

Informe proyecto Data Analytics

Integrantes: Alan Serraiocco, Alexis Lania, Andrés Muñoz Pampillón.

Objetivo del proyecto:

Divulgar información básica, aspectos clave, brindar ejemplos claros desglosando información que a priori puede ser difícil de procesar, y facilitar la comprensión del mundo de las criptomonedas. Despejando dudas, explicando con ejemplos, y proveyendo consejos de seguridad.

Alcance:

Confeccionar un tablero interactivo que satisfaga los objetivos del proyecto, en el plazo estipulado para la entrega del mismo, utilizando las herramientas y conocimientos adquiridos durante el cursado y la investigación correspondiente.

Descripción de la temática de datos:

Para este proyecto se analizarán las principales criptomonedas, este estudio toma como fuente los datos recolectados por la API de blockchain y un Google Trend.

Cada criptomoneda cuenta con un registro diario desde que se introdujo al mercado de valores digital, el cual nos indica diferentes cotizaciones y volumen de transacciones en el mercado registradas en blockchain. Además, como extra para complementar la cantidad de datos cualitativos se agrega un historial de popularidad en internet según el nivel mensual de búsquedas en Google de cada una de las criptomonedas.

Hipótesis:

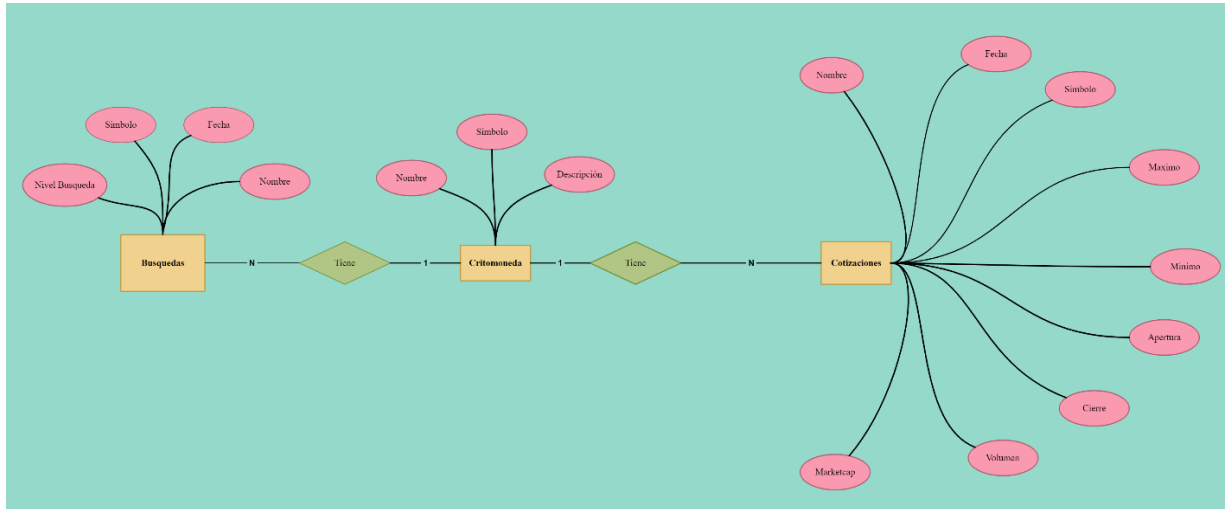
Analizar como varia la tendencia de valores y volumen de transacciones entre las diferentes criptomonedas teniendo en cuenta su popularidad en internet, su registro histórico de precios y como el valor de las principales criptomonedas afecta la tendencia de las otras.

Usuario final y nivel de aplicación del análisis:

El tablero está diseñado para poder ser comprendido por cualquier persona que se interese en saber más de la temática de las criptomonedas por lo que evita el uso de terminología técnica o excesivamente complicada para enfocarse más en la absorción de conocimiento a través de gráficos generando disparadores de curiosidad al poder variar los mismos, incluyendo imágenes y texto para su fácil comprensión. Por lo que se define el nivel de aplicación como informativo o educativo y se quiere evitar completamente que el usuario tome esto como una guía de inversión ya que se pretende despertar la curiosidad en el usuario, no orientarlo a invertir en cierta moneda virtual.

Diagrama entidad relación:

A continuación, se detalla el diagrama entidad relación creado a partir de la base de datos en SQL Server.



Listado de tablas:

A continuación, se mencionarán las tablas creadas en la base de datos, se añadirá una breve descripción y se detallarán sus claves primarias y foráneas si es que tienen.

Índice:

Contienen un listado del nombre, símbolo y una descripción de cada una de las criptomonedas, actúa de tabla maestra en la base de datos.

- **PK:** Nombre

Nivel búsqueda:

Esta tabla provee un nivel de búsqueda en internet según Google Trends para cada criptomoneda mes a mes.

- **PK compuesta:** (Nombre, fecha)
- **FK:** Símbolo.

Cotizaciones:

Provee un registro de diferentes valores de mercado de interés día a día para cada una de las diferentes criptomonedas.

- **PK compuesta:** (Nombre, fecha)
- **FK:** Símbolo.

Listado de columnas por tablas:

A continuación, se mencionará por cada tabla el tipo de dato y clave que tienen las diferentes columnas.

Índice		
Campo	Tipo de dato	Tipo de clave / Constraint
Nombre	varchar(20)	PK
Símbolo	varchar(20)	-
Descripción	Text	-

Nivel Búsqueda		
Campo	Tipo de dato	Tipo de clave / Constraint
Nombre	varchar(20)	PK compuesta
Fecha	nvarchar(50)	PK compuesta
Símbolo	varchar(20)	FK
Nivel Búsqueda	nvarchar(50)	-

Cotización		
Campo	Tipo de dato	Tipo de clave / Constraint
Nombre	varchar(20)	PK compuesta
Fecha	datetime2	Pk compuesta
Símbolo	varchar(20)	FK
Máximo	numeric(37,17)	-
Mínimo	numeric(37,17)	-
Apertura	numeric(37,17)	-
Cierre	numeric(37,17)	-
Volume	numeric(37,17)	-
Marketcap	numeric(37,17)	-

Transformaciones realizadas:

Las transformaciones realizadas a través del editor de consultas de Power Bi sobre el dataset importado desde SQL Server fueron:

- Conversión de tipo long date a short date para las fechas importadas desde SQL server de la tabla vista cotizaciones.
- Conversión de formato long date a formato mm-yyyy para la columna fecha de la tabla vista nivel búsqueda.
- Cambio de numero decimal a formato moneda para los valores de las cotizaciones diarias de la tabla vista cotización

Medidas calculadas generadas y fórmulas:

En este apartado describiremos las medidas calculadas junto a sus fórmulas donde tenemos:

- Creación de la tabla calendario con sus columnas correspondientes.
- Conocer el valor medio entre el máximo y el mínimo
Promedio = (Vista_Cotizacion[Mínimo] + Vista_Cotizacion[Máximo]) / 2
- Medida para tener el día actual, FechaHoy = FORMAT(TODAY(),"dd/mm/yyyy")
- Medida para incorporar en una tarjeta el día actual
Fecha actual = "Fecha actualización "& FORMAT([FechaHoy],"dd/mm/yy ")

- Medida para conocer la fecha de la 1ra cotización de cada criptomoneda
Fecha Inicio = MIN(Vista_Cotizacion[Fecha])
- Obtener el año de la 1ra cotización Inicio = year(Medidas[Fecha Inicio])
- Conocer la fecha de la última cotización Última Fecha = max(Vista_Cotizacion[Fecha])
- Recuperar la primera cotización
Valor Inicio = CALCULATE(min(Vista_Cotizacion[promedio]),
FIRSTDATE(Vista_Cotizacion[Fecha]))
- Recuperar la última cotización Ultimo Valor =
CALCULATE(MIN(Vista_Cotizacion[Promedio]),LASTDATE(Vista_Cotizacion[Fecha]))
- Crecimiento neto entre 1ra y última cotización Neto = ((Medidas[Ultimo Valor] /
Medidas[Valor Inicio]) -1)
- Generar un KPI respecto al neto KPI = if([Neto] >=0, UNICHAR(9650),unichar(9660))
- Conocer la máxima cotización para cada criptomoneda
Maxima Cotizacion = MAX(Vista_Cotizacion[Maximo])
- Conocer la mínima cotización para cada criptomoneda
Minima Cotizacion = MIN(Vista_Cotizacion[Minimo])
- Conocer el máximo valor de apertura para cada criptomoneda
Maxima Cotización Apertura = MAX(Vista_Cotizacion[Apertura])
- Conocer el máximo valor de cierre para cada criptomoneda
Maxima Cotizacion Cierre = MAX(Vista_Cotizacion[Cierre])
- Recuperar la mínima cotizacion de cierre Minima Cotizacion Cierre =
MIN(Vista_Cotizacion[Cierre])
- Obtener el promedio de los valores de cierre Promedio Valor Cierre =
AVERAGE(Vista_Cotizacion[Cierre])
- Obtener el valor promedio de cierre respecto al mes anterior
Promedio Valor Cierre MA = CALCULATE([Promedio Valor
Cierre],DATEADD(Calendario_Cotizacion[Fecha],-1,MONTH))
- También se generó una tabla con las columnas nombre y la imagen asociada al logo de
cada criptomoneda

Segmentaciones elegidas:

Para el filtrado se decidió usar el nombre de las criptomonedas y su logo asociado para que fuera más didáctico y poder seleccionar que moneda deseamos ver en profundidad en las visualizaciones.

Como así también se añadieron barras de tiempo para que el usuario pueda hacer zoom o ampliar el periodo de tiempo que se representa en los gráficos dispuestos por las solapas.

Descripción de las solapas:

A continuación, se mencionarán las solapas que integran la visualización y se hará una descripción de la información que presentan.

Portada

Es la primera solapa que aparece al abrir la el tablero de Power Bi, nos informa tema que se va a tratar a través de su título, la fecha de actualización y una barra de navegación hacia las solapas siguientes la cual se mantiene en las siguientes solapas.

Información

Brinda información importante relacionada al mundo de las criptomonedas las cuales son importantes comprender en especial para la gente que está ingresando en el mismo.

Inicios y últimos registros

Le da al usuario un primer contacto con valores numéricos, aquí podemos encontrar una comparativa histórica respecto a la primera y última cotización de las criptomonedas más importantes, acompañado de un valor porcentual el cual nos muestra cuanto crecieron junto a su KPI para una mejor visualización. También podemos ver en qué año se lanzaron cada una y un top 5 criptomonedas según su marketcap.

Cotización

En esta solapa el usuario puede filtrar por criptomoneda haciendo click al símbolo de la misma. Además, puede seleccionar el rango de años e incluso el mes que quiere analizar, esto ofrece cierta interactividad por parte del usuario para “jugar” con las variables. Dependiendo de los parámetros escogidos se modificarán los gráficos que aparecen en la solapa teniendo:

- Una tarjeta la cual muestra el mayor valor o precio alcanzado por la moneda seleccionada en el periodo escogido.
- Otra tarjeta la cual muestra el mínimo valor alcanzado.
- Un gráfico que nos indica el promedio de los valores de cotización al cierre para los filtros aplicados.
- Un gráfico de barras el cual indica la variación mes a mes de la variación porcentual para esa criptomoneda en el año elegido, mostrándose en verde cuando el precio subido y en rojo cuando bajo mes a mes.
- Una tabla la cual indica en números la cotización promedio de un mes y la del mes pasado para cada mes del año seleccionado.

Volumen

Aquí el usuario mediante la barra de tiempo puede seleccionar el rango de años, con este modificada el grafico de embudo el cual muestra el volumen de compra venta de las criptomonedas.

Además, junto con la selección de las mismas mediante su logo modificara el grafico de líneas que muestra el marketcap por año y criptomonedas pudiendo hacer zoom en diferentes subdivisiones de tiempo como cuatrimestre, trimestre, semestre, semestre, mes y día.

Búsqueda

En esta solapa podemos ver el nivel y los picos de popularidad de las diferentes criptomonedas basados en su nivel de búsqueda en internet a lo largo del tiempo.

Conclusiones

En la solapa final de la presentación se muestra información importante a modo de cierre, junto a ciertas recomendaciones de seguridad que junto a los conceptos aprendidos en las solapas anteriores serán más fáciles de entender y recordar.

Hoja Ruta Proyecto Final Data Analytics

- Como primer paso para definir la temática del proyecto final criptomonedas, se hizo una lectura previa, investigación y un primer contacto mediante links en internet en busca de definir cuestiones conceptuales del tema.
- Después de un primer contacto con el entorno se inició una búsqueda de dataset en fuentes de confianza, llegando a los datos recolectados por la API de Blockchain los cuales cuentan de diferentes archivos .csv que provén diferentes cotizaciones y volumen para cada criptomoneda siendo registro diario desde que se introdujo al mercado de valores digital.
- Debido a que hasta este momento la mayor parte de los datos obtenidos para el trabajo eran del tipo cuantitativo y a que se busca hacer un análisis desde varios enfoques para las diferentes criptomonedas, se optó por añadir otro dataset el cual indica el nivel de búsquedas mensuales para cada una en Google, siendo cada archivo .csv un registro mensual para cada termino de búsqueda. Para esto se usó la herramienta Google Trends teniendo como intervalo de tiempo el mismo que en el otro dataset.
- El siguiente paso fue unificar los diferentes archivos .csv en 2 archivos, uno el cual provee todos los niveles de búsqueda para cada criptomoneda mensualmente y otro con las cotizaciones registradas diariamente para cada criptomoneda. Para ello se usó Excel con el fin de limpiar algunas celdas y eliminar columnas duplicadas.
- Llegado a este punto se visibilizo que para conectar ambas tablas en una base de datos sería necesario una tabla maestra, para confeccionar la misma se usó el nombre, símbolo y descripción de cada criptomoneda extraídos desde la página de blockchain.
- Una vez se tuvieron las tablas listas se procedió a crear la base de datos en Microsoft SQL Server debido a que Excel por la cantidad de filas dejaba de funcionar en muchas ocasiones. Se importaron los datos de los archivos .csv a la base de datos, se añadieron claves primarias, se conectaron las tablas mediante claves foráneas, se le dio un formato adecuado a cada columna junto con restricciones de ser necesarias.
- A la par se hizo el diagrama entidad relación para visualizar correctamente el modelo propuesto.
- Para la realización del tablero lo primero que se hizo fue importar los datos desde la base de datos creada anteriormente en SQL Server.
- Una vez cargadas las tablas se verifico la integridad de los mismos y mediante la edición de consulta se dio el formato necesario a cada columna para después conectar las tablas mediante las relaciones correspondientes.
- Como de antemano se sabe que el trabajo realizado requiere de un análisis temporal se decidió realizar la tabla calendario con las columnas respectivas.
- Una vez la transformación de datos y su carga estuvo correctamente realizada el siguiente paso fue definir el tipo de tablero que se iba a confeccionar, debido a que este es una guía sobre conceptos y análisis de criptomonedas orientado a todo tipo de personas, se bosquejaron el tipo de solapas y la complejidad de conceptos que se desarrollarían así como el nivel de explicación que se daría de cada uno para que la legibilidad estuviera al alcance de un público general siendo el tablero dinámico e interactivo para facilitar la comprensión.
- Para iniciar la confección de cada solapa se analizaron el tipo de graficas que se incluirían en las mismas, dando esto un grupo de variables y medidas que serían de utilidad, las cuales se agruparon todas juntas.

- Una vez se tuvo todo lo necesario para empezar a desarrollar los gráficos estos se hicieron priorizando en una primera instancia su correcto funcionamiento y la inclusión de los filtros necesarios para ello.
- Cuando todas las solapas tuvieron los gráficos insertados con un correcto funcionamiento se dio paso a unificar colores, cambio de fuente y tamaño de texto para su correcta visualización, agregar imágenes y logos, confeccionar una barra de navegación, diseño de una portada e incorporación de tooltips.
- Se confecciona un informe final detallando resultados obtenidos e información importante.

Iniciativas a futuro

Como complemento del proyecto a futuro al ser un tablero informativo para todo público es necesario darle visibilidad. Para ello se publicará el tablero en sitios web relacionados al ámbito financiero, blogs, páginas de ciberseguridad y sitios online que ayuden a divulgar información sobre el mundo de las criptomonedas utilizando el tablero de Power Bi integrado en el código de la página para que los lectores puedan interactuar y obtener los conocimientos que este brinda. Pudiendo así ayudar a gente que desea invertir en el ámbito de las criptomonedas teniendo un primer contacto a través de la información plasmada, visibilizando los hitos importantes a tener en cuenta para detectar un proyecto o criptomonedas fraudulentas.

Conclusión

Lo expuesto a lo largo de este trabajo y el proceso de realización del mismo junto con el de investigación previo permite arribar a las siguientes conclusiones.

1. En cuanto al concepto planteado como hipótesis podemos afirmar que las principales criptomonedas tienen un efecto importante en proyectos emergentes, siendo estos en muchos casos arrastrados por la tendencia del momento. Otros factores que afectan el crecimiento de una criptomoneda es su popularidad en internet donde por razones inherentes al funcionamiento y comercialización de estas se encuentra la mayor cantidad de personas interesadas en ellas como empresas que ofrecen servicios relacionados y donde se publicitan nuevas criptomonedas.
2. Respecto al tablero informativo conformado para ofrecer respuestas a preguntas comunes y enseñar términos específicos a personas que no están familiarizadas con el mundo de las cripto finanzas es claro que es necesario mayor difusión del tema, como la desarrollada en esta ocasión, ya que la falta de la misma genera mucha confusión, también el acceso a esta debe ser interactiva y subjetiva permitiendo prevenir a los posibles inversores caer en proyectos poco confiables.

Siendo esto consecuencia directa del crecimiento en poco tiempo de las criptomonedas lo cual muchas veces hace que la generación de guías e información básica junto a su divulgación no sea lo suficientemente rápida para su correcta temprana llegada a los consumidores de estos servicios.