntando a mostrar información de todos los años, pero también teniendo la opción de dividirlo por mes, ya que en la solapa anterior se mostro un analisis más amplio, apuntando al análisis de la muestra.

1. El **primer indicador** nos muestra el precio de apertura del período.

Según la información de la base de datos, este indicador nos brinda el precio de la última transacción de un día de negociación. (trading)

2. El **segundo indicador** nos muestra el precio de cierre del período.

Según la información de la base de datos, este indicador nos brinda el precio de la primera transacción de un día de negociación. (trading)



3. El **tercer indicador** nos muestra el precio alto del período.

Según la información de la base de datos, este indicador nos brinda el precio mínimo en un día de negociación . (trading)

4. El **cuarto indicador** nos muestra el precio de bajo del período.

Según la información de la base de datos, este indicador nos brinda el máximo en un día de negociación . (trading)

5. Se emplea un **segmentador de datos** para poder filtrar la información del proyecto según el periódo de tiempo que se desee analizar. En este caso tenemos el segmentador por Año.

6. Se emplea un **segmentador de datos** para poder filtrar la información del proyecto según el periódo de tiempo que se desee analizar. En este caso tenemos el segmentador por Mes.

7. Por último contamos con dos **gráficos de barra** los cuales nos indican por un lado la relación entre el precio de cierre y el precio de apertura del bitcoin, por otro lado la diferencia entre el precio más alto y más bajo

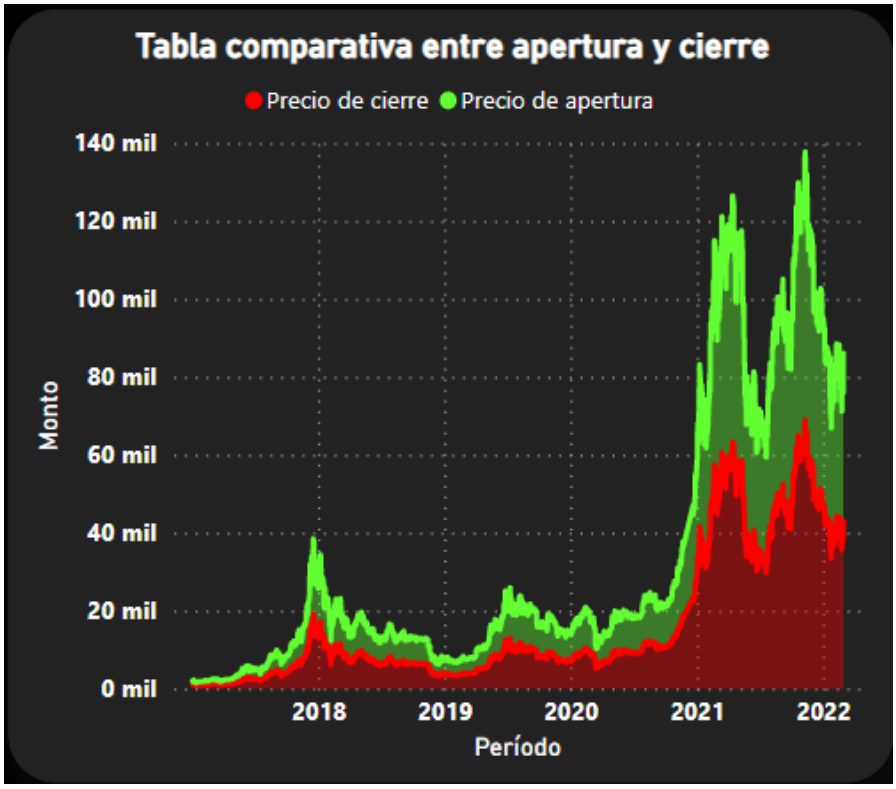
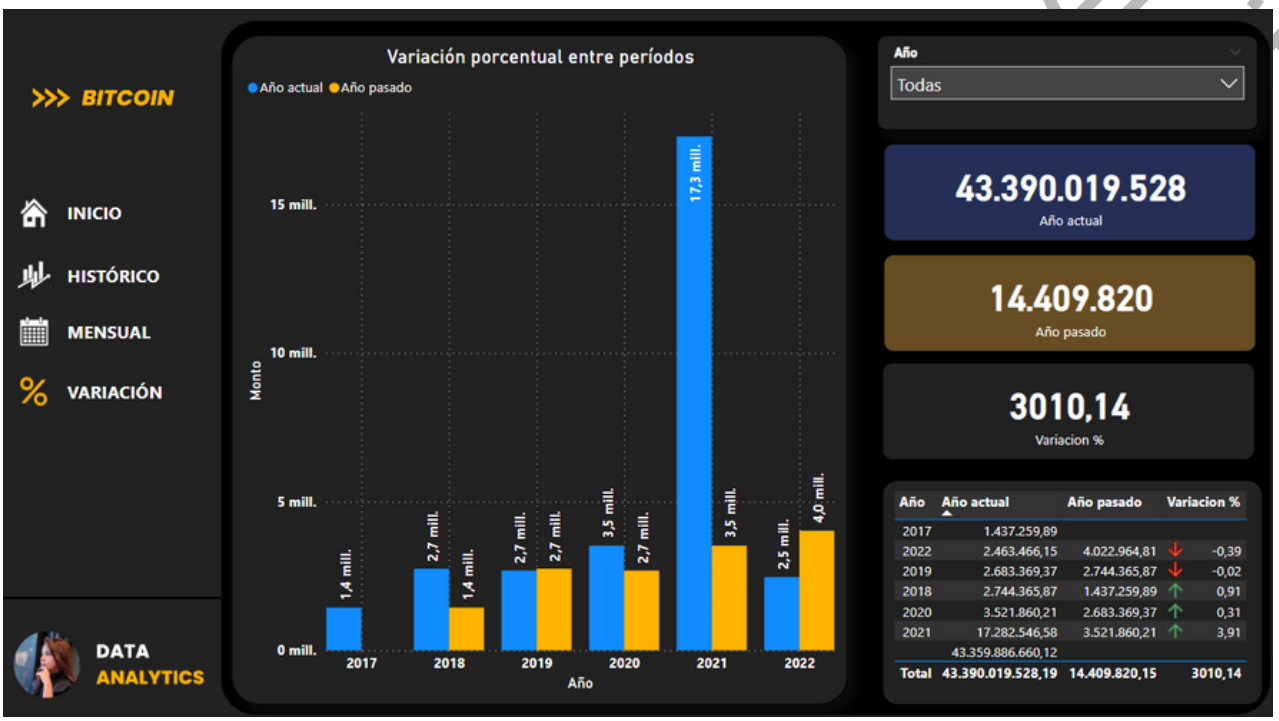


Tabla de relacion entre precio de apertura y cierre medido en tiempo y por monto en USD

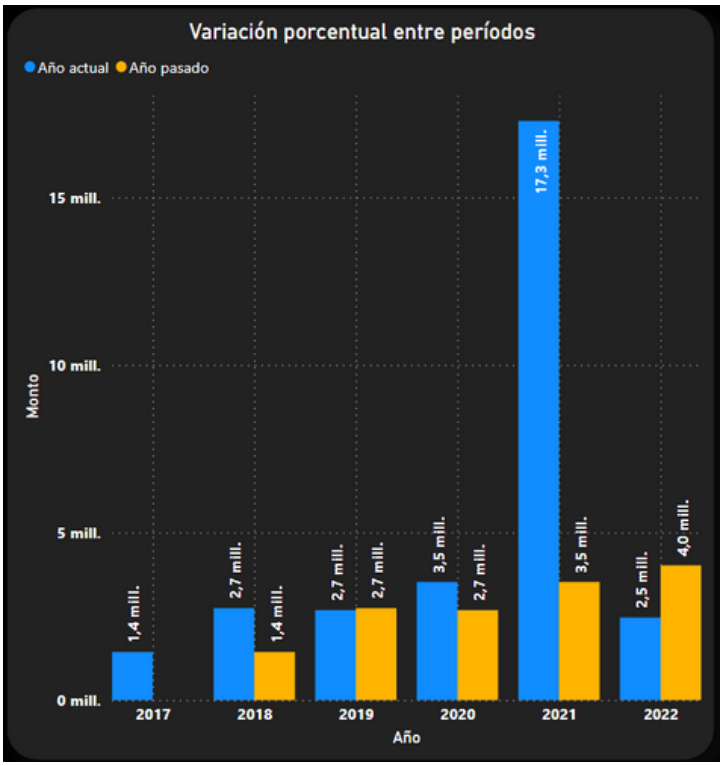
Tabla de relacion entre precio alto y bajo medido en tiempo y por monto en USD

Sección 3: Variación porcentual

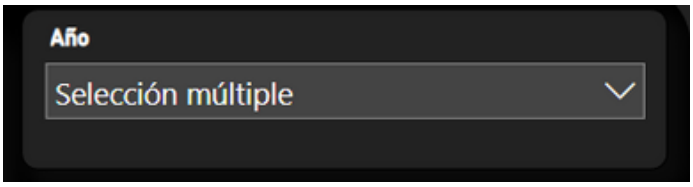


Esta última sección es una analisis a nivel total, donde se compara la variación en porcentaje de un año en comparación con el anterior, para saber cuanto es el incremento o el decremento del mismo.

1. En el lado izquierdo vemos un **grafico de barra** en donde se ve la variacion que tiene el bitcoin por año, mostrando las del año actual de un color **celeste** y de un color **amarillo** el monto del año pasado, de esta menera es sencillo realizar una comparación de ambos.

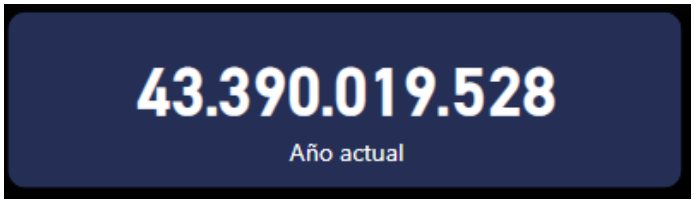


2. En el lado derecho, se colocó un segmentador de datos en forma de lista desplegable.

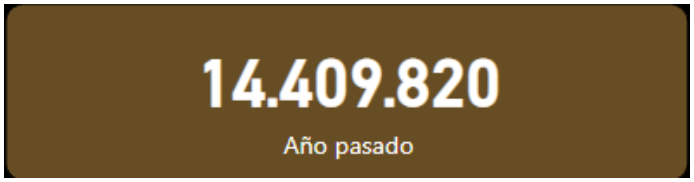


Con la finalidad de analizar un período de tiempo específico.

3. En segundo lugar tenemos una **tarjeta**, en donde se puede ver el monto total en dólares del período actual. El mismo tiene un ligero color azul, para que sea más fácil de compararlo con el gráfico de barra.



4. En tercer lugar también se colocó una **tarjeta**, en donde se puede ver el monto total en dólares del período anterior. El mismo tiene un ligero color amarillo, para que sea más fácil de compararlo con el gráfico de barra.



5. En cuarto lugar se colocó una **tarjeta** que muestra el total de variación porcentual

6. Por último se colocó una **tabla** en donde se puede visualizar la variación que se produjo entre un año y su anterior, sumado a un condicional que muestra una **flecha en verde** para indicar si hubo una subida en comparación al año anterior o si hubo una baja (se ve con una **flecha roja**)

Año	Año actual	Año pasado		Variacion %
	43.359.886.660,12			
2017	1.437.259,89			
2018	2.744.365,87	1.437.259,89	↑	0,91
2019	2.683.369,37	2.744.365,87	↓	-0,02
2020	3.521.860,21	2.683.369,37	↑	0,31
2021	17.282.546,58	3.521.860,21	↑	3,91
2022	2.463.466,15	4.022.964,81	↓	-0,39
Total	**43.390.019.528,19**	**14.409.820,15**		**3010,14**

Mira el proyecto por BEHANCE

behance.net/lindapazw

Futuras líneas

Como futuras líneas de investigación están resumidos aquellos aspectos que resultaron ser potencialmente muy interesantes para desarrollar en trabajos complementarios y que deberán ser analizados próximamente, ya que su alcance escapa a los primariamente planificados.

Además, como ya se mencionara anteriormente, cada nuevo análisis de las ideas aquí expresadas podrá ser fuente de nuevas interpretaciones o soluciones a diversos problemas que hasta aquí no se han presentado y por lo tanto, la lista que se presenta a continuación podrá no incluir a la totalidad de las mismas.

- En un futuro seria relevante la ampliación de la base de datos, con el fin de completar los periodos faltantes en la investigación realizada y los futuros.
- Realizar una solapa en donde se pueda realizar un análisis a nivel diario del bitcoin, donde se vea su variación en porcentaje con el día anterior y la diferencia en volumen de BTC y de USD que tuvo con respecto al día anterior.
- Encontrar la relación que une la magnitud de los corrientes valores de los parámetros y tratar de dar una breve explicación sobre cuáles fueron las cuales de esa fluctuación que provoco la subida o la bajada tan amplia de la criptomoneda en un tan breve periodo de tiempo.