

UACM

Universidad Autónoma
de la Ciudad de México

Nada humano me es ajeno

Diseño de Data Warehouse sobre la criptomoneda Bitcoin(BTC)

Descripción breve

Diseño de un Data Warehouse,
sobre la criptomoneda Bitcoin()

(13-003-0093)

Ascencio Hernández Tomás

(13-003-0883)

Moreno Navarrete Ixchel Giuseppe

Especialidad de base de datos II
2019-I

ÍNDICE

I-Introducción

II-Alcance

III-Propósito

IV-Objetivos

V-Análisis

VI-Análisis de requerimientos

VII-Análisis OLTP

VIII-Nivel de granularidad

IX-Modelo conceptual ampliado

X-Modelo lógico del Data Warehouse

XI-Unión

XII-Integración de datos

XIII-Consultas

XIV-Conclusiones

XV-Manual De la Aplicacion

I-Introducción

En la actualidad lo más común es que empresas tengan sistemas heterogéneos de aplicaciones, sistemas operativos, bases de datos, hardware y software, con ello, se sabe que entorno a las aplicaciones, su naturaleza es integrar un determinado manejador de base de datos.

Sin embargo, para el usuario final esto es completamente transparente y así como el negocio que forman las empresas. De tal manera que están soportadas por Bases de datos distribuidas, estructuradas o NoSQL e incluso el origen de la información que se procesa no necesariamente su fuente es una base de datos. Lo cual es importante no solo por de donde proviene, ni donde se genera, si en donde almacena dicha información.

Se sabe que anteriormente el enfoque de los primeros sistemas de bases de datos era operativo, es decir, procesamiento transaccional. Sin embargo, en los últimos años ha surgido una noción más sofisticada de la base de datos, la cual responde a las necesidades operacionales y otra que responde a necesidades Informativas o Analíticas. Pero con la evolución del propio Internet, o el surgimiento de Bases de Datos que rompen los esquemas tradicionales de cómo gestionar la información, surge la necesidad de Almacenamiento de todo ese cumulo de información de origen diverso.

Según Bill Inmon, conocido como “El padre del concepto de Data Warehouse” los datos, la tecnología e incluso los usuarios para el procesamiento

operacional son totalmente de los datos, la tecnología y usuario para soportar necesidades informativas o analíticas, es decir, la misma información almacenada puede ser útil para realizar cálculos o generar reportes y para otros usuarios les será útil para la toma de decisiones.

Siendo un Data Warehouse o “Almacén de datos”, un depósito de información reunida de varios orígenes, almacenada bajo un esquema unificado en un solo sitio. Obteniendo datos actuales e históricos de varias fuentes ya sea documentos de texto, videos, audios, gráficos, etc., a través de una organización y el mundo exterior, para facilitar el reporte y análisis para los usuarios del negocio.

Por ello, este proyecto se realizará el desarrollo de un Data Warehouse en el cual se permita obtener reportes y análisis para el negocio, entorno a criptomoneda Bitcoin estudiando su comportamiento entorno al paso del tiempo;

II-Alcance

Dado a que únicamente pudimos obtener los DataSet de la criptomoneda Bitcoin, se determine que los principales aspectos que cubrirá nuestro DW será sobre los valores máximos y mínimos así como el volumen de transacciones y la depreciación de cada o ganancia respecto al anterior día mes o año;

III-Propósito

Construir un Data Warehouse que contenga la información sobre la criptomoneda Bitcoin. Para realizar estudios de alto nivel, entorno a diferentes eventos que se susciten para obtener la mayor información de cambios, en valor y tiempo.

Se utilizará un dataSet que proporciona Investing, tomando como referencia estos datos se proporcionaran los reportes y análisis.

IV-Objetivos

- Construir un DW de la criptomoneda Bitcoin usando un DataSet de Investing
- Generar reportes de comportamiento en cuestión del tiempo
- Analizar los reportes, por medio de estadísticas que se puedan obtener de estos datos y obtener conclusiones sobre ellos
- Usar la metodología Hefesto para documentar el proceso para el DW

V-Análisis

Para construir el Data Warehouse se utilizaran los datos que proporciona el sitio Investing, usando el manejador de base de datos Oracle en su versión 11g el cual se comportara como se explica a continuación

Siendo el producto criptomoneda Bitcoin a comprar, generando reportes de caída

estabilidad e incremento. Se planea extraer estudios estadísticos como la media el promedio el valor máximo y mínimo, así como el volumen transaccional, siendo estas nuestras métricas, para obtener las diferentes dimensiones de la base de datos.

Para comenzar el proceso se usará los pasos que proporciona la metodología Hefesto, para la documentación de este proceso, con ello se desea establecer un orden de en el equipo de trabajo, para obtener un procedimiento a seguir, esta metodología cuenta con las siguientes características:

- Los objetivos y resultados esperados en cada fase se distinguen fácilmente y son sencillos de comprender.
- Se basa en los requerimientos de los usuarios, por lo cual su estructura es capaz de adaptarse con facilidad y rapidez ante los cambios en el negocio.
- Reduce la resistencia al cambio, ya que involucra a los usuarios finales en cada etapa para que tome decisiones respecto al comportamiento y funciones del DW.
- Utiliza modelos conceptuales y lógicos, los cuales son sencillos de interpretar y analizar.
- Es independiente del tipo de ciclo de vida que se emplee para contener la metodología.
- Es independiente de las herramientas que se utilicen para su implementación.

- Es independiente de las estructuras físicas que contengan el DW y de su respectiva distribución.
- Cuando se culmina con una fase, los resultados obtenidos se convierten en el punto de partida para llevar a cabo el paso siguiente.
- Se aplica tanto para Data Warehouse como para Data Mart.

1) ANÁLISIS DE REQUERIMIENTOS

- ➡ a) Identificar preguntas
- ➡ b) Identificar indicadores y perspectivas
- ➡ c) Modelo Conceptual

2) ANÁLISIS DE LOS OLTP

- ➡ a) Conformar indicadores
- ➡ b) Establecer correspondencias
- ➡ c) Nivel de granularidad
- ➡ d) Modelo Conceptual ampliado

3) MODELO LÓGICO DEL DW

- ➡ a) Tipo de Modelo Lógico del DW
- ➡ b) Tablas de dimensiones
- ➡ c) Tablas de hechos
- ➡ d) Uniones

4) INTEGRACIÓN DE DATOS

- ➡ a) Carga Inicial
- ➡ b) Actualización

VI-Análisis de requerimientos

Identificar preguntas

Para identificar la pregunta se utilizará y se basará en lo que se quiere llegar con los objetivos descrito anteriormente, en los que se plasma que es lo que se quiere obtener con la construcción del Data WareHouse. De esta manera obtener los reportes necesarios para ello, por tanto, las preguntas que se quieren contestar son:

1.- ¿Cómo se comporta la criptomoneda Bitcoin en un rango de tiempo establecido?

2.- ¿Cuál es el volumen transaccional máximo en un intervalo de tiempo?

3.-¿El valor máximo alcanzado entre los años 2016 y 2019?

4.- ¿El valor mínimo alcanzado entre los años 2016 y 2019?

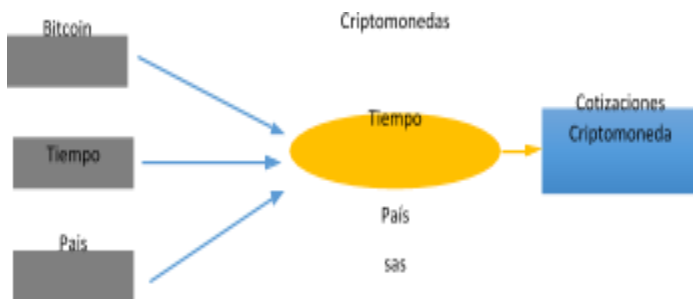
5.- ¿Obtener los promedios en un determinado tiempo?

Identificar indicadores y perspectivas

Para la identificación de indicadores se tienen:

- Criptomoneda
- Tiempo
- País

Modelo Conceptual



VII-Análisis OLTP

Conformar indicadores

Estos primeros indicadores van entorno a las diferentes divisas existentes en la base de datos que se tiene para trabajar.

- “Máximo Bitcoin” vs “Tiempo”
 - Hecho: Máximo de la criptomoneda Bitcoin
 - Función en manejador de BD MAX, realiza una comparativa de los datos hasta encontrar los máximos en un rango de tiempo, por ejemplo, meses, años, trimestres, o semestres.
- “Mínimo Bitcoin” vs “Tiempo”
 - Hecho: Mínimo de la criptomoneda Bitcoin

- Función en manejador de BD MIN, realiza una comparativa de los datos hasta encontrar el mínimo en un rango de tiempo, por ejemplo, meses, años, trimestres, o semestres.
- “Promedio Bitcoin” vs “Tiempo”
 - Hecho: Promedio de la criptomoneda Bitcoin

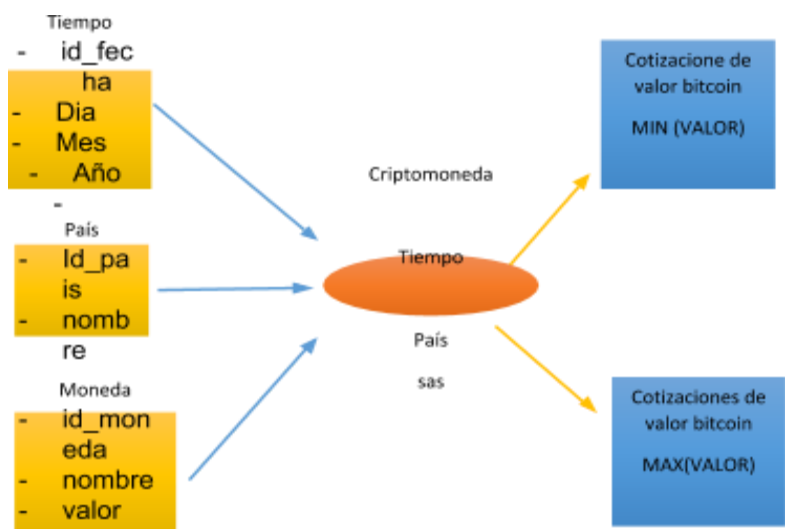
VIII-Nivel de granularidad

Con respecto a la perspectiva Tiempo, es importante definir mediante el cual se agruparán los datos, ya que se puede por semana, mes, día o años, etc. Por tanto, por este proyecto se realizará por año y a lo más mes, por tanto, esta perspectiva, tiene las fechas del 01/01/2016 y columna 2/01/2019.

Para la perspectiva de país, es sumamente importante ya que, al obtener los máximos, mínimos y promedios, se puede obtener un entendimiento como se comportado la criptomoneda en torno a las diferentes criptomonedas.

Moneda, esta perspectiva que se generará basado en el nombre e importancia de la Criptomoneda generando un máximo, mínimo del valor de la criptomoneda.

IX-Modelo conceptual ampliado.



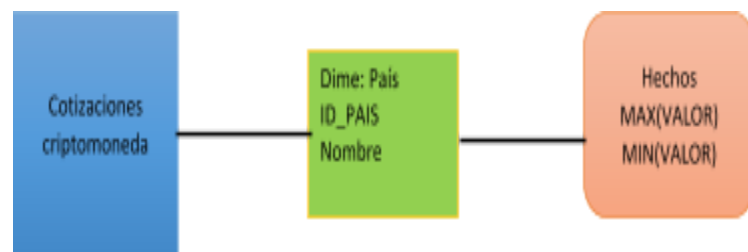
X-Modelo lógico del Data Warehouse

Tipo de modelo lógico del Data Warehouse

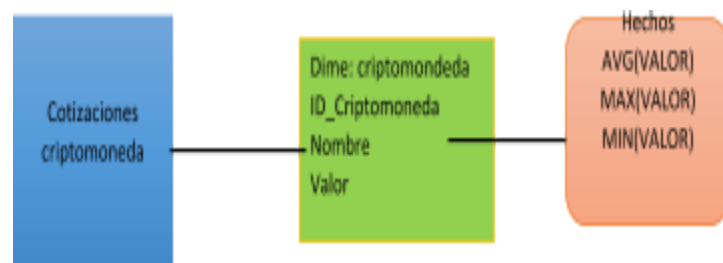
El modelo estrella, consta de una tabla centrar de “Hechos” y varias “dimensiones”, incluida una dimensión de “Tiempo”. Lo que caracteriza a este modelo es que solo existe una tabla de dimensiones para cada dimensión, es decir, que la única tabla que tiene relación con otra es la de hechos, de tal manera que toda la información quede relacionada en una dimensión, o mejor dicho en una sola tabla.

Tabla de dimensiones

- Perspectiva tipo de criptomoneda por país



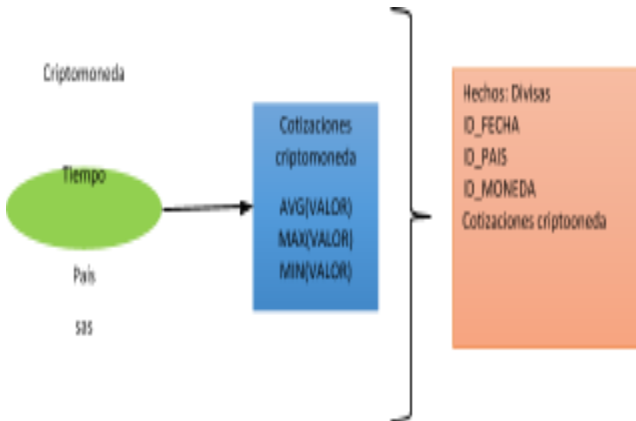
- Perspectiva tipo de criptomoneda por moneda(Bitcoin)



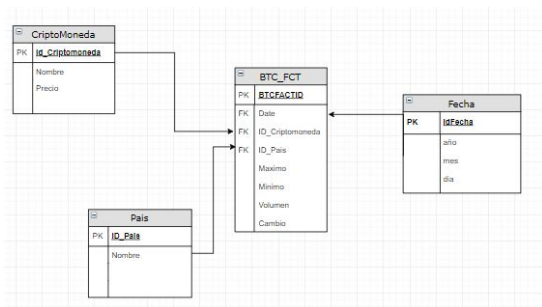
- Perspectiva tipo de divisa por tiempo



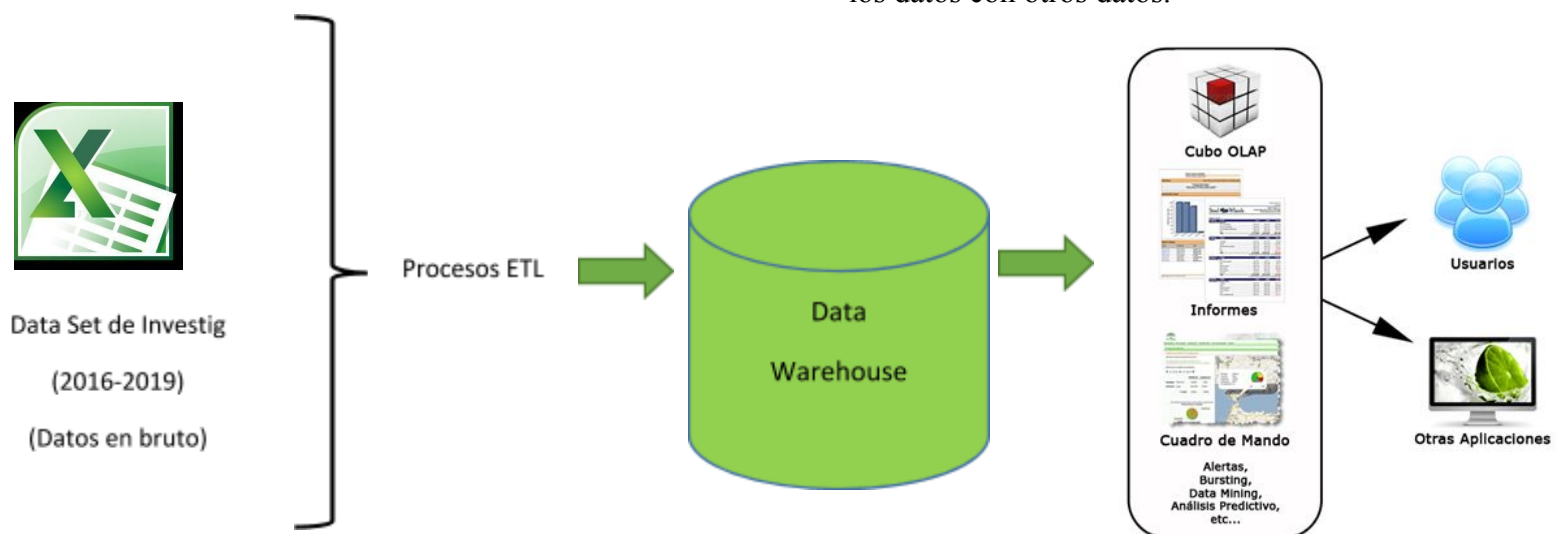
- Tabla de hechos



XI-Unión



Principales componentes del Data Warehouse



Data Warehouse ETL

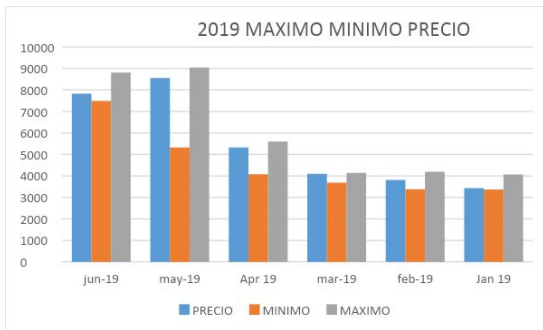
Al proceso de extracción de datos de los sistemas de origen, o fuentes de información, se le denomina comúnmente ETL, en nuestro caso nuestra fuente de información fue el reporte de Banxico.

ETL implica las siguientes tareas:

Extracción de datos: Es lo primero que hace una herramienta ETL. Se trata de obtener la información de las distintas fuentes de origen, tanto internas como externas. Durante la extracción, se identifica los datos deseados y se extrae de muchas fuentes diferentes, incluyendo los sistemas de bases de datos y aplicaciones. Después de la extracción de datos, tienen que ser transportados físicamente al sistema de destino o a un sistema intermedio para su posterior procesamiento y/o transformación.

Transformación: Es el filtrado, limpieza, depuración, homogeneización y agrupación de la información. Incluye la agrupación de los datos de las diferentes fuentes. La transformación se produce mediante el uso de reglas o tablas de consulta o mediante la combinación de los datos con otros datos.

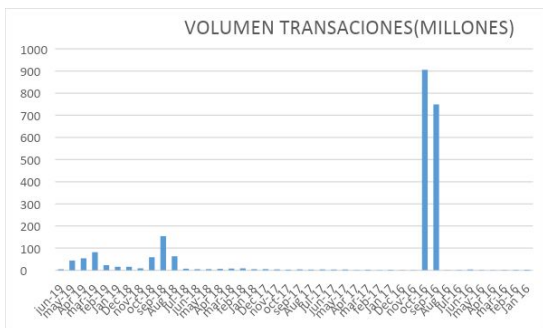
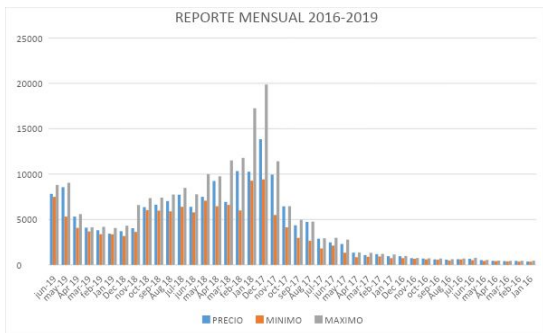
Jun 19	7828.5	7494	8807.7
May 19	8558.3	5319.5	9045.9
Apr 19	5320.8	4076.8	5594.4
Mar 19	4102.3	3681.8	4138.1
Feb 19	3816.6	3382.5	4194.2
Jan 19	3437.2	3368.2	4070.5



en los años 2016-2019 siendo En el mes de Diciembre 2017 donde alcanzó el precio máximo de \$19870.6 dólares

XV-Manual De la Aplicacion

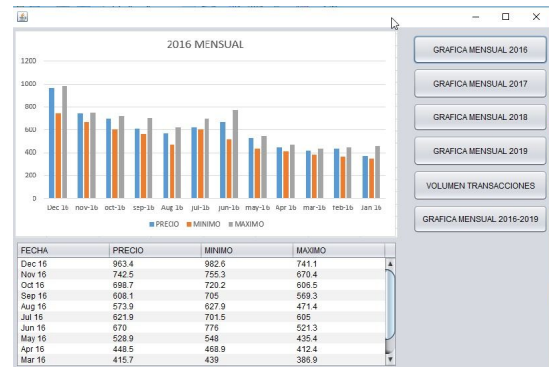
Para el uso de esta aplicación se mostrar con diferentes capturas de pantalla que acción realiza en cada botón de la aplicación mostrando así los datos mensuales de cada año tomando un rango del 2016 al 2019.



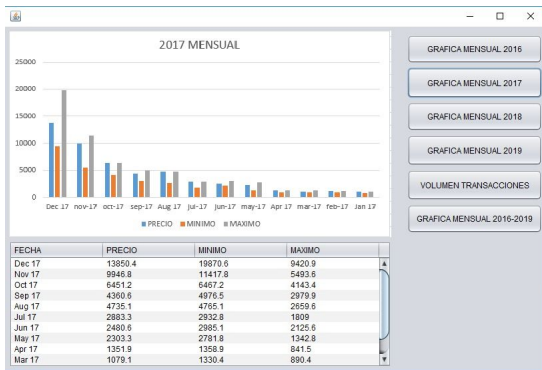
XIV-Conclusiones

La criptomoneda Bitcoin es una moneda muy volátil ya que es descentralizado es decir no está respaldado por ningún gobierno y esto de paso a que su precio sufran muchas especulaciones

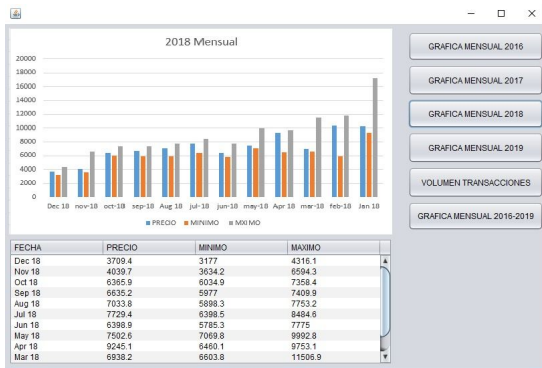
Y con el Data Warehouse nos permite visualizar mas fácil mente el precio minimo que o máximo que llego alcanzar



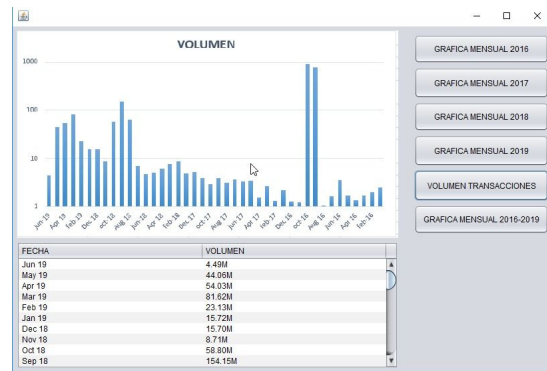
La aplicación al darle click en el primer botón de gráfica mensual 2016 mostrar la en la parte inferior izquierda los datos de ese año la fecha que consta del mes el precio y el valor máximo y mínimo de ese año y mes, al mismo tiempo en la parte superior izquierda se mostrará la gráfica respecto a esos datos del año de los valores ya mencionados.



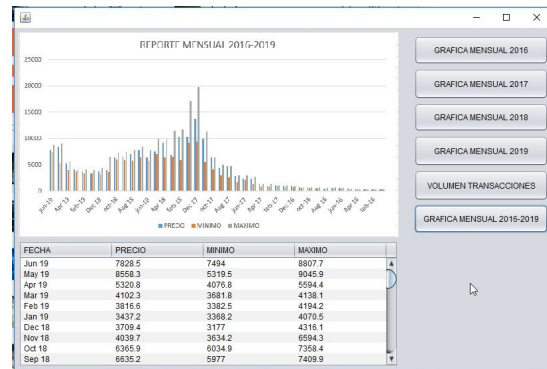
para el botón mensual 2017 serían las mismas acciones del paso anterior.



De la misma manera será para el botón mensual 2018 mostrando sus datos de ese año con su respectiva fecha precio valor máximo y mínimo.y por último este botón del mismo tipo que muestra los datos y la gráfica del año 2019.



El botón volumen transacciones mostrará en la parte inferior izquierda los datos de ese año la fecha que consta del mes y el volumen de cada año y su mes, al mismo tiempo en la parte superior izquierda se mostrará la gráfica respecto a esos datos del año de los valores ya mencionados entre un rango de 0 a 1000.



y por ultimo tenemos el boton grafica mensual 2016-2019 en este boton mostrar los mismo datos como en los primeros que es por año solo que este juntara todos los datos del año 2016-2019 para ver cómo es que se comporta el Bitcoin.