
Análisis de Factores de Correlación en el Precio del Bitcoin

Analysis of correlation factors in the Bitcoin price



Máster en Economía: Instrumentos del Análisis Económico
Facultad de Ciencias Económicas y Empresariales
Universidad de Cantabria

AUTOR: Javier Torre Saiz

TUTOR: Dr. Daniel Diaz Fuentes

04/09/2023

ABSTRACT

We are facing a transformative new landscape for global finance, where Blockchain technology has supported the development of new digital exchange assets, cryptocurrencies, where Bitcoin has emerged and gained traction in recent years as a pioneering digital currency, captivating global attention and challenging traditional conceptions of currency and value. In this paper, we will undertake a comprehensive examination of the complex dynamics governing the valuation of the Bitcoin price. The methodology we will employ to achieve this involves the implementation of a multiple linear regression model, through which we determine the interaction of variables that potentially influence the market price. Our research encompasses a number of relevant independent variables, including mining difficulty, Bitcoin (BTC) mining activity, Bitcoin ATM count, the federal funds rate, the money supply (M1) and the overall inflation rate. These independent variables are chosen for their theoretical ability to impact Bitcoin valuation across multiple economic channels by looking for statistically significant correlations that impact the Bitcoin price. This research is relevant as it contributes substantially to the understanding of the core economic dynamics driving the cryptocurrency ecosystem. By determining the impact of these variables, we provide results that inform economic theories and shed light on the complexities inherent in this growing digital asset landscape. This work has implications, moreover, not only for investors navigating this novel financial asset class; but also, for policymakers seeking to understand the evolving environment of the financial ecosystem in a digital age in order to make socio-economic and investment decisions that ensure greater advantages for private enterprise and society at large.

Keywords: Bitcoin, economic dynamics, digital assets, and multiple linear regression.

RESUMEN

Nos encontramos ante un nuevo horizonte transformador para las finanzas globales, en el que la tecnología Blockchain ha soportado el desarrollo de nuevos activos digitales de intercambio, las criptomonedas, donde el Bitcoin ha surgido y cobrado fuerza en los últimos años como una moneda digital pionera, cautivando la atención mundial y desafiando las concepciones tradicionales de moneda y valor. En este trabajo, vamos a abordar un examen exhaustivo de las complejas dinámicas que rigen la valoración del precio del Bitcoin. La metodología que vamos a emplear para conseguirlo consiste en la implementación de un modelo basado en una regresión lineal múltiple, mediante el cual

se determina la interacción de variables que potencialmente ejercen influencia sobre el precio del mercado. Nuestra investigación abarca una serie de variables independientes pertinentes, que incluyen: la dificultad de minería, la actividad minera del Bitcoin (BTC), el recuento de cajeros automáticos de Bitcoin, la tasa de fondos federales, la oferta monetaria (M1) y la tasa de inflación global. Estas variables independientes son elegidas por su capacidad teórica para impactar en la valoración del Bitcoin a través de múltiples canales económicos buscando correlaciones estadísticamente significativas que impacten en el precio del Bitcoin. Esta investigación es relevante, ya que contribuye sustancialmente a la comprensión de las dinámicas económicas principales que impulsan el ecosistema de las criptomonedas. Al determinar el impacto de estas variables, ofrecemos resultados que fundamentan teorías económicas y arrojan luz sobre las complejidades inherentes a este creciente panorama de activos digitales. Este trabajo tiene implicaciones, además, no solo para inversores que navegan en esta novedosa clase de activos financieros; sino también para formuladores de políticas que buscan comprender el entorno en evolución del ecosistema financiero en una era digital para tomar decisiones socioeconómicas y de inversión que garanticen mayores ventajas para la iniciativa privada y la sociedad en general.

Palabras claves: Bitcoin, dinámicas económicas, activos digitales y regresión lineal múltiple.

ÍNDICE DE CONTENIDOS:

1. INTRODUCCIÓN	1
1.1. CONTEXTO Y MOTIVACIÓN	3
1.2. OBJETIVOS DE LA INVESTIGACIÓN	5
1.2.1. Objetivo General	5
1.2.2. Objetivos Específicos	5
1.3. JUSTIFICACIÓN DE LA ELECCIÓN DE VARIABLES	5
1.4. ESTRUCTURA DEL TRABAJO	6
2. MARCO CONCEPTUAL	8
2.1. FUNDAMENTOS DEL BITCOIN Y SU MERCADO	8
2.2. ANTECEDENTES AL ESTUDIO	10
2.3. FACTORES DE INFLUENCIA EN EL PRECIO DEL BITCOIN	11
3. METODOLOGÍA	13
3.1. REGRESIÓN LINEAL MÚLTIPLE	13
3.2. DEFINICIÓN DE VARIABLES: DEPENDIENTES E INDEPENDIENTES	16
3.3. RECOPIACIÓN Y PROCESAMIENTO DE DATOS	19
3.4. VALIDACIÓN DE DATOS PARA EL MODELO DE REGRESIÓN	20
3.4.1. Homocedasticidad	21
3.4.2. Valores atípicos	21
3.4.3. Test y Train	22
3.4.4. Multicolinealidad	22
4. ANÁLISIS DE RESULTADOS	24
4.1. PRESENTACIÓN DE RESULTADOS DE LA REGRESIÓN	25
4.2. INTERPRETACIÓN DE LOS COEFICIENTES DE REGRESIÓN	27
4.3. VERIFICACIÓN DE LOS SUPUESTOS DEL MODELO	27
4.3.1 Heterocedasticidad	28
4.3.2 Normalidad	29
4.3.3 Multicolinealidad	30
4.3.4 Independencia de los errores	31
4.4. PREDICCIÓN DEL PRECIO DE BITCOIN	32
5. DISCUSIÓN	33
5.1. CONTEXTO DE LOS RESULTADOS EN LA LITERATURA EXISTENTE	33
5.2. IMPLICACIONES PRÁCTICAS DE LOS HALLAZGOS	36
5.3. LIMITACIONES DEL ESTUDIO Y FUTURAS INVESTIGACIONES	38
6. COMPARACIÓN CON OTROS MÉTODOS DE PREDICCIÓN	39

6.1. MODELOS DE APRENDIZAJE AUTOMÁTICO: FLEXIBILIDAD Y RIESGO DE SOBREAJUSTE	40
6.2. ANÁLISIS DE SERIES DE TIEMPO: TENDENCIAS TEMPORALES ...	41
6.3. MODELOS BASADOS EN SENTIMIENTOS Y FUENTES EXTERNAS	42
6.4. REFLEXIONES Y CONSIDERACIONES FINALES	44
7. CONCLUSIONES	44
8. BIBLIOGRAFÍA	46
9. APARTADOS ANEXOS	49
Código R	50

1. INTRODUCCIÓN

En el panorama actual del siglo XXI, surge un fenómeno disruptivo que ha emergido en el ámbito financiero global: las criptomonedas, monedas virtuales que emplean una tecnología descentralizada y criptográfica a la que conocemos como Blockchain, la cual nos abre un nuevo abanico de posibilidades de intercambio económico que garantiza la realización de transacciones financieras basadas en la confianza transaccional brindando seguridad y eficiencia. Uno de sus mayores objetivos es frenar la producción de papel moneda por parte de los gobiernos, siendo teóricamente inmune a la interferencia o manipulación del mismo y eliminando el rol de los intermediarios, por ejemplo, del sistema bancario, en el ecosistema financiero. Echeveste (2018), afirma que este hecho ha significado una importante evolución de las monedas digitales, hasta tal punto que durante el primer semestre de 2018 se llegaron a registrar más de 2.000 tipos de criptomonedas a nivel mundial, aunque una gran cantidad de éstas no lograron ser exitosas debido a que no aportaron grandes innovaciones.

La última gran crisis económico-financiera que estalló en 2008 evidenció la vulnerabilidad de los sistemas bancarios y, sumado a esto, generó un aumento significativo de la desconfianza de la población respecto de las instituciones financieras. Adicionalmente, Cadenas (2021) aporta que los ajustes económicos y las políticas monetarias implementadas por parte de los Estados, algunas todavía vigentes, desencadenaron problemas como, pérdida de riqueza, devaluación de las monedas y problemas de hiperinflación o reducción de beneficios sociales, entre otros. En el año 2008 fue el primer momento en el que se materializó la innovación de un sistema de efectivo descentralizado y criptográfico a través de la criptomoneda Bitcoin publicada de forma anónima por Satoshi Nakamoto (2008), quien afirmó que el comercio electrónico había llegado a depender por completo de las instituciones financieras que actúan como terceros confiables para realizar pagos por internet, recalando que los comerciantes deben pedir una cantidad de información desbordante a sus clientes para garantizarles seguridad, violentando su confianza y, finalmente, debiendo asumir un cierto porcentaje de fraude como inevitable.

El carácter innovador y el potencial disruptivo de las criptomonedas implicó un crecimiento explosivo en el transcurso de 2017 y principios de 2018, momentos en los que los generosos retornos que se generaron provocaron muchas especulaciones sobre los activos de las criptomonedas (Lakomski, 2020). Sin embargo, las monedas digitales no son aceptadas como medio de pago en muchos países. El efectivo implica una serie de leyes fiscales que

regularizan la producción del mismo; mientras que las criptomonedas, funcionan mediante un soporte electrónico descentralizado que no puede ser rastreado, lo cual también ha favorecido movimientos ilícitos en los últimos años, al poder realizarse transacciones de manera anónima, lo que podría generar distorsiones en la economía o incluso alguna especie de burbuja, según (Echeveste, 2018). Este vacío regulatorio, en combinación con la alta y repentina popularidad y la falta de garantía institucional, hace que el mercado de las criptomonedas sea tan volátil que incluso ha llegado a identificarse como el "salvaje oeste".

Pese a que las monedas virtuales no están reguladas por ningún tipo de organización gubernamental, tenemos que informar de que las criptomonedas también suponen para muchos una alternativa financiera, por ejemplo, como forma de intercambio en países con regímenes donde se controla la economía y no se puede acceder a los bienes más básicos (Duin, 2021). Además, algunas criptomonedas tienen una base caritativa y de compromiso con el medio ambiente (Baldrés, 2018). Es así como las criptomonedas en estos años han logrado crear un sistema globalizado para la realización de medios de intercambios transparentes y responsables (Sandoval et. Al, 2020). La prevención, evitando el doble gasto y la garantía de almacenamiento distribuido, garantiza que todos los actores del mercado se protejan del fraude.

En este contexto, nuestro estudio investiga la correlación entre el precio del Bitcoin y una serie de variables identificadas. Los factores que vamos a analizar abarcan la dificultad de minería, la actividad minera de Bitcoin, el recuento de cajeros automáticos de Bitcoin, la tasa de fondos federales, la oferta monetaria (M1) y la tasa de inflación global. Estas variables han sido seleccionadas estratégicamente debido a su potencial teórico para influir en la valoración de la criptomoneda en cuestión a través de intrincados caminos económicos, buscando correlaciones estadísticamente significativas que impacten en el precio del Bitcoin.

Para encontrar los factores influyentes debemos recurrir a enfoques analíticos rigurosos que nos permitan modelar y cuantificar las relaciones subyacentes, para ello, empleamos un modelo de regresión lineal múltiple, una metodología estadística sólida que considera de forma simultánea los efectos de múltiples variables en la variable dependiente. En nuestro caso particular, el precio del Bitcoin. Este modelo se propone cuantificar las relaciones entre las variables independientes y el precio del Bitcoin, permitiéndonos en nuestro avance discernir qué variables tienen correlaciones estadísticamente significativas, contribuyendo así a una comprensión más precisa de su influencia real.

Los hallazgos de nuestra investigación revelan que la dificultad de minería, la actividad minera de Bitcoin y la tasa de inflación global, muestran correlaciones estadísticamente significativas

con el precio del Bitcoin. Estas variables se posicionan como pivotes influyentes que arrojan luz sobre las dinámicas que impulsan la valoración del Bitcoin. Sin embargo, veremos con nuestro análisis que la oferta monetaria (M1) y la tasa de fondos federales no presentan correlaciones estadísticamente significativas con el precio del Bitcoin.

El valor añadido por el modelo de regresión lineal múltiple es significativo, ya que nos permite aislar el impacto de variables individuales cuando consideramos la interacción de múltiples factores, lo que desvela un enfoque que refleja el intrincado ecosistema en el que opera el Bitcoin. Al resaltar cuáles son las variables que influyen de manera más significativa en el precio del Bitcoin, proporcionamos percepciones que soportan teorías económicas y certifican la naturaleza multifacética de los activos digitales. Esta investigación es relevante y se posiciona más allá de los inversores que incursionan en esta novedosa clase de activos, extendiéndose a los formuladores de políticas que buscan comprender el panorama financiero en evolución de la era digital.

Pese a ello, es necesario señalar que nuestros hallazgos indican correlaciones en lugar de causas. Aun habiendo sido capaces de identificar variables pivotes, existen otros factores que también podrían contribuir a la valoración del Bitcoin. Nuestra investigación, no obstante, contribuye de forma significativa a una comprensión más amplia de las intrincadas dinámicas económicas que impulsan el ecosistema de las criptomonedas. Al ser capaces de determinar el impacto de estas variables, tenemos la oportunidad de ofrecer conclusiones que refrendan teorías económicas y desvelan las complejidades que subyacen en este floreciente panorama de activos digitales

1.1. CONTEXTO Y MOTIVACIÓN

En el cambiante horizonte financiero de hoy en día, las criptomonedas han emergido como una nueva frontera de inversión, pero también como una fuente de incertidumbre en los mercados globales. En este contexto, proponernos un estudio exhaustivo de los factores que podrían influir en el precio del Bitcoin, supone un reto apasionante y de enorme interés. La volatilidad inherente y la falta de organismos o métodos tradicionales que respalden estas monedas digitales ha planteado desafíos importantes para inversores, reguladores y analistas en estos tiempos. Es en este cruce entre innovación tecnológica y finanzas tradicionales, es donde surge la motivación para desentrañar los verdaderos elementos que impulsan las fluctuaciones en el precio del Bitcoin. A través de la investigación y el análisis de estas variables independientes, se busca encontrar respuestas sobre los mecanismos económicos, tecnológicos y psicológicos que podrían estar en juego en la alteración del precio de la criptomoneda objeto de estudio. Por ende, es imperativo proporcionar una base sólida para la

toma de decisiones informadas en un entorno cada vez más dinámico, complejo y con más adeptos.

Ante el creciente uso de monedas virtuales y la inminente necesidad de encontrar alternativas eficientes para facilitar las transacciones comerciales a nivel global, surgió un impulso hacia sistemas financieros más ágiles y rentables. Esta tendencia tomó fuerza a partir de 2008, cuando los retos económicos mundiales del momento incentivaron la búsqueda de soluciones que redujeran los tiempos y los costes de transacción, mientras se buscaba que aportaran agilidad y eficiencia en las remesas internacionales. Este cambio se reflejó en el aumento del uso de monedas digitales.

Mencionando los costes de transacción como una de las barreras principales que consigue superar la tecnología Blockchain, *“la Teoría de la economía de costos de transacción en la economía, ha sido un paradigma teórico dominante en el estudio de la gestión y la organización”*, (Chiles y McMackin, 1996 ; David y Han, 2004 ; Williamson, 2002) desde la publicación del libro seminal de Williamson, *Market and Hierarchies* (Williamson, 1975).

La tecnología Blockchain puede reducir significativamente estos costes de transacción para las partes interesadas en el ecosistema empresarial, como empresarios, ángeles y capitalistas de riesgo. La reducción de costes principal se logra al reducir el costo de búsqueda, siendo capaz de eliminar intermediarios de terceros en el sistema. Además, el Blockchain puede abordar muchos de los problemas que obstaculizan el desarrollo de un ecosistema empresarial sólido a nivel local para el desarrollo económico, según (Mahto et al., 2018; Mahto y McDowell, 2018).

Con estos hechos, resulta esencial emprender una exhaustiva investigación sobre el Bitcoin, dado que existen comportamientos y factores regulativos que a menudo no se toman en consideración o se desconocen por parte de los inversores.

Este estudio, no solo ofrece una mejor comprensión de los factores que afectan el precio del Bitcoin, sino que también proporciona una perspectiva más completa y fundamentada para los inversores, reguladores y analistas. El conocimiento profundo de estas variables podría contribuir positivamente a reducir la incertidumbre que rodea a las criptomonedas y, potencialmente, a mejorar su estabilidad en los mercados financieros existentes. En última medida, esta investigación tiene el potencial de beneficiar tanto a los actores individuales como a la totalidad del sistema financiero en su conjunto, al aportar información valiosa, basada en el análisis económico, con el objetivo de favorecer la toma de decisiones

informadas y la navegación exitosa en un entorno financiero innovador y en constante evolución.

1.2. OBJETIVOS DE LA INVESTIGACIÓN

1.2.1. Objetivo General

Nos proponemos llevar a cabo un análisis de los factores que influyen en la variación del precio del Bitcoin, mediante un enfoque basado en la regresión lineal múltiple, con el fin de proporcionar una visión integral de las relaciones económicas y tecnológicas presentes que afectan de manera inherente su valor en los mercados financieros.

1.2.2. Objetivos Específicos

Podemos categorizarlos en cuatro fases esenciales:

- Identificación y recopilación de datos históricos relevantes sobre el precio del Bitcoin y las variables independientes seleccionadas, incluyendo la dificultad de minería, la actividad minera de Bitcoin, el recuento de cajeros automáticos de Bitcoin, la tasa de fondos federales, la oferta monetaria (M1) y la tasa de inflación global desde junio de 2017 a diciembre de 2021.
- La realización de un análisis descriptivo de los datos recolectados para comprender las tendencias, distribuciones y relaciones preliminares entre las variables, estableciendo una base exploratoria para el estudio.
- La aplicación de un modelo de regresión lineal múltiple que utilizamos para cuantificar las relaciones estadísticas entre las variables independientes y el precio del Bitcoin, desentrañando qué factores presentan correlaciones significativas y cómo contribuyen a las variaciones en su valor.
- Interpretación de los resultados del modelo de regresión lineal múltiple, analizando el impacto relativo de cada variable independiente en el precio del Bitcoin y llevando a cabo una discusión de las implicaciones económicas y tecnológicas de estas relaciones para el mercado criptográfico, o de las criptomonedas.

1.3. JUSTIFICACIÓN DE LA ELECCIÓN DE VARIABLES

Las variables independientes que han sido consideradas para el análisis se han seleccionado estratégicamente, dado su potencial inherente para influir en la valoración del Bitcoin mediante una serie de mecanismos económicos. Cada una de estas variables tiene una conexión conceptual con aspectos críticos del ecosistema del Bitcoin y su relación con el mercado financiero.

De esta forma, la dificultad de minería y la actividad minera de Bitcoin están íntimamente relacionadas con las limitaciones de producción y suministro de Bitcoin, lo que puede afectar

directamente a su escasez percibida y, en última instancia, a su valor estimado en el mercado. Por su parte, el recuento de cajeros automáticos de Bitcoin podría señalar la adopción real y el acceso público al Bitcoin, siendo una variable capaz de influir en la percepción y la demanda de la criptomoneda. La tasa de fondos federales, mientras tanto, puede indicar la dirección general de las tasas de interés y las condiciones macroeconómicas, lo que supone un impacto en las preferencias de inversión y asignación de capital hacia activos alternativos como las criptomonedas. Hablando de la oferta monetaria M1 y la tasa de inflación global, diremos que están conectadas con la percepción de la moneda fiduciaria y la idea de preservación del valor en relación con las criptomonedas, lo cual influye potencialmente en la demanda de Bitcoin como valor de resguardo. A través de nuestro enfoque analítico, buscamos desenmascarar la compleja red de interacciones entre estas variables estratégicas y el precio del Bitcoin, con el objetivo de comprender cuáles son los factores determinantes que afectan a su valor y comportamiento en el mercado financiero.

1.4. ESTRUCTURA DEL TRABAJO

A continuación, se presenta la estructura de este trabajo con el objetivo de proporcionar una guía clara y organizada para la comprensión de la investigación que proponemos en torno a los factores que influyen en el precio del Bitcoin. Cada sección está diseñada de manera estratégica para brindar una visión completa y detallada de los diferentes aspectos que abordamos en el estudio.

En esta sección, se establece la base teórica sobre la cual se construye la investigación. Se exploran los conceptos fundamentales del Bitcoin y su mercado, proporcionando una comprensión profunda de la criptomoneda y su contexto financiero. Adicionalmente, se examinan los antecedentes de estudios relacionados, lo que nos permite ubicar esta investigación dentro del marco más amplio de la literatura existente. También son sujeto de análisis las teorías que han surgido en relación con los factores que influyen en el precio del Bitcoin, sentando las bases para la exploración y análisis en las secciones posteriores.

La metodología empleada para abordar nuestra investigación se detalla en esta sección. Se describe el enfoque de regresión lineal múltiple que se va a utilizar para modelar las relaciones entre las variables dependientes e independientes. Con ello, se define y justifica la elección de las variables específicas que se pretenden estudiar e interpretar con relación al precio del Bitcoin. Por su parte, la recopilación y procesamiento de datos también se presentan aquí, haciendo hincapié en la importancia de la calidad y relevancia de los datos utilizados. La preparación y validación posterior de los datos para el modelo de regresión se detallan de forma exhaustiva, incluyendo aspectos como la homocedasticidad, valores atípicos o la división en conjuntos de entrenamiento y prueba, y la multicolinealidad.

Los resultados obtenidos mediante el modelo de regresión lineal múltiple se exponen en la sección de análisis de resultados. En ella, se presentan de manera concisa los hallazgos relacionados con la relación entre las variables independientes y el precio del Bitcoin, variable dependiente del modelo. La interpretación de los coeficientes de regresión se analiza en detalle, permitiendo una comprensión profunda de cómo cada variable puede afectar al precio y en qué medida de significancia estadística. Con ello, se realiza una prueba de los supuestos del modelo, asegurando la robustez y validez de los resultados. También se lleva a cabo la predicción del precio del Bitcoin, proporcionando una visión de las estimaciones futuras basadas en el modelo desarrollado.

La sección de discusión, pretender contextualizar los resultados obtenidos en la literatura existente. Aquí se comparan y contrastan los hallazgos con los estudios previos, permitiendo una comprensión más profunda de la contribución única de esta investigación. Añadido a esto, se analizan las implicaciones prácticas de los resultados, y se explora cómo nuestros hallazgos pueden ser relevantes para inversores y formuladores de políticas. Se identifican igualmente las limitaciones del estudio y se ofrecen recomendaciones para futuras investigaciones, teniendo por objetivo avanzar en la comprensión de los factores que influyen en el precio del Bitcoin y sus implicaciones con la economía y la sociedad.

En la sección de reflexión y comparación con otros métodos de predicción, nos adentramos en el contraste de los métodos utilizados en este estudio con otras aproximaciones de análisis y predicción posibles. Igualmente, exploramos los modelos de aprendizaje automático y su flexibilidad y posibilidades en la captura de patrones complejos, así como el riesgo potencial de sobreajuste en estas técnicas. Buscamos también conexiones con el análisis de series de tiempo como una herramienta para capturar tendencias temporales en los datos financieros. Por último, los modelos basados en sentimientos, factores externos y noticias son discutidos en detalle, dado su interesante potencial de análisis, razonaremos cómo los factores externos pueden influir en el precio del Bitcoin. Finalmente, se reflexiona sobre las conclusiones generales y se cierra la investigación con consideraciones finales que destacan la importancia de este estudio en el panorama de las criptomonedas.

En resumen, la estructura este trabajo está diseñado para llevar al lector a través de un viaje completo de investigación, análisis y reflexión en torno a los factores que influyen en el precio del Bitcoin. Cada sección se construye sobre la anterior para proporcionar una comprensión completa y matizada de este emocionante y dinámico campo financiero.

2. MARCO CONCEPTUAL

2.1. FUNDAMENTOS DEL BITCOIN Y SU MERCADO

El Bitcoin es hasta el momento la criptomoneda pionera de la actual era digital, posicionándose en el contexto económico, social y político, y despertando en el inversor una nueva concepción de interacción con el dinero y su valor, además de contribuir a desafiar algunas de las convicciones económicas de nuestra cultura. En esta nueva era de tecnología e información, las criptomonedas trascienden las fronteras y marcan un nuevo hito en la transformación de las finanzas y la sociedad del siglo XXI. En su breve historia, el precio del Bitcoin se ha percibido con una gran volatilidad. En 2011, alcanzó su punto máximo histórico hasta ese momento, de \$31,56, para posteriormente caer en 2014 y, de nuevo, resurgir en 2021 a un nuevo máximo histórico muy superior, de \$68.992. Esta volatilidad intuimos que proviene de una interacción compleja de variables, que incluyen cambios y ambigüedades regulatorias, liquidez limitada y especulación desenfrenada (Baldrés, 2018). Hay una ausencia de cooperación en la regulación del Bitcoin, que expone a su mercado al fraude y a la manipulación, mientras que el número restringido de compradores y vendedores generan una la liquidez limitada, que acentúa aún más las fluctuaciones de los precios. La especulación desmesurada fomenta un clima en el que los inversores compran Bitcoin con la esperanza de revenderlo a un precio más elevado en el futuro, incrementando aún más la volatilidad de los precios. En otros análisis (Auer et al, 2022) se concluye que los pequeños inversores tienden a especular con la compra de activos digitales de Bitcoin y otras criptomonedas cuando su valor empieza a subir, con la esperanza de que continúe el crecimiento en los sucesivos ejercicios; mientras que los grandes inversores conciben este momento para la venta de activos, siendo estos últimos los que obtienen los mejores rendimientos económicos.

En la actualidad, el mercado de criptomonedas revela un crecimiento en ascenso que repercute positivamente en la confianza de los inversores, pese a la amplia variedad de monedas virtuales disponibles. Desde su lanzamiento, el Bitcoin ha destacado y acaparado la atención de las mayorías, mostrándose, además, resistente a la elevada volatilidad. Ha sufrido considerables fluctuaciones y pérdidas, como la del 80% de su valor tras alcanzar su máximo histórico de 20,000 USD en 2017. Aun así, el Bitcoin sigue siendo la criptomoneda favorita y de mayor valor añadido para los inversores. El uso de tecnología Blockchain ha contribuido de forma notable a su aceptación, ofreciendo una mayor seguridad y confianza.

Debido al auge de las criptomonedas, ha nacido el interés por parte de la academia de estudiar y comprender su valor, comportamiento en el mercado y riesgo asociado, puesto que, además

de facilitar las operaciones cambiarias, se ha convertido en una fuente alternativa de financiación e inversión para particulares y empresas. Con la tendencia asociada hacia la digitalización de monedas y unidades de pago, la aceptación de las criptomonedas como medio de pago digital también aumentó (Müller y Ong, 2020).

Algunas referencias importantes nos sugieren que, si consideramos el empleo de las criptomonedas y su influencia sobre la economía mundial se puede observar que cada día estamos más cerca al uso frecuente de pagos a través de monedas digitales (Carrera et Al., 2020). Las criptomonedas son los activos financieros con mayor volatilidad que se pueden comercializar actualmente en los mercados financieros y sus precios son muy variantes respondiendo a las tendencias del mercado que afectan las decisiones de los inversores (García,2018). El Bitcoin es fuertemente volátil, pero dicha volatilidad podría ser menos pronunciada si los bancos dieran un paso al frente y aceptaran realizar intercambios de Bitcoin por otras monedas o activos, reduciendo así su inestabilidad en el precio (Báez, 2019).

Supone un reto de enorme interés el plantearnos una investigación profunda del Bitcoin y de las circunstancias y variables que afectan a su valor cada día. Se evidencian comportamientos y tendencias regulatorias que frecuentemente pasan desapercibidos o no son comprendidos en su totalidad por parte de los inversores. El hecho de no estar estrictamente reguladas o respaldadas por entidades gubernamentales y financieras, además de su carencia de validez como medios de pago oficiales, conlleva que estas monedas virtuales estén percibiendo la dificultad a la hora de garantizar su continuidad y estabilidad en los mercados financieros conocidos.

Considerando estos hechos, la propuesta de investigación sobre los factores que podrían influir y determinar el precio del Bitcoin, cobra una vital importancia.

A partir de variables como: la actividad minera de Bitcoin, la dificultad de minería, el recuento de cajeros automáticos de Bitcoin, la tasa de fondos federales, la oferta monetaria (M1) y la tasa de inflación global, encontramos la necesidad de determinar las relaciones subyacentes que desencadenan las fluctuaciones de precio de nuestra criptomoneda objeto de estudio. Como hemos referido anteriormente, cada una de estas variables está teóricamente contemplada para ejercer una determinada influencia sobre la valoración del Bitcoin a través de diversos mecanismos económicos y tecnológicos. Por todo ello, es esencial ampliar nuestra comprensión sobre cómo estas variables interactúan entre sí y con el mercado en conjunto.

2.2. ANTECEDENTES AL ESTUDIO

Los factores que influyen en el precio del Bitcoin suponen un fuerte punto de interés para la comunidad financiera y académica de la actualidad. Estudios precedentes han analizado esta cuestión desde distintas perspectivas, a través de enfoques estadísticos y análisis de datos para explorar la relación entre el precio del Bitcoin y diversos factores financieros, macroeconómicos y tecnológicos, principalmente. Estos antecedentes son de gran utilidad y nos proponen un marco de análisis para comprender la complejidad de los factores que intervienen en los cambios de tendencia y la volatilidad del Bitcoin en los mercados financieros.

El autor, Kristoufek (2013) nos ilustra con su análisis de regresión para evaluar cómo factores de carácter económico, financiero y de mercado, pueden influir en la variación del precio del Bitcoin. Mediante este estudio, se identificaron los elementos que tienen un impacto significativo en la dinámica de precios del Bitcoin. Otro estudio de relevancia es el de Bouoiyour y Selmi (2016), quienes llevaron a cabo un análisis de series temporales mediante el método ARDL para examinar las relaciones entre el precio del Bitcoin y diversas variables de tipo financiero y macroeconómico. Mediante este análisis, se hallaron patrones y tendencias de influencia para los cambios de valor y la volatilidad de la Criptomoneda objeto de estudio. Si hablamos, por su parte, de los factores regulatorios, Cheah y Fry (2015) se adentraron en un examen exhaustivo para medir la correlación entre el precio del Bitcoin y determinados aspectos regulatorios. Utilizaron un modelo de regresión lineal múltiple para comprender la relación existente entre la posición regulatoria de los gobiernos, entidades públicas y de las empresas, con el precio del Bitcoin. Sus resultados revelaron la existencia de una correlación estadísticamente significativa entre la postura regulatoria de los gobiernos y el precio del Bitcoin, un hecho que puede confirmar, a nuestro juicio, la percepción limitante que pueden tener algunos inversores a la hora de utilizar Bitcoin cuando se sienten o no, respaldados por las entidades oficiales de referencia en materia monetaria y financiera.

Bianchi (2020) investigó si las criptomonedas pueden considerarse una clase de activos, y qué factores macroeconómicos están generalmente relacionados con los precios de las criptomonedas, incluido el Bitcoin. Este estudio se sirvió de técnicas de análisis de series temporales y regresión para determinar la relación entre estas variables y el precio de la criptomoneda.

Por su parte, Dubey (2022) estudió los factores que tienen influencia en los retornos de Bitcoin. Su análisis abordó las carencias existentes de investigación en torno a los determinantes con poder predictivo sobre el precio de Bitcoin. En la investigación, se exploraron factores de

cuatro categorías: financieros, macroeconómicos, técnicos y fundamentales. El análisis utilizó datos de panel de cinco países para generar una regresión de efectos aleatorios, pruebas de causalidad de Granger y otras pruebas de diagnóstico para validar las hipótesis. Los hallazgos indicaron que, en el largo plazo, la oferta de Bitcoin, el precio del petróleo, el volumen de operaciones y la capitalización de mercado, obtuvieron un impacto significativo para el precio del Bitcoin. Mientras, en el corto plazo, los retornos de Bitcoin fueron relacionados solo con la capitalización de mercado y el precio del petróleo. Posteriormente, se encontró evidencia de que los retornos de Bitcoin aumentaron su influencia para los inversores, los rendimientos del S&P 500, la capitalización de mercado y el volumen de operaciones a corto plazo.

Sun, Y., Amanda, C. y Centana, B.C. (2023) investigaron los elementos que impactaron en el rendimiento del precio de Bitcoin antes y durante la pandemia del COVID-19. Se sirvieron de variables, como *hashrate*, volumen de transacciones, redes sociales y otras variables macroeconómicas, siendo capaces de aplicar el modelo de corrección de errores vectoriales (VECM) para analizar las posibles relaciones a corto y largo plazo entre estas variables. Los resultados de esta investigación destacaron que, tanto la red social, Twitter, como el precio del oro, afectaron al precio del Bitcoin en el corto plazo antes de la pandemia, mientras que *hashrate*, volumen de transacciones, Twitter y el índice de estrés financiero, impactaron, según sus evidencias, en el largo plazo. Esta investigación ha contribuido a una mayor comprensión de los factores influyentes del precio del Bitcoin, así como su evolución a través de diversos contextos de incertidumbre, como la pandemia mundial del Coronavirus.

Con todo ello, estos antecedentes nos brindan un panorama variado y diverso de los enfoques utilizados para comprender los factores que influyen de manera significativa en el precio del Bitcoin. Pese a que no todas las investigaciones se centran únicamente en la correlación entre variables, todas contribuyen a una comprensión más elevada de los factores determinantes que impactan en la fluctuación del Bitcoin en los mercados financieros.

2.3. FACTORES DE INFLUENCIA EN EL PRECIO DEL BITCOIN

Como para cualquier activo financiero conocido, el precio del Bitcoin puede ser influenciado por una serie de factores aleatorios y complejos. En los últimos años, los analistas y expertos en la materia han creado diversas teorías para tratar de explicar los determinantes del precio de nuestra criptomoneda objeto de estudio, el Bitcoin. Las teorías señalan, desde fundamentos económicos hasta factores sociales, tecnológicos y psicológicos y proporcionan una base para entender las dinámicas que subyacen en el mercado del Bitcoin. A continuación, exponemos algunas de las teorías más destacadas:

- **Teoría de la oferta y demanda:** Esta teoría, de base en la economía, es fundamental también en los mercados financieros y nos ayuda a entender que el precio del Bitcoin se rige por la interacción entre la oferta y la demanda. En aquellas situaciones donde la demanda excede la oferta, es lógico y común observar un incremento en el precio; de manera inversa, cuando la oferta supera la demanda, se tiende a registrar una disminución en el precio. Este mecanismo, aplicado al contexto de las criptomonedas, se fundamenta en la ley económica de la oferta y la demanda y, como es posible predecir, es extensible a casi todo tipo de bienes, servicios y activos. Por su parte, los eventos que afectan a la percepción de escasez, como la reducción a la mitad de la recompensa por bloque minado, pueden influir en el precio.
- **Teoría de la adopción:** Nos enseña que el incremento en la adopción y el uso de Bitcoin por cuenta de los individuos y empresas puede aumentar su precio. Cuantas más personas adopten Bitcoin como medio de intercambio económico o como una reserva de valor, la demanda aumentará, lo que implica generar un impacto al alza en el precio del activo financiero. Esta teoría está íntimamente relacionada con la anterior.
- **Teoría de la tecnología Blockchain:** La tecnología tras el Bitcoin es aquella conocida como Blockchain y, según estudios, juega un papel fundamental en su precio. El continuo avance y la innovación en esta tecnología de cadena de bloques, puede afectar la percepción del consumidor/inversor del valor a largo plazo de Bitcoin y, por ello, afecta a su precio en los mercados.
- **Teoría de factores macroeconómicos:** Los factores y movimientos macroeconómicos, como la inflación, las tasas de interés, y las políticas gubernamentales, pueden tener una influencia en el precio del Bitcoin. Un ejemplo se da en momentos de incertidumbre económica, cuando algunos inversores se deciden a recurrir a novedosos activos, como el Bitcoin, como un refugio seguro, lo que puede impulsar su precio al alza. Este era uno de los factores que mencionábamos al principio, cuando nos referimos al impacto de las crisis económico-financieras, que implican en muchos casos la búsqueda de nuevos activos de valor en la economía por parte de agentes e individuos.
- **Teoría de sentimiento del mercado:** El sentimiento del mercado juega un papel fundamental en las fluctuaciones del valor de los activos del mercado, como se ha observado siempre en los activos bursátiles que, hoy día, con mucha más fuerza, se ven influenciados por noticias, redes sociales y eventos actuales, pudiendo jugar un papel determinante en el precio del Bitcoin en el corto plazo. Las reacciones

emocionales de los inversores a eventos positivos o negativos pueden generar inesperados movimientos de precios. Este efecto, podemos intuir que es aún mayor cuando hablamos de activos económicos nuevos, que además son digitales, descentralizados y que no están respaldados por entidades financieras oficiales.

- **Teoría de la regulación:** Los procedimientos legales y regulatorios de los gobiernos y las instituciones financieras en torno al Bitcoin puede ser otro motivo de variación de su precio. Regulaciones más o menos favorables pueden aumentar la volatilidad en el mercado, afectando así en la confianza de los inversores finales.
- **Teoría de la competencia criptográfica:** Como sabemos, Bitcoin no es la única criptomoneda existente en el mercado, aunque sí la de mayor valor actualmente. La competencia entre distintas monedas digitales puede influir significativamente en su precio. Si en algún momento nuevas criptomonedas ganan popularidad y adopción por parte de inversores, esto podría afectar a la demanda de Bitcoin en el futuro.

3. METODOLOGÍA

3.1. REGRESIÓN LINEAL MÚLTIPLE

En este proyecto, vamos a abordar nuestro análisis desde la perspectiva de la regresión lineal múltiple (MLR), también denominada regresión múltiple. Podemos recordar que estamos ante un enfoque estadístico empleado con el objetivo de anticipar el resultado de una variable, basándose en el valor de dos o más variables. La variable que se pretende predecir es designada como la variable dependiente, mientras que las variables que son utilizadas para estimar su valor se conocen como variables explicativas o independientes. Para una variable dependiente, su valor se modifica en función de otra variable, denominada variable independiente. Desde una perspectiva científica, la variable dependiente es aquella que es sometida a prueba. Por su parte, las variables dependientes pueden llamarse variables de resultado, variables de respuesta o variables en el lado izquierdo. La regresión lineal múltiple, se contextualiza en un método estadístico que emplea dos o más variables independientes para pronosticar la variable dependiente. Los economistas y analistas utilizan esta técnica para determinar la variabilidad del modelo y la contribución relativa singular de cada variable independiente en la variabilidad total. Las dos formas principales que puede adoptar la regresión múltiple son la regresión lineal y la regresión no lineal.

En el año 2006, Montgomery et al. reportó que la regresión lineal múltiple es el modelo donde interviene más de una variable independiente o regresora. El modelo se describe acorde a la siguiente expresión:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon$$

Donde, para $i = n$ observaciones:

y_i variable dependiente

x_i variable explicativa

β_0 intercepto (término constante)

β_p coeficientes de pendiente para cada variable explicativa del modelo

ε término de error del modelo (también llamado residuo)

El principio fundamental del modelo de regresión múltiple se sostiene en una serie de supuestos clave. El punto de partida descansa bajo la premisa de que existe una relación lineal entre las variables dependientes y las variables independientes. Junto a este enunciado, se establece que las variables independientes no deben presentar una excesiva correlación entre sí. Por otra parte, asumimos que la muestra de observaciones se elige de forma independiente y al azar de la población disponible en cuestión. Otro factor necesario es que los residuos, lo que es lo mismo, las diferencias entre los valores observados y los valores predichos sigan una distribución normal con media igual a cero y una varianza σ para asegurar la validez del análisis.

En cuanto a las métricas de análisis, el coeficiente de determinación (R-cuadrado, o R^2) es una métrica estadística utilizada para cuantificar la proporción de la variación en el resultado que se puede explicar mediante la variación de las variables independientes. Cuantos más predictores agreguemos al modelo de regresión lineal múltiple (MLR), el valor de R-cuadrado tiende a aumentar y mejorar su capacidad para explicar el modelo. Sin embargo, este valor, por sí solo, no permite identificar qué predictores deben incluirse o excluirse del modelo. Su rango oscila en una amplitud de entre 0 y 1, donde 0 indica que ninguna de las variables independientes elegidas puede prever el resultado; mientras que 1 indica que el resultado puede predecirse en su totalidad a partir de las variables independientes. Para interpretar los resultados del análisis, entendemos que los coeficientes beta son significativos en la medida en que todas las otras variables se mantengan constantes. Podemos representar el resultado de la regresión múltiple, tanto horizontalmente, como una ecuación, o en forma tabulada verticalmente, ofreciendo una comprensión visual detallada de las relaciones entre las variables.

Es de vital trascendencia que, en los modelos de regresión lineal múltiple, los predictores exhiban independencia entre sí y eviten la presencia de colinealidad. La colinealidad se da cuando un predictor revela una relación lineal con uno o varios de los demás predictores del modelo. También cuando uno de los predictores se convierte en una combinación lineal de los demás. Esta situación nos conduce a una falta de precisión a la hora de identificar los efectos individuales de las variables colineales sobre la variable respuesta. Por consiguiente, se amplía la varianza de los coeficientes de regresión estimados, lo que dificulta la determinación de su significación estadística. Además, hemos de tener en cuenta que pequeñas variaciones en los datos pueden generar cambios significativos en las estimaciones de los coeficientes. Cierto es que la colinealidad total solo se evidencia cuando el coeficiente de correlación simple o múltiple entre algunas de las variables independientes es igual a 1. Lo cierto es que esta situación no se corresponde con la realidad. Aun así, es común encontrarse con el fenómeno de casi-colinealidad o multicolinealidad no perfecta. Al no existir un enfoque estadístico definitivo para determinar la multicolinealidad o colinealidad entre los predictores de un modelo de regresión, sí encontramos evidencia de diversas reglas empíricas formuladas para evaluar su impacto en la estimación y el contraste del modelo. Entre los pasos lógicos establecidos para abordar esta problemática, se encuentran: observar si el coeficiente de determinación R^2 es alto (cercano a 1) sin que otros predictores resulten significativos, lo que podría sugerir colinealidad; generar una matriz de correlación para medir las relaciones lineales entre los pares de predictores, habida cuenta de que, incluso sin evidencia de altas correlaciones, no se descarta la presencia de multicolinealidad; calcular modelos de regresión lineal simple para cada predictor en función del resto. Si el coeficiente de determinación R^2 es elevado en alguno de ellos, indicaríamos posible colinealidad; y finalmente, utilizar métricas adicionales como la tolerancia (TOL) y el Factor de Inflación de la Varianza (VIF), que cuantifican la relación entre la variabilidad de un predictor con el resto de los predictores, evidenciando de esta manera posibles indicios de colinealidad.

$$VIF_{\beta_j} = \frac{1}{1 - R^2}$$

El cálculo de R^2 se obtiene a través de la regresión del predictor X_j en relación con los demás predictores. Este enfoque es el más oportuno, ya que se emplean umbrales de referencia usuales para la métrica VIF (Factor de Inflación de la Varianza) de la siguiente manera:

- Si $VIF = 1$: Indica la ausencia total de colinealidad.
- Si $1 < VIF < 5$: Puede sugerirnos cierta colinealidad que podría afectar la regresión del modelo.

- Si $5 < VIF < 10$: Sugiere un motivo de preocupación por una posible colinealidad.
- La tolerancia, expresada como el inverso del VIF ($1/VIF$), tiene umbrales recomendables entre (0,1 y 1).

$$Tolerancia_{\beta_j} = \frac{1}{VIF_{\beta_j}}$$

Cuando advertimos la presencia de colinealidad entre los predictores del modelo, nos servimos de dos enfoques para afrontar la situación. El primer enfoque consiste en excluir uno de los predictores problemáticos, manteniendo el que, según el juicio del investigador, ejerce una influencia real sobre la variable respuesta (Hernández Martín, 2012). Esta medida no suele afectar significativamente a la capacidad predictiva del modelo, ya que la colinealidad nos indica que la información que aporta uno de los predictores es redundante en presencia del otro. La segunda forma de superar la colinealidad busca consolidar las variables colineales en un único predictor, aunque esto conlleva el riesgo de omitir cierta interpretación individual. Si buscamos establecer relaciones causa-efecto, la existencia de colinealidad puede llevarnos a interpretaciones erróneas, haciéndonos creer que una variable es la causa cuando, realmente es la otra la que está ejerciendo influencia sobre ese predictor. Por todo ello, es imprescindible abordar la colinealidad de manera profunda y científica para poder garantizar resultados sólidos y confiables en el análisis de regresión múltiple.

3.2. DEFINICIÓN DE VARIABLES: DEPENDIENTES E INDEPENDIENTES

En esta sección, establecemos la distinción entre las variables dependientes e independientes elegidas para ser sometidas a análisis en el marco de nuestra investigación. La variable dependiente, en este contexto, es el precio del Bitcoin en el mercado financiero, que se posiciona como el factor central de interés y observación. El precio del Bitcoin representa el valor cuyo comportamiento buscaremos explicar y prever a través del modelo de regresión lineal múltiple.

Precio Bitcoin (US): Cuando nos referimos al "Precio Bitcoin (US)" entendemos que expresa el valor en dólares estadounidenses (US) de una unidad de la criptomoneda Bitcoin en un momento específico del tiempo. El precio del Bitcoin es nuestra variable dependiente y presenta un coeficiente de volatilidad elevado debido a factores que rigen la oferta y la demanda en los mercados de intercambio, a eventos socioeconómicos y políticos, a la adopción pública y privada, y otros acontecimientos variados relacionados con el contexto de las criptomonedas. El precio del Bitcoin es un indicador crucial en el análisis y seguimiento de

las tendencias del mercado, así como para la toma de decisiones de inversión y estrategias de negocio.

Por otra parte, presentamos las variables independientes, también conocidas como explicativas, que buscan reunir el conjunto de factores que influyen potencialmente en el precio del Bitcoin. A continuación, ofrecemos una descripción de las variables dependientes a considerar en nuestro estudio.

Dificultad de minería: Desde el mes de mayo de 2021, la extracción de nuevos Bitcoin (BTC) en la cadena de bloques ha experimentado un aumento en su dificultad, alcanzando un máximo histórico de cerca de 22 billones. Esta cifra refleja ante todo la cantidad de poder de cómputo necesario que se aplica para extraer la criptomoneda, y se calcula cada 14 días en función del “poder de hash” que entra en competencia por las recompensas en la red durante este período bisemanal. *Hashing* se refiere a la cantidad de potencia de procesamiento que utilizan los PC para construir la cadena de bloques. Cuantos más bloques de transacciones verificadas se procesan o "hash", más Bitcoin o recompensa por este proceso de verificación, se extrae. Si aumenta la dificultad de minería implica que un mayor número de mineros digitales compite para obtener las criptomonedas (by Raphael Auer, 2023). En su comienzo, se estimó que el suministro máximo de BTC se alcanzaría en el año 2024, aunque ya casi el 90 por ciento de este suministro se alcanzó en marzo de 2021, lo que influye en la dificultad de la minería.

Rentabilidad de minería: El proceso de minería de Bitcoin generó cada vez más beneficios a finales de 2020, pero el crecimiento de las ganancias pareció detenerse durante el mes de marzo de 2021. Cuando se ejecuta el proceso de minería de criptomonedas, una computadora está tratando de resolver problemas lógicos complejos para verificar las transacciones en la cadena de bloques. Una vez completado este proceso, el minero recibe criptomonedas como recompensa en bloque con un valor comparable a su esfuerzo de minado. Parece que es probable que las computadoras con mayor tasa de hash o poder de cómputo, resuelvan más acertijos o problemas lógicos y, por lo tanto, extraigan más criptomonedas. Para que un minero digital pueda ganar dinero con este proceso, se han de considerar varios costes, como el consumo de electricidad durante el proceso de extracción, las tarifas de transacción, o si el hardware utilizado es eficiente o no, entre otras.

Recuento de cajeros automáticos BTC EE. UU: A 1 de noviembre de 2020, había unos 9.096 cajeros automáticos de Bitcoin en el mundo. Podemos encontrar dos tipos principales de cajeros automáticos de Bitcoin: los básicos, que tan solo permiten a los usuarios comprar Bitcoins, y los complejos, que permiten a los usuarios comprar y vender dinero virtual. El mayor número de cajeros automáticos de Bitcoin se registró en los Estados Unidos.

Tasa de fondos federales: La tasa efectiva de los fondos federales de EE. UU. se redujo drásticamente entre febrero y abril de 2020 (IMF, 2023). Cayó del 1,58 por ciento en febrero de ese año, al 0,65 en marzo y al 0,05 en abril. La reducción de esta tasa se entiende como una consecuencia a la pandemia mundial del COVID-19, como respuesta a la flexibilización cuantitativa de la Reserva Federal durante el mismo período. Tras algunos cambios leves en la tasa efectiva desde ese momento, se fijó en 0,08 por ciento en septiembre de 2021, y se mantuvo invariable hasta marzo de 2022, cuando se incrementó a 0,2. Desde mayo del 2022, la tasa efectiva de fondos federales de EE. UU. alcanzó el 0,77 por ciento.

La tasa efectiva de los fondos federales de EE. UU. determina la tasa de interés que pagan las instituciones de depósito, como bancos y cooperativas de crédito, que prestan saldos de reserva a otras instituciones de depósito (IMF, 2023). Cuando se cambia la tasa efectiva en tiempos de crisis, se busca una forma común de estimular la economía, ya que tiene un impacto significativo en toda la realidad económica, como el crecimiento económico, la inflación o el empleo.

Fueron muchos los países, a parte de Estados Unidos, los que ajustaron su tasa efectiva, u otras tasas de interés de depósito, como respuesta a los efectos económicos desencadenados por la pandemia del Coronavirus. En todo el mundo, los gobiernos y los bancos centrales tomaron medidas para minimizar la crisis económica. La gran mayoría de los países redujeron las tasas de sus bancos centrales a principios de 2020, y no fue hasta un año después que la mayoría de los países comenzaron a aumentarlas lentamente de nuevo.

M1 Oferta monetaria: Previamente a mayo de 2020, momento en el que se produce un punto de inflexión en la economía, el concepto de M1 englobaba tres elementos primordiales. En primera instancia, la moneda en circulación fuera del Tesoro de EE. UU, las entidades bancarias de la Reserva Federal y las cajas fuertes de las instituciones de depósito. En segundo lugar, los depósitos a la vista presentes en bancos comerciales, excluyendo los montantes bajo la custodia de instituciones de depósito, el gobierno estadounidense, los bancos extranjeros y las instituciones oficiales, y ajustándolos por las sumas en efectivo en trámite de recolección y la reserva federal flotante (Banco Mundial, 2023). En tercer lugar, abarcaba otros depósitos a la vista (OCD), que pueden referirse, tanto las cuentas de orden de retiro negociable (NOW) como las de servicio de transferencias automáticas (ATS) en instituciones de depósito, junto con depósitos a la vista en instituciones de ahorro y cuentas de participación en cooperativas de crédito (Banco Mundial, 2023).

A partir de mayo de 2020, la definición de M1 consta de los siguientes componentes. Primero, incorpora la moneda fuera del Tesoro de los Estados Unidos, los bancos de la Reserva Federal

y las bóvedas de las instituciones de depósito. En segundo lugar, cuenta con los depósitos a la vista presentes en bancos comerciales, excluyendo los montos que están bajo la custodia de instituciones de depósito, el gobierno estadounidense, bancos extranjeros e instituciones oficiales, siendo ajustados por las sumas en efectivo en trámite de recolección y la reserva federal flotante. En tercer término, comprende diversos depósitos líquidos, que engloban tanto los OCD como los depósitos de ahorro, e incluyen las cuentas de depósito del mercado monetario (Banco Mundial, 2023). El M1 desestacionalizado es construido a partir de la suma de la moneda, los depósitos a la vista y los OCD (previo a mayo de 2020) u otros depósitos líquidos (a partir de mayo de 2020), necesariamente desestacionalizando cada uno de estos componentes por separado (IMF,2023).

Tasa de inflación mundial: Este dato estadístico muestra la tasa de inflación comparada con el año anterior. En el año 2020, la tasa de inflación a nivel mundial ascendió a aproximadamente 3,18 por ciento en comparación con el período anterior.

Según el Banco Mundial, la inflación experimentó un aumento considerable, especialmente entre los años 2021 a 2022. *“La razón principal fue el precio de la energía, que no sólo afecta directamente a los consumidores a través de los costes de electricidad y calefacción, sino que también encarece la producción de bienes. En 2022, la guerra de Ucrania fue un factor adicional, que provocó un aumento repentino de los costes de calefacción como consecuencia de los cortes de suministro de gas”.* (Banco Mundial, 2023)

3.3. RECOPIACIÓN Y PROCESAMIENTO DE DATOS

La fase de búsqueda, recopilación y procesamiento de datos es considerada como una parte crítica y fundamental en el proceso analítico llevado a cabo en esta investigación. Los datos de precios del Bitcoin, un elemento central en este estudio, se obtuvieron de (CoinGecko), una plataforma digital de alto reconocimiento que ofrece información exhaustiva sobre el rendimiento de una amplia gama de criptomonedas. Este sitio web, junto con su aplicación móvil conforman una fuente confiable y actualizada de datos para comprender las variaciones de precios y las tendencias que adopta cada día el mercado de las criptomonedas. Añadido a lo anterior, para poder integrar las variables de dificultad y rentabilidad de la minería del Bitcoin en el modelo, se recurrió al explorador de BTC.com. Esta completa herramienta proporciona estadísticas detalladas que relacionan los datos de bloques, transacciones y direcciones, lo cual nutre el análisis con información crítica sobre la cadena de bloques y su funcionamiento. Con el objetivo de enriquecer aún más la investigación, encontramos información precisa sobre el recuento de cajeros automáticos de Bitcoin a través de (Coin ATM Radar). Esta fuente es capaz de rastrear y localizar las ubicaciones de los cajeros de criptomoneda,

proporcionando datos esenciales para entender la infraestructura física del Bitcoin y su alcance global. Junto con ello, fueron recopilados también datos relevantes de fuentes de gran reputación, como St. Louis Fed, la Reserva Federal de St. Louis. Esta institución es fundamental para la economía de Estados Unidos y aporta datos relevantes sobre las tasas de fondos federales y la oferta monetaria M1. Por su parte, la tasa de inflación global fue obtenida del Fondo Monetario Internacional (FMI), entidad de renombre internacional que nos brinda información básica para ayudarnos a comprender el contexto económico global. Con esta meticulosa recopilación de estos datos y su consecutivo procesamiento garantizamos la rigurosidad y la base sólida sobre la cual se cimentará el análisis de la relación entre las variables seleccionadas y el precio del Bitcoin.

3.4. VALIDACIÓN DE DATOS PARA EL MODELO DE REGRESIÓN

Validar los supuestos de regresión supone una etapa crucial de la metodología en este estudio, dada su intención de garantizar la solidez y confiabilidad de los resultados obtenidos mediante el análisis de regresión lineal múltiple. En esta fase llevaremos a cabo una evaluación exhaustiva de los supuestos fundamentales que dan sustento a esta técnica estadística. La teoría nos ilustra y nos indica que la regresión lineal múltiple (RLM) se basa en determinadas premisas, tales como la relación lineal entre las variables, la independencia y aleatoriedad de las observaciones, además de la normalidad en la distribución de los residuos y la ausencia de multicolinealidad entre las variables independientes. El hecho de validar estos supuestos es esencial para asegurar que los resultados sean adecuadamente interpretados y utilizados, y que las conclusiones extraídas puedan ser respaldadas por una base fiable y muy sólida.

A continuación, en esta fase, aplicaremos diversos métodos y pruebas estadísticas para concluir si los supuestos se cumplen para el conjunto de datos analizados. Procederemos a examinar si existe una relación lineal adecuada entre las variables, y evitaremos así aplicar la regresión en casos donde no se cumpla este hecho. También verificaremos la independencia y aleatoriedad de las observaciones, con el objetivo de asegurar que los resultados no presentan sesgos. Por su parte, la normalidad de los residuos será analizada a través de pruebas estadísticas que confirmen si los residuos siguen una distribución normal con una media de cero y una varianza constante. Por último, interpretaremos si existe o no multicolinealidad entre las variables independientes, con el objetivo de evitar la presencia de relaciones lineales fuertes entre las variables que puedan distorsionar los resultados finales. Con la validación de estos supuestos garantizamos la integridad, además de la robustez del análisis de regresión múltiple, ofreciendo una interpretación lo más precisa y confiable posible sobre las relaciones entre las variables independientes y el precio del Bitcoin.

3.4.1. Homocedasticidad

Uno de los supuestos que han de cumplirse en nuestro análisis es la homocedasticidad. Para asegurar su existencia, tenemos que dar un paso fundamental en el proceso de construcción del modelo, que es aplicar una transformación logarítmica al precio del Bitcoin. La homocedasticidad es la propiedad bajo la cual la dispersión de los residuos, es decir, las diferencias entre los valores observados y los predichos, se mantiene constante a lo largo del rango de valores pronosticados del modelo. Este hecho ha de producirse para garantizar la validez y la confiabilidad de los resultados obtenidos a través del análisis de regresión. Como sabemos, la transformación logarítmica es una técnica muy utilizada en modelos de regresión y análisis estadísticos para moderar la varianza de los datos y cumplir con el supuesto de homocedasticidad.

En nuestro contexto de análisis del precio del Bitcoin, el hecho de aplicar la transformación logarítmica tiene fundamentos sólidos. Considerando la naturaleza volátil y tantas veces exponencial de los cambios de precio en el mercado de las criptomonedas, es bastante lógico y común que los datos nos revelen una distribución asimétrica (Figura 1). En cambio, cuando aplicamos el logaritmo, los valores extremos tienden a atenuarse notablemente, lo que simplifica la amplitud de los cambios más drásticos y reduce la dispersión de los datos en un rango más manejable para continuar con la investigación.

Hemos de destacar que, aunque vayamos a conservar el precio original como referencia en nuestro análisis, no lo utilizaremos como variable independiente en el modelo de regresión lineal. En su lugar, emplearemos la columna de registros transformados por logaritmo del precio. Esta columna representará la nueva variable independiente en nuestro análisis. Hemos tomado esta elección dada la capacidad del logaritmo para reducir las diferencias extremas en los valores del precio de la criptomoneda, lo cual mejora su interpretación, además de la linealidad en el modelo de regresión, contribuyendo cumplir con mayor rigurosidad los supuestos requeridos.

3.4.2. Valores atípicos

Es fundamental que identifiquemos y gestionemos las observaciones atípicas resultantes para el modelo de análisis de regresiones lineales múltiples. Estas observaciones distan de la tendencia general del conjunto de datos, y tienen potencial para ejercer una influencia dañina en el modelo, pudiendo distorsionar los resultados. Podemos emplear varias diferentes técnicas para detectarlas. Una de las más efectivas es a través del análisis de los residuos. Para que podamos asegurar la robustez y validez de los resultados, es necesaria la eliminación de valores atípicos. En este sentido, podemos servirnos de la técnica de distancia de Cook, la cual evalúa tanto el residual como el apalancamiento de cada observación, lo que

nos proporciona una medida de la influencia que tiene cada punto de datos en el modelo. Cuando calculamos la distancia de Cook, cuantificamos cuánto y de qué forma se altera el modelo de regresión, si una determinada observación es eliminada. Hemos de considerar principalmente aquellas observaciones con distancias de Cook que superen un umbral, normalmente establecido como tres veces la media de todas las distancias.

En nuestra investigación, hemos identificado tres valores atípicos en la columna de valores atípicos agregados (Anexo.Tabla 1). Estos puntos destacados tienen el potencial para alterar la fuerza y la dirección de las relaciones halladas del modelo de regresión múltiple. Por esta razón, con el objetivo de ser rigurosos y garantizar la integridad y validez de los resultados, procedemos con la eliminación de los valores atípicos identificados en el conjunto de datos. Con esta acción, buscamos minimizar la influencia negativa que podrían desencadenar estos valores sobre la calidad de las conclusiones extraídas de nuestro análisis. La correcta gestión de observaciones atípicas refuerza la confianza en la interpretación de los resultados del modelo de regresión múltiple.

3.4.3. Test y Train

Hemos procedido con la creación de conjuntos de entrenamiento (train) y de validación (test) como un paso de valor para asegurar la calidad y la generalización del modelo estudiado. El conjunto de entrenamiento es utilizado para el desarrollo y ajuste del modelo, mientras que, por su parte, el conjunto de validación lo utilizaremos posteriormente para evaluar la capacidad predictiva del modelo analizado. Con ello, procederemos a generar los conjuntos de entrenamiento y validación y una vez construido el modelo, utilizando los datos del conjunto de entrenamiento, lo aplicaremos al conjunto de validación para medir su eficacia y desempeño en datos no vistos con anterioridad.

Tras completar el proceso de construcción del modelo utilizando los datos de entrenamiento, volveremos a evaluar el conjunto de validación para comprender cómo el modelo se comporta ante datos distintos. Con el objetivo de garantizar una distribución adecuada de los datos, evitando todo tipo de sesgos, asignaremos el 70% de los datos al conjunto de entrenamiento, mientras que el 30% restante, irá al conjunto de validación. De esta forma pretendemos asegurar la representación significativa de la información en ambos conjuntos. Con esta división estratégica, evaluaremos de forma precisa la capacidad del modelo para generalizar nuevas observaciones no utilizadas en la fase de desarrollo.

3.4.4. Multicolinealidad

La multicolinealidad representa uno de los supuestos más relevantes en el análisis de regresión. Determina la presencia de una posible correlación significativa entre múltiples variables independientes. Cuando ocurre este fenómeno, se pueden generar problemas con

el análisis, que a priori cuenta con que las variables predictoras sean independientes entre sí con el fin de evitar problemas de precisión e interpretación en los coeficientes estimados en la investigación.

Contextualizado a nuestro estudio, hemos encontrado la forma de observar la correlación de las principales variables independientes: BTC_Dificultad_Mineria, BTC_ATM_US, M1_Oferta_Monetaria, y Tasa_Inflacion_Global, y hemos concluido, como se puede apreciar a continuación, que se puede percibir una correlación que va desde moderada a alta con otras variables del conjunto (Figura 1). Este hallazgo inicial, aunque no concluyente, genera preocupación, ya que, como hemos enunciado anteriormente, la correlación alta entre variables independientes puede dificultar la interpretación individual de su efecto en la variable dependiente.



Figura 1. Matriz de correlación variables independientes. Elaboración propia.

Con el fin de abordar este reto y tomar las decisiones óptimas sobre la inclusión o no de ciertas variables, vamos a emplear la métrica de Factor de Inflación de Varianza (VIF, por sus siglas en inglés). El VIF nos permite cuantificar la intensidad de la colinealidad y medir la relación entre la varianza de un coeficiente estimado y la varianza cuando la variable predictora es eliminada del modelo. En general, como hemos descrito al principio del capítulo, es necesario

que el VIF sea menor a 10, ya que valores superiores a este, indican una multicolinealidad significativa.

Variable	VIF
BTC_Dificultad_Mineria	5.160656
BTC_Rentabilidad	2.311268
BTC_ATM_US	4.805880
Fed_Tasa_US	4.974582
M1_Oferta_Monetaria	8.509970
Tasa_Inflacion_Global	3.940872

Tabla 1. Resultados del Factor de Inflación de la Varianza (VIF). Elaboración propia.

El análisis realizado muestra los resultados donde se calcularon los valores de VIF para cada variable independiente. Los resultados indican que todas las variables están por debajo del umbral de 10, lo que sugiere que no existe una multicolinealidad severa entre ellas. Pese a que habían sido observadas correlaciones en un nivel superficial, el análisis del VIF nos proporciona cierta tranquilidad en lo que se refiere a la multicolinealidad, y nos permite conservar todas las variables expuestas en el análisis por el momento. Es determinante recordar que la mitigación de la multicolinealidad es necesaria para asegurar la interpretación adecuada y precisa de los resultados de regresión, contribuyendo a garantizar la robustez del modelo ante nuevos datos.

4. ANÁLISIS DE RESULTADOS

El precio del Bitcoin (BTC) experimentó un nuevo hito histórico en 2021, superando valores que alcanzaron los 65.000 USD en los meses de febrero, abril y noviembre de dicho año (Figura 2). Las dos primeras subidas se debieron a eventos que involucraron a Tesla y Coinbase, respectivamente; mientras que el último, fue de nuevo debido al anuncio de Tesla en el que compartía que había adquirido la moneda digital por valor de 1.500 millones de dólares estadounidenses. Con ello, contemplamos la salida a bolsa del mayor intercambio de cifrado de EE. UU., hecho que contribuyó a alimentar nuevamente el interés masivo sobre el Bitcoin.

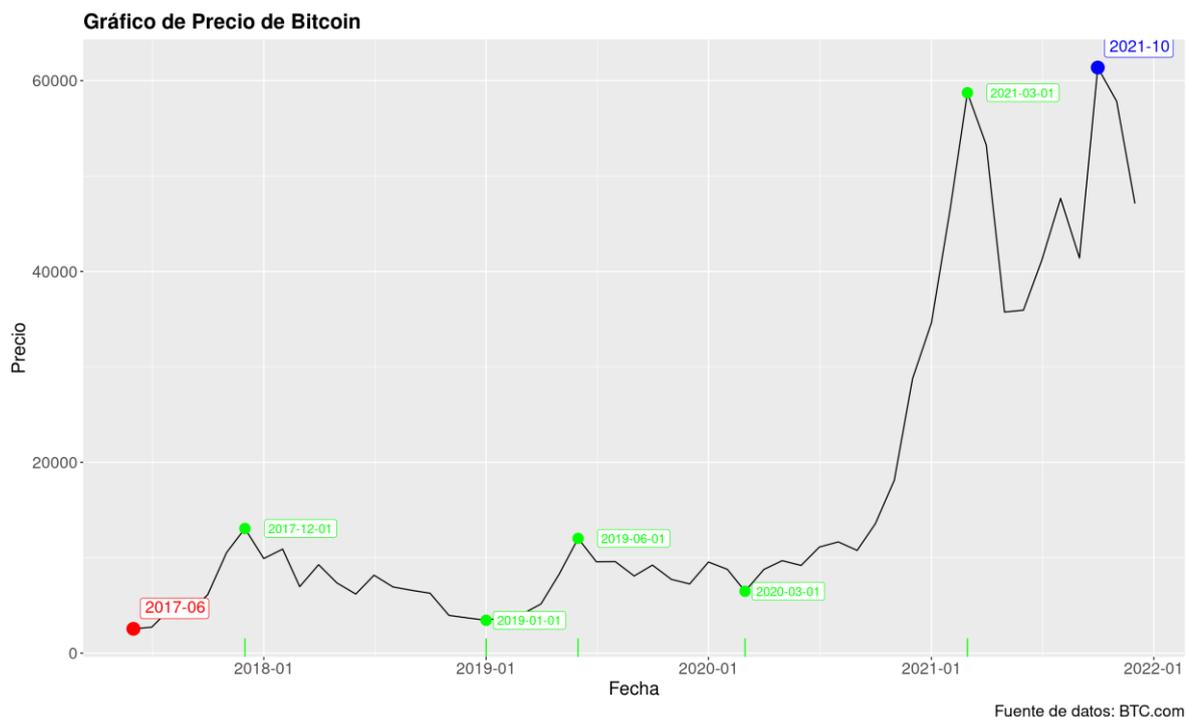


Figura 2. Serie de tiempo sobre la evolución del precio del Bitcoin (USD). Fuente: BTC.com

Estamos ante la criptomoneda más conocida y mediática del mundo que, no obstante, sufrió una importante corrección en abril tras las especulaciones que se sucedieron en aquellas semanas sobre la regulación gubernamental. Estos hechos se relacionan directamente con la percepción del valor de la criptomoneda por parte de los usuarios e inversores. Otro motivo de su corrección, según los expertos, fue un apagón en la región de Xinjiang en China, lo que condujo a una disminución en el hashrate de Bitcoin (cuántos Bitcoins se extraen) y, lo que generó un cambio de actitud por parte de bastantes inversores que los llevó a vender sus activos. Según los datos basados en las direcciones IP de los llamados “hashers”, o extractores de Bitcoin, quienes utilizaban ciertos grupos de minería de Bitcoin, se estima que más de la mitad de toda la minería de Bitcoin se produjo en China. La tercera subida de precios parece que estuvo vinculada con el lanzamiento de un ETF de Bitcoin en los Estados Unidos.

4.1. PRESENTACIÓN DE RESULTADOS DE LA REGRESIÓN

Hemos procedido con el cálculo del modelo de regresión lineal con el objetivo de analizar las relaciones entre las variables seleccionadas en nuestro análisis, aplicando los principios de la regresión lineal al conjunto de datos. La variable dependiente fue examinada en relación con el resto de las variables independientes. Cuando hemos procedido con el cálculo del modelo de regresión lineal, hemos recogido algunas ideas sobre cómo los cambios generados en las variables independientes influyen en la variación de la variable dependiente. El análisis trata de revelar la fuerza, dirección y significancia de estas relaciones, abriendo un marco

cuantitativo para analizar y entender los factores que contribuyen a las fluctuaciones en la variable objetivo. Mediante este proceso de modelado, nuestro estudio busca ofrecer ideas valiosas sobre la dinámica de los fenómenos observados, contribuyendo a la mejor comprensión de los elementos que nos llevan a los resultados de la investigación.

Después de proceder con un análisis exhaustivo de las variables predictoras en nuestro modelo, hemos podido identificar que algunas de estas variables seleccionadas no presentan una relación estadísticamente significativa con la variable dependiente. En este sentido, hemos decidido proceder con la eliminación gradual de aquellas variables que obtuvieron p-valores mayores a 0,05, lo cual implica que no aportan una contribución significativa para el modelo objeto de análisis.

El orden de nuestra investigación nos ha llevado a, primero, observar que *BTC_ATM_US*, la variable con el p-valor más alto, carecía de relevancia en la predicción del resultado. Por ello, se procedió con su eliminación y se volvió a ejecutar el proceso. Sin embargo, en el segundo análisis, se encontró que la siguiente variable con menor significancia era *Fed_Tasa_US*, lo que nos llevó, nuevamente, a su eliminación en el siguiente paso.

De nuevo, repetimos el proceso y encontramos que *M1_Oferta_Monetaria*, tampoco estaba contribuyendo de manera significativa a explicar el fenómeno en estudio, debido a su elevado p-valor. Por lo tanto, se decidió también descartar esta variable y realizar una nueva iteración. Finalmente, se llegó también a una decisión ejecutiva con respecto a la variable *Tasa_Inflacion_Global*. Esta variable, pese a tener un p-valor que podría intuirse como atractiva para el modelo, se reconoció que no era confiable, a causa, principalmente, de la limitada granularidad de los datos. La tasa de inflación global, que se había obtenido a partir de un promedio anual, no estaba capturando adecuadamente las fluctuaciones mensuales que podrían influir en el fenómeno analizado y que son tan relevantes para la variable dependiente. Por este motivo, se decidió eliminar esta variable del conjunto de variables predictoras del modelo.

Tras estas comprobaciones, la iterativa evaluación y selección de variables llevada a cabo, concluyó con la determinación de que las variables *BTC_Dificultad_Mineria* y *BTC_Rentabilidad* son las únicas variables predictoras que se consideraron significativas y confiables para explicar el modelo. Esta decisión se tomó con el objetivo de construir un modelo sólido que pudiera generar *insights* significativos y relevantes sobre los factores que influyen en el precio del Bitcoin, fenómeno bajo investigación. El modelo resultante quedaría como:

$$Y_0 = \beta_0 + \beta_1 x_i + \beta_2 x_i + \varepsilon_i$$

4.2. INTERPRETACIÓN DE LOS COEFICIENTES DE REGRESIÓN

Tras un minucioso proceso de selección y evaluación de variables, acabamos determinando cuáles son las variables independientes que presentan una significancia estadística en nuestro modelo de regresión lineal múltiple. Las variables que conservaremos para proceder con nuestro análisis son: BTC_Dificultad_Mineria y BTC_Rentabilidad.

En términos de capacidad explicativa y ajuste del modelo, el coeficiente de determinación R² es de 0,7129. Esto significa que el modelo que hemos creado es capaz de explicar el 71,29% de la variabilidad presente en el conjunto de datos de prueba generados.

El modelo se traduce a la siguiente ecuación:

$$\text{PrecioLog} = 7.69 + 0.12X_{\text{BTC}_{\text{DificultadMineria}}} + 0.78X_{\text{BTC}_{\text{Rentabilidad}}} + \varepsilon$$

Donde, la constante en el origen es 7,69 y representa el valor de la variable dependiente, Precio_Log cuando todas las demás variables independientes son iguales a cero.

Observamos que un incremento de 0,12 en la variable BTC_Dificultad_Mineria implica un aumento de 1 unidad en la variable Precio_Log, siempre que mantengamos constante el efecto de las otras variables. Del mismo modo, un aumento de 0,78 en la variable BTC_Rentabilidad se corresponde con un aumento de 1 unidad en la variable dependiente Precio_Log, manteniendo siempre constantes el resto de las variables predictoras. En conclusión, estos coeficientes explican cómo un mínimo cambio en una de las variables independientes puede impactar en el valor estimado de la variable dependiente, mientras que las otras variables permanecen constantes.

Los resultados anteriores nos proporcionan una comprensión más detallada y cuantitativa de cómo las variables seleccionadas están influyendo en el precio logarítmico del Bitcoin. Los coeficientes que hemos obtenido ofrecen una medida concreta de la relación existente entre las variables independientes y la variable dependiente en el contexto del modelo de regresión lineal múltiple desarrollado.

4.3. VERIFICACIÓN DE LOS SUPUESTOS DEL MODELO

La verificación de los supuestos del modelo de regresión supone un paso determinante en el análisis estadístico, debido a que nos garantiza la validez y la confiabilidad de los resultados obtenidos. Los supuestos son las condiciones que deben cumplirse para que el modelo se adecúe a los datos en cuestión. El modelo es sometido a pruebas de supuestos que tienen por objetivo evaluar si las premisas subyacentes, como son entre otras, la linealidad, la homocedasticidad y la independencia de errores, son satisfechas por los datos observados.

La verificación de supuestos proporciona al investigador una base sólida de la que partir para la interpretación y generalización de los resultados, permitiendo a los investigadores comprender en qué medida y con qué limitaciones el modelo se ajusta a la realidad. Con todo ello, este proceso verificador nos permite adoptar de enfoques correctivos o refinamientos del modelo en caso de que alguno de los supuestos no se cumpla a priori, mejorando de esta forma la calidad y fiabilidad de los análisis de regresión.

Para poder verificar y contrastar la validez de los supuestos en un modelo de regresión, debemos recurrir a una combinación de pruebas gráficas y de teóricas que nos ofrecen una visión integral para nuestra posterior evaluación. Las pruebas gráficas de residuos versus valores ajustados o gráficos Q-Q, nos proporcionan una representación visual de la dispersión real de los residuos y su distribución, además de permitirnos identificar patrones poco usuales o no lineales en los propios residuos, así como verificar si la varianza de los mismos es constante a lo largo del rango de valores ajustados, esto se conoce como el supuesto clave de la homocedasticidad. Por último, las pruebas gráficas nos pueden señalar adicionalmente la presencia de valores atípicos o la influencia excesiva de ciertas observaciones en el modelo.

4.3.1 Heterocedasticidad

En las gráficas presentadas a continuación, podemos observar cómo, al menos, aparentemente, aunque no hayan sido proporcionados muchos datos en el conjunto de entrenamiento (35 observaciones), los datos parecen ajustarse y cumplir con el supuesto, mostrando una dispersión aparentemente constante a través de la salida residual (Figura 3). Por otro lado, las pruebas teóricas se basan en su mayoría en análisis estadísticos y pruebas formales para evaluar la adecuación de los residuos a las distribuciones esperadas teóricamente.

Por su parte, para evaluar la homocedasticidad, la prueba de Breusch-Pagan es esencial en un modelo de regresión, ya que es capaz de detectar si la varianza de los residuos es constante en diferentes niveles de las variables predictoras. Si la prueba resulta significativa, estaremos ante una violación de la homocedasticidad, lo que podría hacer que la validez de las inferencias estadísticas del modelo se viera afectada.

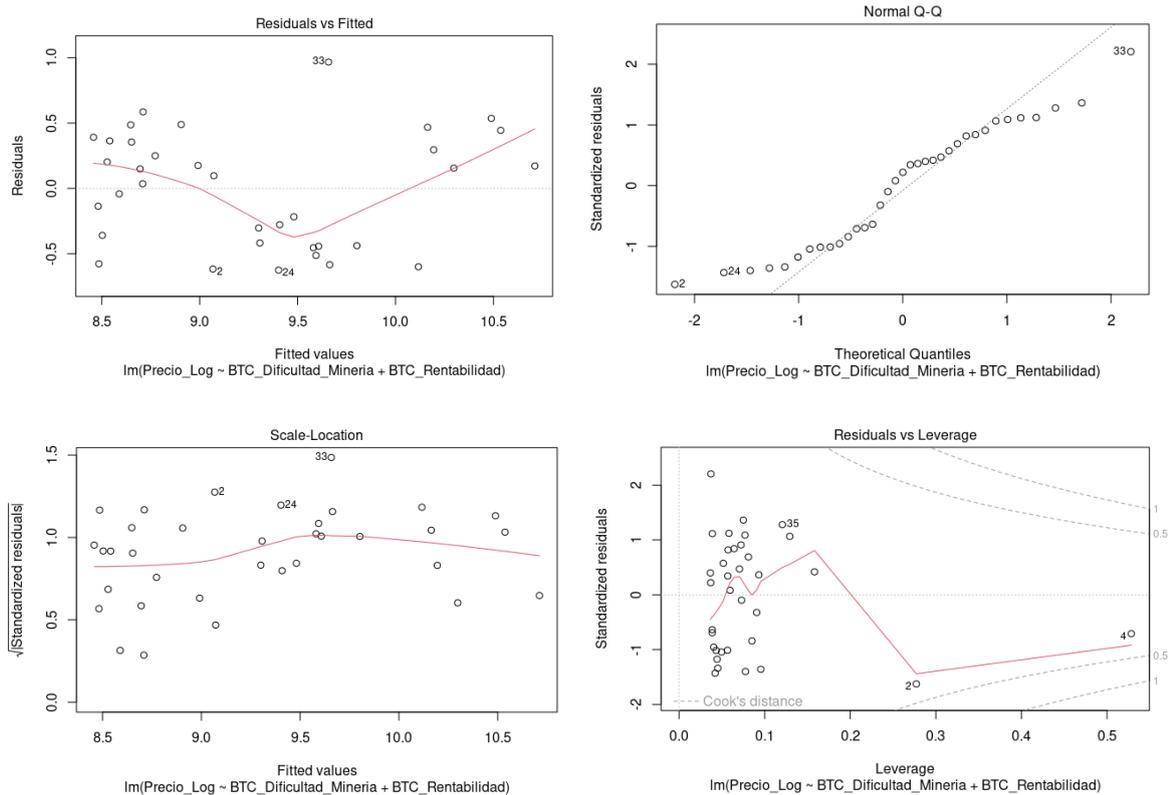


Figura 3. Pruebas gráficas validación de supuestos. Elaboración propia.

La hipótesis nula de la prueba de Breusch-Pagan enuncia que hay heteroscedasticidad y el valor de $p > 0.05$ ($p\text{-value} = 0.4878$) (es no significativo), por lo que rechazamos la hipótesis nula (H_0) y concluimos que los residuos del modelo son homocedásticos.

4.3.2 Normalidad

La normalidad es uno de los supuestos más relevantes en la regresión lineal, la cual establece que los residuos del modelo deben tener una distribución normal. Esto es esencial para poder garantizar la validez y eficacia de las pruebas de significación y los intervalos de confianza de los coeficientes del modelo. La normalidad se puede verificar mediante pruebas estadísticas como la prueba de normalidad de Shapiro-Wilk o, también desde una perspectiva gráfica mediante el análisis de los residuos en un gráfico de cuantiles normales (Hernández Martín, 2012). Si los residuos no siguen una distribución normal, los resultados pueden resultar sesgados, llevándonos a conclusiones potencialmente incorrectas.

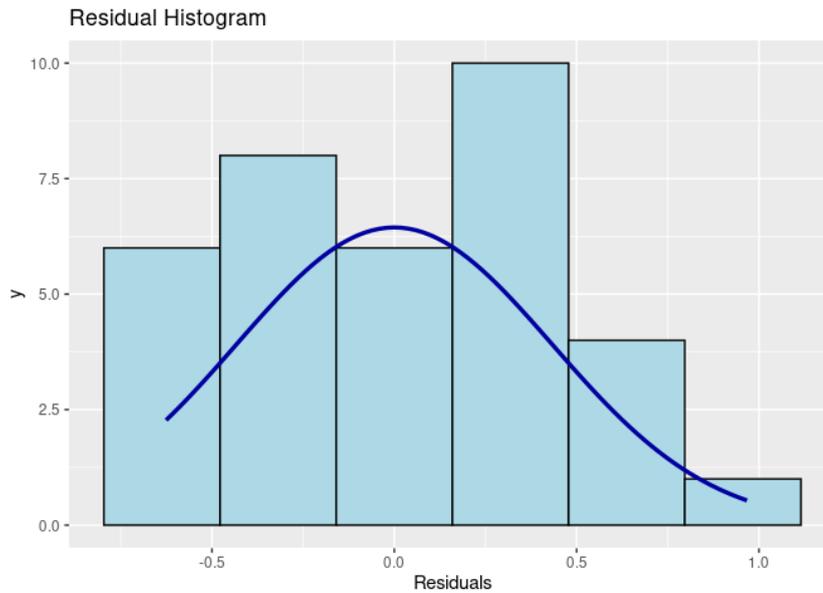


Figura 4. Pruebas gráficas para la normalidad de los residuos. Elaboración propia.

La mayor parte de los residuos tienen un resultado alrededor de cero y, junto con ello, el número de observaciones percibido en las colas del histograma es bajo. Atendiendo a nuestro modelo de regresión, podemos afirmar que tiene una distribución bastante normal (Figura 4).

Prueba	Estadístico	P-valor
Shapiro-Wilk	0.9379	0.0482
Kolmogorov-Smirnov	0.12	0.6507
Cramer-von Mises	4.0129	0.0000
Anderson-Darling	0.77	0.0411

Tabla 2. Pruebas teóricas normalidad de residuos. Elaboración propia.

Como podemos observar en la tabla precedente, en el análisis de normalidad de los residuos han sido aplicadas varias pruebas estadísticas, incluyendo test de Shapiro-Wilk, Kolmogorov-Smirnov, Cramer-von Mises y Anderson-Darling. Dado que el tamaño de la muestra en el conjunto de los datos de prueba es menor a 50, elegimos la prueba de Shapiro-Wilk. Los resultados han revelado que el p-valor obtenido es inferior a 0,05, lo cual nos lleva a rechazar la hipótesis nula (H_0). Este resultado nos está indicando que los residuos están distribuidos normalmente, lo que supone un elemento crucial para las inferencias estadísticas del modelo. Pese a que la normalidad se cumple de manera marginal, este hecho continúa siendo favorable para los propósitos del análisis.

4.3.3 Multicolinealidad

Cuando variables predictoras en un modelo están altamente correlacionadas entre sí, decimos que se está produciendo multicolinealidad entre ellas. Este hecho nos dificulta a la hora de

atribuir efectos individuales a las variables. En el contexto del modelo previo, este tipo de interrelaciones pueden afectar a la interpretación y estabilidad de los coeficientes de regresión.

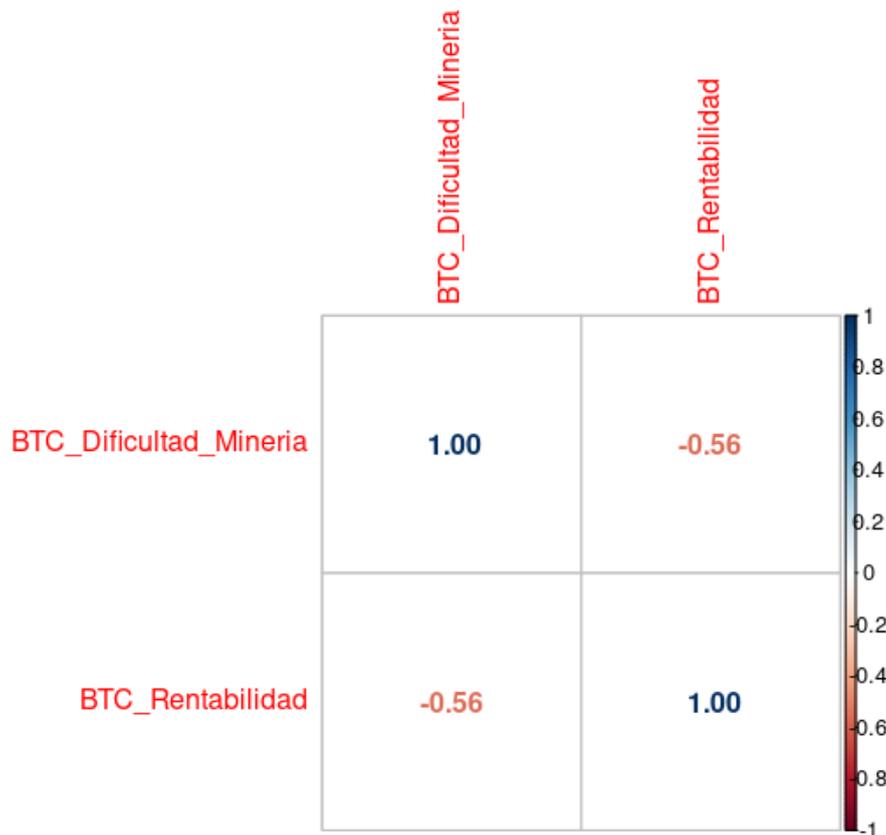


Figura 5. Matriz correlación del modelo resultante. Elaboración propia.

Confirmamos con esta prueba que no hay presencia de multicolinealidad sobresaliente con las variables restantes en el modelo. De hecho, se da una correlación moderadamente negativa. Con ello determinamos que una disminución en la dificultad minera nos lleva a un aumento en la rentabilidad minera, y viceversa. Un aumento en la dificultad minera indica una disminución en la rentabilidad minera.

4.3.4 Independencia de los errores

Es esencial para asegurar que el modelo de regresión sea confiable y que sus estimaciones sean válidas, que se dé la independencia de los errores. Para este contexto, los residuos no pueden mostrar patrones de autocorrelación con potencial para sesgar las inferencias. Si confirmamos la ausencia de multicolinealidad, estamos minimizando la posibilidad de que los errores estén relacionados entre sí. En nuestro caso, la falta de patrones discernibles en los residuos sugiere que no hay presencia de estructuras sistemáticas en los errores, lo que valida las suposiciones de independencia y de robustez del modelo en su predicción e interpretación.

4.4. PREDICCIÓN DEL PRECIO DE BITCOIN

En esta fase es donde aplicaremos el modelo de conjunto de entrenamiento en el conjunto de datos de prueba con el objetivo de validar si se ajusta o no a los datos. La raíz del error cuadrático medio, “*Root Mean Square Error*” (RMSE por sus siglas en inglés), determina si el modelo se ajusta bien a los datos reservados. El RMSE es la raíz cuadrada de la media del cuadrado de todo el error. El uso del Error Cuadrático Medio Raíz (RMSE) es ampliamente extendido y se reconoce como una métrica de error altamente efectiva y versátil para evaluar predicciones numéricas (Hernández Martín, 2012).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (S_i - O_i)^2}$$

donde O_i son las observaciones, S_i los valores predichos de una variable y n el número de observaciones disponibles para el análisis. El RMSE es una buena medida de precisión, enfocada principalmente en comparar errores de pronóstico de diferentes modelos para una variable particular, aunque no entre variables, ya que depende de la escala. Se ejecuta la función de predicción con el modelo desarrollado por el conjunto de entrenamiento.

El análisis (RMSE) y su comparación con el error estándar residual supone una evaluación esencial de la capacidad predictiva del modelo. En este contexto, el RMSE representa la raíz cuadrada de la media de los errores cuadrados entre los valores pronosticados por el modelo y los valores reales del conjunto de reserva. Al haber obtenido un valor de RMSE de 0,38, que es menor que el error estándar residual previamente calculado, que resultaba en 0,44 se sugiere que el modelo que ha sido construido en el conjunto de entrenamiento está mostrando una capacidad superior para predecir el conjunto de reserva a la del conjunto de entrenamiento.

El resultado obtenido ha de interpretarse como que el modelo generaliza bien en datos no vistos, ya que un RMSE más bajo nos indica que los errores en las predicciones del modelo en el conjunto de reserva son relativamente pequeños. Lo que es lo mismo, el modelo no solo se ajusta bien a los datos de entrenamiento, sino que también logra una buena adaptación a datos nuevos y no utilizados durante la fase de entrenamiento. Este hecho es fundamental para poder evaluar la capacidad de predicción del modelo en escenarios sucesivos del mundo real. En conjunto, el hecho de que el RMSE sea inferior que el error estándar residual, nos sugiere que el modelo tiene la capacidad de capturar relaciones y patrones entre las variables que son transferibles a datos no observados previamente.

En el mes de enero de 2022, la dificultad de minería de BTC se situó en aproximadamente 23,17, mientras que la rentabilidad de minería de BTC alcanzó alrededor de 0,19 (CoinGecko).

Si queremos calcular la predicción del precio, es esencial aplicar la función logarítmica inversa al resultado obtenido. Este proceso resulta imprescindible para transformar el valor predicho nuevamente a Bitcoin en USD, tal como habíamos realizado previamente la transformación logarítmica de los datos originales durante la fase de limpieza de datos.

$$PrecioLog = 7,69 + 0,12(24,17) + 0,78(0,19)$$

$$e^{PrecioLog} = 40.888,33$$

Hemos calculado el intervalo de confianza (IC) del 95% siendo este [\$8.197,59; \$14.289,80]. El precio real de Bitcoin en enero de 2022 estaba entre \$ 37.188,11 y \$ 38.483,13, por lo que, en conclusión, el modelo predice que el precio de Bitcoin será de 40.888,33 dólares en enero de 2022, con un intervalo de confianza del 95% de [8.197,59 dólares, 14.289,80 dólares]. De nuevo, encontramos que el R2 es 0,7129 y el error estándar es 0,4468. De forma intuitiva, dada la volatilidad de Bitcoin en el último año, observamos una diferencia entre la predicción del modelo y el precio real, de alrededor de \$2.405,2 a \$3.700,22. Con ello, podemos concluir satisfactoriamente que el modelo asume un rango razonable para extrapolar a corto plazo el precio futuro de Bitcoin.

5. DISCUSIÓN

5.1. CONTEXTO DE LOS RESULTADOS EN LA LITERATURA EXISTENTE

Cuando hablamos de la "Dificultad de minería" nos estamos refiriendo a un concepto esencial en el mundo de las criptomonedas, con mayor relevancia en el contexto del Bitcoin. Es la medida que refleja la dificultad o complejidad necesarias en el proceso de minería de nuevas unidades de criptomoneda. La minería es el proceso a través del cual los mineros digitales de la red resuelven complejas ecuaciones matemáticas con el objetivo de validar y agregar nuevas transacciones a la cadena de bloques. Al mismo tiempo, esto garantiza la seguridad e integridad de la red.

La dificultad de minería es un valor que se ajusta automáticamente por la red de Bitcoin aproximadamente cada dos semanas. Este ajuste se realiza con el objetivo de mantener constante el tiempo promedio que se necesita para encontrar un nuevo bloque en la cadena de bloques, alrededor de 10 minutos (BTC.com). Cuantos más mineros se unen a la red contribuyendo al aumento de la potencia de cómputo, más se incrementa la dificultad para generar nuevos bloques y más lento se vuelve el proceso de minado. De forma opuesta, si la

potencia de cómputo disminuye, la dificultad se ajusta a la baja para mantener el ritmo constante.

La discusión sobre la influencia de la dificultad de minería en el precio del Bitcoin está sobre la mesa. Diversos autores han planteado perspectivas fundamentadas en investigaciones concretas. Por ejemplo, (Antonopoulos 2014), discute cómo la dificultad de minería está relacionada con la oferta y la demanda de Bitcoin, y cómo los cambios en dicha dificultad pueden influir en el equilibrio entre la producción y la adopción de Bitcoin. El autor argumenta que la dificultad de minería puede también suponer un impacto para la percepción del valor de Bitcoin, ya que afecta directamente a la cantidad de recursos que son necesarios para producir nuevas unidades monetarias criptográficas y a la validación de transacciones en la red. Antonopoulos explora también, cómo los ajustes en la dificultad de minería pueden contribuir positivamente a la estabilidad del suministro y cómo los mineros digitales pueden responder a estos cambios tratando de mantener su rentabilidad. Aunque no solo nos ilustra sobre la influencia de la dificultad de minería en el precio, el autor presenta algunos argumentos sobre cómo de este factor pueden derivar ramificaciones en la dinámica de mercado y, por consiguiente, en el precio del Bitcoin.

Guizani y Nafti (2019), analizan la relación existente entre la dificultad de minería y el precio del Bitcoin. Estos autores sostienen que la dificultad de minería es una especie de "indicador adelantado" sobre el precio del Bitcoin, ya que suele aumentar antes de que el precio suba. Por su parte, Lindland (2020) expone en su tesis que la dificultad de minería desempeña un papel crucial en el precio del Bitcoin. A través de su análisis enuncia que la dificultad de minería tiende a aumentar paralelamente con el precio del Bitcoin. La razón es que los mineros son incentivados a invertir en hardware más potente para obtener mayores recompensas. Zhang et. al. (2022) exponen la importancia de la dificultad de minería como "factor clave" en la determinación del precio del Bitcoin. Nos argumentan que esta dificultad tiene un impacto directo en la oferta del Bitcoin, lo que a su vez puede determinar la dinámica de los precios.

En conclusión, muchos de los principales autores que han afrontado el tema que estamos analizando, concluyen argumentos como que la dificultad de minería es imperativa para mantener la estabilidad y la seguridad de la red de Bitcoin, además de su clara vinculación con el precio y sus fluctuaciones. Cuantos más mineros compiten por resolver los desafíos matemáticos, un mayor poder de cómputo y energía son requeridos, lo que a su vez contribuye a la seguridad y la inmutabilidad de las transacciones.

En resumen, podemos afirmar que la dificultad de minería es un indicador clave que arroja luz sobre el nivel de competencia y esfuerzo necesarios para participar en la minería de Bitcoins, lo cual impacta en la emisión y la seguridad de la criptomoneda.

Analizando la literatura sobre nuestra otra variable principal, la "Rentabilidad de minería" nos damos cuenta de que es un concepto fundamental en el ámbito de las criptomonedas y, principalmente en el contexto de la minería de Bitcoin. La Rentabilidad de minería se refiere al nivel de ganancias que pueden obtener los mineros cuando participan en la validación y verificación de transacciones en la red de criptomonedas, mediante el proceso de minería. La rentabilidad de la minería se ve influenciada por varios elementos clave, como son el precio actual de la criptomoneda, la dificultad de minería, los costes operativos (la electricidad y el hardware) y la recompensa obtenida por bloque minado. Los mineros digitales tratan además de averiguar si el ingreso potencial que pueden obtener por la emisión de nuevas monedas y las tarifas de transacción será mayor que los costes asociados de la energía y el equipo necesario para llevar a cabo el proceso de minería.

Para el caso del Bitcoin, su rentabilidad de minería puede fluctuar ampliamente dada la volatilidad del precio de la criptomoneda y los ajustes periódicos producidos en la dificultad de minería. Si el precio de Bitcoin aumenta, los mineros pueden obtener mayores ingresos, pero si el precio cae, la rentabilidad es probable que disminuya. Por otra parte, si la dificultad de minería es ajustada al alza, será requerido más poder de cómputo y, por ello, más energía, lo que puede suponer un determinante para la rentabilidad de la minería, la cual es esencial para el funcionamiento de la red de criptomonedas porque incentiva a los mineros a contribuir con sus recursos y poder de cómputo para además mantener la seguridad e integridad de la cadena de bloques.

Un autor que ha abordado el tema de la influencia de la rentabilidad de la minería en el precio del Bitcoin es Saifedean Ammous en su libro "*The Bitcoin Standard: The Decentralized Alternative to Central Banking*". En su libro, Ammous explora cómo la rentabilidad de la minería efectivamente puede afectar a la oferta y la demanda de Bitcoin, y cómo ello puede influir causalmente en el precio. El autor sostiene cómo la rentabilidad de la minería está directamente relacionada con la recompensa que los mineros reciben a la hora de verificar transacciones y agregar nuevos bloques a la cadena. Parece factible que esta recompensa, junto con las llamadas tarifas de transacción, puede afectar la oferta circulante de Bitcoin y, por ende, tener un impacto en su precio final. Ammous discute además cómo la rentabilidad de la minería afecta a la percepción del valor de Bitcoin por parte de los mineros y cómo estos actores pueden tomar decisiones en función de la rentabilidad y la competencia en el mercado. El autor proporciona *insights* sobre cómo esta variable puede interactuar con otros elementos económicos para acabar influyendo en el precio del Bitcoin.

Tenemos ya claro que la influencia del precio del Bitcoin en la rentabilidad de la minería de BTC es innegable, debido a que existe una relación directa entre el valor del BTC y los

beneficios obtenidos por los mineros. Se ha observado que, en periodos de mercados bajistas, los mineros están especialmente atentos a las fluctuaciones del precio del BTC, ya que enfrentan riesgo de pérdida si su valor cae por debajo de un determinado umbral crítico. Adicionalmente, la rentabilidad minera se ve condicionada por la recompensa por bloque, que representa la cantidad de BTC aportada a los mineros por cada bloque minado en la cadena de BTC. Podemos añadir que la recompensa por bloque de Bitcoin viene experimentando también varias reducciones a lo largo del tiempo, alcanzando su valor actual de 6.5 BTC tras múltiples *halvings* de recompensa. Estos *halvings*, se producen cada cuatro años aproximadamente y tienen por objetivo limitar la entrada de nuevas monedas a la red, por lo que tienen como consecuencia un papel clave en la relación entre la rentabilidad de la minería y el precio del Bitcoin. En conjunto, estos factores demuestran cómo la rentabilidad de la minería ejerce una influencia significativa en la dinámica del precio del Bitcoin (Partz, 2022). En resumen, podemos exponer que la rentabilidad de la minería es un indicador esencial capaz de determinar si la actividad de minería de criptomonedas es financieramente viable o no para los participantes. Por su parte, los costes operativos y la dificultad de minería, juegan un papel fundamental en el mantenimiento y la seguridad de la red de criptomonedas.

5.2. IMPLICACIONES PRÁCTICAS DE LOS HALLAZGOS

En este capítulo, vamos a exponer cómo los resultados obtenidos en este proyecto tienen una serie de implicaciones prácticas que se interrelacionan en múltiples ámbitos del mundo financiero e inversores y participantes activos del ecosistema de las criptomonedas. Las relaciones entre las variables seleccionadas nos aportan un enfoque cuantitativo sólido para entender los factores clave que afectan a la valoración del Bitcoin. A continuación, exponemos algunas de las implicaciones prácticas más relevantes:

Inversores y Stakeholders Los resultados y hallazgos obtenidos tienen el potencial de aportar nueva información de valor, además de funcionar como guía ante las decisiones de inversión en el mercado de criptográfico. La relación positiva y estadísticamente significativa encontrada entre la rentabilidad de minería, la dificultad de minería, y el precio del Bitcoin, nos lleva a entender que los inversores pueden beneficiarse si tienen en cuenta estas variables al cuando evalúan el potencial de inversión del Bitcoin. Una comprensión más arraigada sobre cómo estos factores influyen el precio logarítmico del Bitcoin, podría proporcionar a los inversores y principales interesados una ventaja estratégica a la hora de tomar decisiones de compra o venta. Los inversores pueden tener en cuenta estos hallazgos a la hora de pronosticar sus estrategias de inversión y gestionar sus carteras de criptomonedas.

Formuladores de Políticas y Reguladores: Para los encargados de formular políticas y regular el mercado de criptomonedas, los resultados ofrecidos también tienen implicaciones

favorables. El hecho de haber identificado variables influyentes al precio del Bitcoin puede ayudar en potencia a los reguladores a comprender mejor las dinámicas del mercado, con el objetivo de diseñar marcos regulatorios más adecuados. Creemos, además, que la consideración de estas variables para la toma de decisiones puede suponer una contribución para la estabilidad y la transparencia del mercado, ayudando a promover un entorno más seguro para los inversores e interesados. Los formuladores de políticas pueden servirse de estos hallazgos para profundizar y estudiar cómo evitar posibles desafíos y riesgos asociados con la volatilidad y la fluctuación de los precios de las criptomonedas.

Analistas Financieros y Económicos: Los resultados expuestos y la metodología empleada pueden resultar de guía y valor para los analistas financieros y económicos que buscan entender las fluctuaciones de este mercado tan particular, así como predecir el comportamiento del Bitcoin. La relación entre las variables independientes y el precio del Bitcoin puede ser incorporada en modelos de pronóstico más sofisticados. Profundizar en estas relaciones con valores de mercado concretos, puede permitir a los analistas realizar proyecciones más precisas sobre el precio futuro del Bitcoin y mejorar la validez de sus recomendaciones. Estos hallazgos también pueden inspirar novedosas investigaciones más detalladas y análisis de la interacción entre las variables con potencial para influenciar y otros indicadores económicos.

Educación y Concienciación Pública: Esta investigación también puede ofrecer sus conclusiones y contribuir a la labor de educación y concienciación pública sobre las criptomonedas. Una vez que hemos comprendido y comprobado cómo determinados factores como la dificultad de minería y la rentabilidad de minería afectan sustancialmente al precio del Bitcoin, aquellos individuos potencialmente interesados en las criptomonedas pueden tomar decisiones más informadas. Esta divulgación puede ser especialmente valiosa para aquellos que están considerando la inversión en criptomonedas por primera vez, aportándoles una comprensión mucho más completa y lógica de los factores que influyen realmente en la valoración de Bitcoin y otras criptomonedas.

En resumen, los hallazgos de nuestra investigación contribuyen, no solo a la comprensión del comportamiento del precio del Bitcoin, sino que sostienen implicaciones prácticas de gran alcance. Al haber proporcionado una visión más precisa y transparente de las variables que influyen en la valoración del mercado del Bitcoin, esta investigación puede ofrecer ayudar de valor a los inversores, reguladores, analistas y al público en general a la hora de tomar decisiones más lógicas y estratégicas en el cambiante y dinámico mundo de las criptomonedas.

5.3. LIMITACIONES DEL ESTUDIO Y FUTURAS INVESTIGACIONES

Las criptomonedas están en constante evolución y, previsiblemente continuarán sus fluctuaciones hasta que sean totalmente reguladas. En este contexto, nuestro estudio se posiciona como una aportación valiosa que ayuda a comprender las complejidades que subyacen a la relación entre variables clave y el precio del Bitcoin. Al estudiar cómo la dificultad de minería, la actividad minera de Bitcoin, la tasa de inflación global y otras variables influyen en mayor o menor medida en la valoración de la criptomoneda líder, nuestra contribución proporciona una base sólida para evaluar la interacción lógica del ecosistema de las criptomonedas.

Cuando comparamos nuestros hallazgos con otras investigaciones precedentes, encontramos objetivos de investigación con características en común en algunos casos, aunque también vemos destacar la relevancia de nuestro enfoque. Si bien algunos estudios anteriores han explorado factores individuales para la determinación del precio del Bitcoin, tomando como referencias el volumen de transacciones o el sentimiento del mercado, nuestro estudio busca un panorama más completo al considerar múltiples variables simultáneamente. Pese a que encontramos conexiones con respecto a la influencia de variables como la dificultad de minería, nuestras conclusiones resaltan cómo la interacción entre estas variables puede ser más significativa de lo que se ha discutido en investigaciones anteriores.

A pesar de los resultados prometedores y de los esfuerzos de este estudio, es necesario reconocer y considerar todas las limitaciones que pueden haber influido en la interpretación de los resultados y en la generalización de las conclusiones del proyecto. Estas limitaciones responden a áreas que merecen un mayor enfoque y un análisis más profundo en futuras investigaciones.

En primer lugar, el presente estudio se ha basado en una muestra específica de datos y período de tiempo definido, lo que puede limitar la generalización de los hallazgos a diferentes contextos y momentos temporales. Como hemos podido confirmar, las dinámicas del mercado de criptomonedas son altamente volátiles y están sujetas a cambios bruscos relacionados directamente con eventos macroeconómicos, tecnológicos y legislativos inesperados, como regulaciones gubernamentales o avances digitales. Por todo ello, la representatividad de los resultados obtenidos podría ser limitada, especialmente si el comportamiento de las variables influyentes cambia en el futuro.

Aunque hayamos utilizado diversas variables predictoras en el modelo de regresión, esto no exime, por supuesto, que otros factores no incluidos en este estudio puedan influir sustancialmente en el precio del Bitcoin. Las criptomonedas son un campo en constante cambio y transformación y es por ello probable que nuevos factores, como avances

tecnológicos, adopción masiva o cambios en la percepción pública o privada, también desempeñen un papel muy relevante en la determinación de los precios del Bitcoin.

Una limitación clave que manifestamos en este estudio ha sido la disponibilidad de los datos utilizados. Los datos históricos sobre criptomonedas pueden verse a menudo sujetos a errores y omisiones, lo que podría afectar la precisión y validez de los resultados. Adicionalmente a esto, algunas variables importantes pueden haber sido excluidas debido a la falta de datos confiables o consistentes de las fuentes estudiadas, lo que podría haber limitado la capacidad del modelo para capturar todos los aspectos potencialmente influyentes.

En lo que respecta a futuras investigaciones, sería oportuno abordar las limitaciones mencionadas ampliando el enfoque en varias direcciones. Por una parte, se podría realizar un estudio longitudinal que analice los cambios en las relaciones entre variables a lo largo del tiempo, teniendo en cuenta la evolución del mercado y los eventos clave. También se podría explorar la influencia de factores no considerados en este estudio, como la adopción masiva por parte de instituciones financieras en un futuro reciente o la aparición de nuevas tecnologías en el sector Blockchain que, por ejemplo, facilitarían las labores de minería. Todo ello podría proporcionar una comprensión más completa del comportamiento del precio del Bitcoin.

Otra cuestión muy interesante para futuras investigaciones podría ser la aplicación de técnicas más avanzadas de análisis como, por ejemplo, los modelos no lineales, las series de tiempo y enfoques de aprendizaje automático. Estas metodologías podrían revelar nuevos patrones y relaciones más sutiles escondidas en los datos que podrían haber quedado sin descubrir en este estudio.

En conclusión, si bien este estudio ofrece información valiosa sobre los factores de influencia en el precio del Bitcoin, es importante reconocer sus limitaciones y considerarlas al interpretar los resultados. Las futuras investigaciones deberían tener en cuenta estas limitaciones y explorar nuevas áreas para continuar mejorando nuestra comprensión de los apasionantes determinantes de los precios de las criptomonedas.

6. COMPARACIÓN CON OTROS MÉTODOS DE PREDICCIÓN

La constante búsqueda por mejorar la eficacia de las predicciones en los precios de las criptomonedas nos conduce a diferentes enfoques y metodologías que han surgido para abordar este desafío. En esta sección, exploramos una comparación entre el enfoque de regresión lineal múltiple utilizado en nuestra investigación y otros métodos populares de predicción en el ámbito de las criptomonedas. Al comprender las fortalezas y debilidades de

cada enfoque, podremos obtener una perspectiva más amplia y completa sobre la eficacia y aplicabilidad del modelo que hemos desarrollado.

6.1. MODELOS DE APRENDIZAJE AUTOMÁTICO: FLEXIBILIDAD Y RIESGO DE SOBREAJUSTE

En el mundo criptográfico, la predicción de precios se ha convertido en un campo de análisis de enorme relevancia. Entre las múltiples metodologías empleadas para abordar este reto, los modelos de aprendizaje automático han surgido como poderosas herramientas prometedoras. A diferencia de otros enfoques tradicionales, los modelos de aprendizaje automático tienen la habilidad de capturar patrones y relaciones complejas en los datos, independientemente de si son lineales o no. Su flexibilidad implícita en estas técnicas permite una capacidad de adaptación y de predicción más avanzadas en los entornos altamente volátiles en los que se mueven las criptomonedas.

Un aspecto diferenciador de los modelos de aprendizaje automático es su capacidad para aprender a partir de los datos. Estos modelos son alimentados a partir de enormes conjuntos de datos históricos a partir de los cuales pueden detectar sutilezas y tendencias que podrían pasar desapercibidas para métodos más tradicionales o simples. Esto es especialmente relevante en el caso de las criptomonedas, donde los patrones de precios pueden verse influenciados por múltiples factores, incluyendo cambios regulatorios, eventos geopolíticos y el comportamiento del mercado en general. Los modelos de aprendizaje automático pueden identificar relaciones no evidentes en apariencia y generar pronósticos más precisos y complejos al incorporar estas y nuevas variables en su análisis.

No obstante, la flexibilidad de los modelos de aprendizaje automático viene acompañada de un desafío importante, conocido como el riesgo de sobreajuste. El sobreajuste ocurre cuando un modelo se ajusta excesivamente a los datos de entrenamiento y es capaz de capturar incluso el ruido o la variabilidad aleatoria. Este hecho puede ser excelente en su capacidad para ajustar los datos utilizados para el entrenamiento, pero al mismo tiempo, puede suponer una falta de generalización a nuevos datos no vistos. En el contexto de las criptomonedas, donde la volatilidad y la incertidumbre son constantes, el riesgo de sobreajuste puede ser especialmente perjudicial.

Para reducir el riesgo de sobreajuste, los practicantes deben aplicar técnicas de regularización y validación cruzada. La regularización implica imponer restricciones en los parámetros del modelo para evitar que se vuelvan excesivamente complejos; mientras que la validación cruzada ayuda a evaluar el rendimiento del modelo en datos no utilizados anteriormente en el entrenamiento. Junto con ello, la selección cuidadosa de las características y la consideración de la existencia de un equilibrio entre flexibilidad y simplicidad son esenciales. Un modelo que

sea demasiado complejo podría memorizar los datos de entrenamiento en lugar de generalizar patrones subyacentes (Vargas, 2018).

En último lugar, los modelos de aprendizaje automático ofrecen un conjunto de herramientas de gran potencial para abordar el desafío de predecir los precios de las criptomonedas. Gracias a su capacidad para adaptarse a patrones complejos y no lineales, estos modelos pueden proporcionar pronósticos más sólidos y precisos (Vargas, 2018). Sin embargo, es importante reconocer que su flexibilidad implica el riesgo de sobreajuste, lo que aviva la importancia de técnicas de regularización y validación cruzada en su implementación. Abordando estas consideraciones, los modelos de aprendizaje automático pueden convertirse en aliados determinantes para la búsqueda de comprensión y capacidad de anticipación ante los movimientos de los precios en el apasionante mundo de las criptomonedas.

6.2. ANÁLISIS DE SERIES DE TIEMPO: TENDENCIAS TEMPORALES

En los estudios sobre los precios de las criptomonedas, el análisis de series de tiempo se posiciona como una herramienta esencial para desenmascarar las complejidades de las variaciones temporales de los datos. Las criptomonedas, como el Bitcoin, muestran patrones de comportamiento que fluctúan a lo largo del tiempo, debido a múltiples factores, como la adopción, la demanda, los eventos del mercado y las noticias relevantes asociadas con tendencias y eventos macroeconómicos. El análisis de series temporales nos brinda un enfoque robusto con el que modelar y predecir las tendencias cambiantes, permitiendo a los investigadores y analistas desentrañar los secretos ocultos en los datos.

Una de las características principales del análisis de series temporales es su foco en la estructura temporal de los datos. Es decir, a diferencia de los modelos de regresión más tradicionales, que asumen que las observaciones son independientes entre sí, los modelos de series de tiempo vienen a reconocer que los valores pasados influirán previsiblemente en los valores futuros y por ello buscan capturar estas dependencias temporales. Este hecho es especialmente condicionante en el contexto de las criptomonedas, cuyos precios a menudo nos muestran patrones cíclicos y recurrentes que pueden ser difíciles de trazar mediante enfoques estáticos.

Una de las técnicas clave en este tipo de análisis es la descomposición, que permite al investigador separar una serie en sus componentes fundamentales: tendencia, estacionalidad y residuo. La tendencia representa la dirección general en la que se mueven los datos a lo largo del tiempo, la estacionalidad refleja patrones repetitivos a intervalos regulares y el residuo captura las fluctuaciones no explicadas por las dos primeras componentes (Mauricio, 2007). Al descomponer la serie, los analistas pueden identificar tendencias a largo plazo y

patrones estacionales, lo que ofrece una comprensión mucho más profunda de los factores subyacentes que influyen en los precios.

Contextualizado en el Bitcoin, el análisis de series de tiempo también puede emplearse para identificar anomalías o eventos inusuales. Los picos o caídas abruptas reflejadas en los precios pueden indicar eventos relevantes en el ecosistema de las criptomonedas, como anuncios de regulaciones, adopción masiva o hackeos, entre otros. Los eventos pueden relacionar patrones distintivos en las series de tiempo, permitiendo a los analistas detectarlos y evaluar su impacto en los precios.

Es interesante destacar que el análisis de series de tiempo también cuenta con sus desafíos implícitos. La volatilidad inherente de las criptomonedas puede introducir ruido en las series de tiempo, dificultando así la identificación de patrones significativos (Mauricio, 2007). Además, la extrapolación a largo plazo puede ser incierta debido a la naturaleza altamente dinámica de las criptomonedas. Sin embargo, a pesar de estos desafíos, el análisis de series de tiempo continúa siendo una herramienta esencial para comprender y predecir los movimientos de precios en este emocionante y cambiante campo (Mauricio, 2007).

En conclusión, el análisis de series temporales emerge como una técnica esencial en la investigación de los precios de las criptomonedas. Cuando capturamos tendencias temporales, patrones estacionales y eventos significativos, este enfoque nos ofrece una visión profunda de las complejidades del mercado criptográfico. A través de la descomposición de series, la identificación de eventos inusuales y la modelización de dependencias temporales, el análisis de series de tiempo se identifica como un aliado útil y de gran potencial para investigadores y analistas en su búsqueda de conocimiento y respuestas sobre los movimientos de precios de las criptomonedas.

6.3. MODELOS BASADOS EN SENTIMIENTOS Y FUENTES EXTERNAS

Para las criptomonedas, la integración de factores externos se ha convertido en una estrategia recurrente con la que entender y predecir los movimientos del mercado. Los modelos basados en sentimiento y fuentes externas, como noticias de índole variada, han surgido con un enfoque innovador que trata de aprovechar la interacción entre los eventos del mundo real, las emociones de los inversores y curiosos del mercado, junto con las fluctuaciones de los precios de las criptomonedas. En esta sección, exploraremos cómo estos modelos pueden albergar una nueva dimensión para el análisis de los precios del Bitcoin y otras criptomonedas. La idea central detrás de estos modelos es que los eventos externos, como anuncios regulatorios, informes económicos, desarrollos tecnológicos y noticias en los medios de comunicación, pueden ejercer una influencia significativa en la percepción del valor de los

activos por parte de los consumidores, influyendo directamente en los precios de las criptomonedas. Estos eventos, como sabemos, pueden desencadenar respuestas emocionales entre los inversores, donde juegan un papel muy importante el miedo o la estimulación ante nuevos beneficios, lo cual se refleja en sus decisiones de compra y venta. El hecho de capturar estas señales emocionales y su impacto en los precios puede proporcionar una ventaja crucial en la predicción del mercado.

El primer paso para construir estos modelos consiste en recopilar y analizar una amplia variedad de datos de fuentes de noticias y de redes sociales. Mediante técnicas de procesamiento de lenguaje natural (NLP), hoy día ya pueden ser analizados de una manera automática ingentes cantidades de datos. Los investigadores pueden evaluar el tono y el sentimiento de los artículos, tweets o publicaciones en línea relacionadas con las criptomonedas. Un sentimiento positivo puede indicar un optimismo general en el mercado, mientras que un sentimiento negativo puede indicar incertidumbre o preocupación.

Una vez capturado el sentimiento, éste se puede incorporar en modelos predictivos para evaluar su impacto en los precios finales. Esto se hace normalmente mediante la construcción de indicadores de sentimiento que cuantifican la proporción de noticias positivas y negativas en un período de tiempo determinado. Con la reciente irrupción de la Inteligencia Artificial (IA), los analistas tienen este trabajo mucho más sencillo, puesto que podemos encontrar modelos de algoritmos pre-entrenados con los que proyectar conjeturas en nuestras bases de datos de una forma casi inminente. Los indicadores mencionados se utilizan como variables independientes en modelos de regresión, aprendizaje automático u otras técnicas estadísticas para predecir los precios futuros.

Es importante señalar que, si bien los modelos basados en sentimiento y noticias ofrecen una nueva perspectiva, también presentan desafíos únicos. La volatilidad y la rapidez con la que cambian las noticias pueden dificultar la recopilación y el procesamiento en tiempo real. Además, la relación entre el sentimiento y los precios puede ser compleja y no lineal, lo que requiere métodos avanzados de análisis y modelización.

En resumen, los modelos basados en sentimientos y noticias abren una nueva puerta hacia el análisis de precios de criptomonedas, desde una perspectiva mucho más social y donde pueden considerarse variables muy interesantes de analizar, como la dependencia digital o la vulnerabilidad presente hoy día ante fenómenos mediáticos. Incorporando factores externos y capturando las reacciones emocionales de los inversores, estos modelos pueden proporcionar una ventaja en la predicción del mercado. Aunque, bien es cierto, que enfrentan desafíos técnicos y teóricos, pero su potencial para mejorar la precisión en las predicciones y para comprender mejor la dinámica del mercado desde una perspectiva de impacto social, es

inegable. Con el mundo de las criptomonedas en constante evolución, los modelos basados en sentimiento y noticias se perfilan como una herramienta también esencial para los investigadores y analistas que buscan una comprensión detallada e integral de los factores que influyen en los precios de las criptomonedas.

6.4. REFLEXIONES Y CONSIDERACIONES FINALES

El habernos permitido la comparación entre diversos modelos y enfoques de predicción, nos permite apreciar la gran diversidad de métodos disponibles, incluso para un activo tan emergente, con capacidad para analizar y predecir los precios de las criptomonedas. El enfoque de regresión lineal múltiple utilizado en esta investigación ofrece una ventaja en términos de interpretabilidad, ya que nos permite discernir la relación entre variables independientes y la variable dependiente. Si bien entendemos que este enfoque puede no capturar todas las complejidades de los movimientos de precios, sí podemos afirmar que proporciona un marco sólido para entender las influencias de las variables seleccionadas.

En última instancia, la elección del enfoque de predicción depende principalmente de los objetivos específicos de la investigación y del nivel de complejidad que se busca alcanzar. Los diferentes métodos tienen sus propias fortalezas y debilidades, de todos podemos aprender y muchos pueden ser complementarios entre sí. En lugar de considerarlos competencia, es posible aprovechar sus características únicas para obtener una visión más completa y precisa de los precios de las criptomonedas en un entorno financiero en constante evolución.

7. CONCLUSIONES

Los hallazgos extraídos de esta investigación arrojan luz sobre las complejidades que afronta la valoración del Bitcoin y su relación con una serie de variables influyentes que se relacionan significativamente con su cambiante precio a lo largo del tiempo. A través de un riguroso análisis de regresión lineal múltiple, se ha logrado identificar qué factores tienen una correlación significativa con el precio de esta criptomoneda. Si bien el modelo desarrollado proporciona una visión inicial de esta relación, es necesario tener en cuenta sus limitaciones y considerar otras metodologías para una comprensión más profunda y precisa.

Uno de los aspectos que debemos resaltar es la demostración de que el precio del Bitcoin es influenciado en mayor medida por factores intrínsecos dentro de su propio ecosistema, en lugar de las fluctuaciones macroeconómicas externas. La significativa relación entre la dificultad de minería y la rentabilidad de la minería con el precio del Bitcoin subraya la importancia de la actividad minera en la determinación de su valor. Estos hallazgos respaldan

la idea de que el enfoque interno de Bitcoin, en lugar de las condiciones económicas generales, tiene un impacto dominante en su precio.

Del mismo modo, se ha evidenciado la limitada relevancia de variables como los cajeros automáticos BTC, la oferta monetaria M1 y la tasa de fondos federales en la predicción del precio del Bitcoin. Estos resultados sugieren que las interacciones y transacciones dentro del propio ecosistema criptográfico tienen un mayor peso en la determinación del precio en comparación con indicadores externos. Este enfoque intrínseco resalta la singularidad de las criptomonedas en su capacidad para operar de forma independiente de los parámetros tradicionales del sistema financiero.

En cuanto a las limitaciones, es importante reconocer que el modelo se basa en una muestra de datos relativamente pequeña, lo que puede afectar la solidez de las predicciones. Por lo tanto, se recomienda la inclusión de un conjunto de datos más amplio para fortalecer el modelo y permitir un análisis más profundo de las relaciones subyacentes del modelo. Adicionalmente, considerar otros enfoques de predicción, como los modelos basados en series temporales y el análisis de sentimiento y noticias, puede ofrecer una perspectiva más completa de los factores que influyen en el precio del Bitcoin.

En última instancia, esta investigación subraya la importancia de comprender la complejidad del mercado criptográfico y su relación con factores internos y externos. El Bitcoin, como líder en el ámbito de las criptomonedas, sigue desafiando las convenciones financieras y demuestra su capacidad para evolucionar en un entorno altamente volátil. Si bien esta investigación ofrece una visión valiosa, es necesario seguir explorando y analizando el mercado de las criptomonedas en constante cambio para comprender cada vez mejor sus tendencias y comportamientos a medida que se va estableciendo como una fuerza influyente en el panorama financiero global.

8. BIBLIOGRAFÍA

- Ammous, S. (2018). *The bitcoin standard: the decentralized alternative to central banking*. John Wiley & Sons.
- Antonopoulos, A. M. (2014). *Mastering Bitcoin: unlocking digital cryptocurrencies*. " O'Reilly Media, Inc."
- Báez León, L. X., & Reyes Molina, J. S. (2019). Bitcoin: análisis del precio desde las teorías de valoración de activos financieros. Trabajo de grado en Economía. Universidad de la Salle. Recuperado de <https://ciencia.lasalle.edu.co/economia/528>
- Baldrés Aguado, A. Monedas digitales: Origen y perspectivas desde un punto de vista social. Universidad Politécnica de Valencia, 2018. Tesis de grado, <https://riunet.upv.es/bitstream/handle/10251/113563/BALDR%C3%89S%20-%20Monedas%20digitales:%20Origen%20y%20perspectivas%20desde%20un%20punto%20de%20vista%20social.pdf?sequence=1>.
- Banco Mundial*. (2023). Tratto da Banco Mundial / *bancomundial.org*. <https://datos.bancomundial.org/indicador/FP.CPI.TOTL.ZG>
- Bianchi, D. (2020). Cryptocurrencies as an asset class? An empirical assessment. *The Journal of Alternative Investments*, 23(2), 162-179. <http://dx.doi.org/10.2139/ssrn.3077685>
- Bouoiyour, J., & Selmi, R. (2015). What does Bitcoin look like?. *Annals of Economics & Finance*, 16(2). <http://www.accessecon.com/Pubs/EB/2016/Volume36/EB-16-V36-I2-P82.pdf>
- Cadenas Jiménez, J. M. (2021). Bitcoin: Uso como medio de pago y posicionamiento en el mercado financiero.
- Carrera-López, J. S., Sánchez-Lunavictoria, J. C., and Loza-Torres, A. G. El uso de las criptomonedas como nueva forma de pago en la economía mundial. *Revista Científica FIPCAEC (Fomento de la investigación y publicación en Ciencias Administrativas, Económicas y Contables)*. ISSN: 2588-090X. Polo de Capacitación, Investigación y Publicación (POCAIP) 5, 5 (2020), 210–223.
- Raphael Auer, G. C. (2023). Crypto trading and Bitcoin prices: evidence from a new database of retail adoption. *BIS: Monetary and Economic Department*, 39.
- DIEGO CAGIGAS, J. C.-F.-G. (2020). Blockchain for Public Services: A Systematic Literature Review. *IEEE Access*.

- Dubey, P. (2022), "Short-run and long-run determinants of bitcoin returns: transnational evidence", *Review of Behavioral Finance*, Vol. 14 No. 4, pp. 533-544. <https://doi.org/10.1108/RBF-02-2022-0040>
- Duin, M. La institucionalización financiera del bitcoin y otras criptomonedas en el mercado económico mundial y su incidencia en el estado venezolano. *Adultus*, 1 (2021).
- Echeveste Parodi, C. Bitcoin y otros: el auge de las criptomonedas. Facultat D'Economia i Empresa, Universitat Rovira i Virgili, noviembre 2018. Tesis de grado, <http://hdl.handle.net/20.500.11797/TFG1693>
- Fry, J., & Cheah, E. T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47, 343-352. <https://doi.org/10.1016/j.econlet.2015.02.029>
- García Hernandez, J. M. (2018). Criptomonedas y Aplicación en la Economía. Trabajo fin de Máster Universitario en Finanzas. Universidad Pontificia Comillas. <http://hdl.handle.net/11531/32886>
- Guizani, S., & Nafti, I. K. (2019). The determinants of bitcoin price volatility: An investigation with ardl model. *Procedia computer science*, 164, 233-238.
- Hernández Martín, Z. (2012). *MÉTODOS DE ANÁLISIS DE DATOS (APUNTES)*. Universidad de la Rioja.
- IMF*. (2023). Tratto da IMF: <https://blogs.imf-formacion.com/blog/mba/masa-monetaria-dinero/>
- Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports*, 3(1), 3415. <https://doi.org/10.1038/srep03415>
- Kuhn, F. L. (2022). What factors drive interest in Cryptocurrencies? A cross-country analysis using Google Trends data. 51.
- Lakomski-Laguerre, O. Monnaie et immortalité: une autre histoire du bitcoin. *Economia History, Methodology, Philosophy*, 10-1 (2020), 145–154. <https://doi.org/10.4000/oeconomia.7702>.
- Lindland, P. (2020). What is driving the Bitcoin market? (Master's thesis, University of Stavanger, Norway).

- Mauricio, J. A. (s.d.). Introducción al Análisis de Series Temporales. *Universidad Complutense de Madrid*, 286.
- Vargas, M. A. (2018). Modelos de análisis de criptomonedas basado en aprendizaje automático y series temporales. *Universidad Central de Venezuela*, 91.
- Müller, Lukas, Ong, M. Aktuelles zum recht der kryptowährungen. *AJP Aktuelle Juristische Praxis* 29, 2 (2020), 198–212.
https://www.alexandria.unisg.ch/259362/1/MUELLER_ONG_AJP_02_2020_198-212.pdf.
- Nakamoto, S. Las criptomonedas: especial referencia al bitcoin, octubre 2008.
<https://nakamotoinstitute.org/bitcoin/>.
- Partz, H. (Julio 24, 2022). Not just Bitcoin price: Factors affecting BTC miner profitability. *Cointelegraph*. <https://es.cointelegraph.com/news/not-just-bitcoin-price-factors-affecting-btc-miner-profitability>
- Sandoval-Guzmán, K. M., López-Ortega, M. I., Domínguez-Rivera, P. E., y Rivera-Díaz, N. E. Perspectivas sobre utilización de criptomonedas para el financiamiento del endeudamiento público de la república de el salvador. *Revista Científica Multidisciplinaria de la Universidad de El Salvador-Revista Minerva* 3, 2 (2020), 92–105
- Sun, Y., Amanda, C. y Centana, B.C. (2023). The effect of hashrate, transaction volume, social media and macroeconomics on Bitcoin before and during the COVID-19 pandemic. *Asian Journal of Accounting Research*, 8(3), 293-306. <https://doi.org/10.1108/AJAR-10-2022-0319>
- Zhang, S., Li, M., & Yan, C. (2022). The Empirical Analysis of Bitcoin Price Prediction Based on Deep Learning Integration Method. *Computational Intelligence and Neuroscience*, 2022.

9. APARTADOS ANEXOS

Fecha	Precio	Precio_Log	Dificultad	Rentabilidad	ATM_US	Fed_Tasa	MI	TIG	cooks	d	outlier
2017-06-01	2547.32	7.842797	0.71	1.21	766.0	1.04	3524.9	3.23	0.1271788		delete
2017-07-01	2718.62	7.907880	0.86	0.88	826.0	1.15	3548.6	3.23	0.0497825		keep
2017-08-01	4676.14	8.450228	0.89	1.62	853.0	1.16	3587.8	3.23	0.0611313		keep
2017-09-01	4207.83	8.344702	1.10	0.84	908.0	1.15	3571.2	3.23	0.0000132		keep
2017-10-01	6130.56	8.721041	1.45	1.11	948.0	1.15	3606.3	3.23	0.0053589		keep
2017-11-01	10542.66	9.263185	1.35	2.08	1052.0	1.16	3631.0	3.23	0.0013759		keep
2017-12-01	13062.15	9.477474	1.87	2.28	1125.0	1.30	3610.6	3.23	0.0068056		keep
2018-01-01	9914.47	9.201751	2.60	0.91	1216.0	1.41	3656.2	3.59	0.0234163		keep
2018-02-01	10903.02	9.296795	3.01	0.85	1319.0	1.42	3619.0	3.59	0.0354220		keep
2018-03-01	6970.36	8.849422	3.46	0.46	1475.0	1.51	3663.4	3.59	0.0141478		keep
2018-04-01	9268.72	9.134401	4.02	0.62	1624.0	1.69	3655.2	3.59	0.0143643		keep
2018-05-01	7363.28	8.904261	4.31	0.44	1752.0	1.70	3650.8	3.59	0.0077092		keep
2018-06-01	6189.31	8.730579	5.08	0.31	1880.0	1.82	3652.2	3.59	0.0010366		keep
2018-07-01	8166.39	9.007782	5.95	0.34	1982.0	1.91	3677.5	3.59	0.0062702		keep
2018-08-01	6938.91	8.844900	6.73	0.28	1988.0	1.91	3693.2	3.59	0.0006102		keep
2018-09-01	6576.61	8.791275	7.15	0.25	2066.0	1.95	3701.8	3.59	0.0000162		keep
2018-10-01	6276.46	8.744561	7.18	0.23	2119.0	2.19	3728.2	3.59	0.0002029		keep
2018-11-01	3952.45	8.282091	6.65	0.15	2183.0	2.20	3705.6	3.59	0.0196632		keep
2018-12-01	3689.56	8.213263	5.62	0.17	2248.0	2.27	3763.4	3.59	0.0246806		keep
2019-01-01	3441.03	8.143526	5.81	0.17	2301.0	2.40	3754.8	3.47	0.0299279		keep
2019-02-01	3772.94	8.235610	6.07	0.16	2428.0	2.40	3762.9	3.47	0.0195361		keep
2019-03-01	4081.22	8.314151	6.38	0.17	2487.0	2.41	3733.6	3.47	0.0142992		keep
2019-04-01	5151.43	8.547030	6.35	0.20	2602.0	2.42	3775.0	3.47	0.0011489		keep
2019-05-01	8287.04	9.022448	7.46	0.27	2760.0	2.39	3783.2	3.47	0.0100899		keep
2019-06-01	12024.08	9.394667	7.93	0.37	2915.0	2.38	3824.4	3.47	0.0462735		keep
2019-07-01	9572.74	9.166675	9.01	0.32	3061.0	2.40	3856.8	3.47	0.0107447		keep
2019-08-01	9590.74	9.168553	10.18	0.25	3271.0	2.13	3858.9	3.47	0.0055603		keep
2019-09-01	8085.71	8.997854	12.76	0.16	3332.0	2.04	3900.0	3.47	0.0016219		keep
2019-10-01	9226.00	9.129781	13.69	0.16	3508.0	1.83	3935.3	3.47	0.0000842		keep
2019-11-01	7729.27	8.952770	12.97	0.14	3752.0	1.55	3955.3	3.47	0.0015848		keep
2019-12-01	7251.28	8.888933	12.95	0.14	3943.0	1.55	4007.1	3.47	0.0041594		keep
2020-01-01	9545.08	9.163781	15.47	0.15	4199.0	1.55	3993.6	3.18	0.0053137		keep
2020-02-01	8778.47	9.080057	15.49	0.13	4548.0	1.58	4002.9	3.18	0.0008524		keep
2020-03-01	6483.74	8.777053	13.91	0.12	4859.0	0.65	4261.9	3.18	0.0008725		keep
2020-04-01	8773.11	9.079447	15.96	0.15	5222.0	0.05	4779.8	3.18	0.0262619		keep
2020-05-01	9688.32	9.178676	15.14	0.08	5535.0	0.05	16232.9	3.18	0.0000022		keep
2020-06-01	9188.06	9.125660	15.78	0.07	5863.0	0.08	16563.5	3.18	0.0035458		keep
2020-07-01	11118.92	9.316403	16.85	0.11	6320.0	0.09	16763.8	3.18	0.0004197		keep
2020-08-01	11657.00	9.363662	17.56	0.09	6692.0	0.10	16896.7	3.18	0.0004485		keep
2020-09-01	10764.28	9.283988	19.31	0.08	7396.0	0.09	17161.5	3.18	0.0192599		keep
2020-10-01	13573.71	9.515890	20.00	0.13	8323.0	0.09	17365.3	3.18	0.0049156		keep
2020-11-01	18114.41	9.804463	19.16	0.12	9096.0	0.09	17626.7	3.18	0.0159129		keep
2020-12-01	28768.84	10.267048	18.60	0.22	26722.7	0.09	17834.5	3.18	0.1374381		delete
2021-01-01	34622.37	10.452255	20.82	0.24	44349.4	0.09	18107.1	4.35	0.0541379		keep
2021-02-01	46642.61	10.750270	21.72	0.27	61976.1	0.08	18367.8	4.35	0.0018537		keep
2021-03-01	58734.48	10.980782	21.87	0.39	79602.8	0.07	18641.4	4.35	0.0024732		keep
2021-04-01	53260.30	10.882946	23.58	0.36	97229.5	0.07	18927.7	4.35	0.0083401		keep
2021-05-01	35749.66	10.484296	21.05	0.23	114856.2	0.06	19259.0	4.35	0.0125123		keep
2021-06-01	35945.79	10.489767	19.93	0.24	132482.9	0.08	19318.8	4.35	0.0045121		keep
2021-07-01	41157.15	10.625153	14.49	0.36	150109.6	0.10	19497.1	4.35	0.0412099		keep
2021-08-01	47663.02	10.771911	17.62	0.35	167736.3	0.09	19746.0	4.35	0.0134071		keep
2021-09-01	41412.61	10.631341	18.99	0.34	185363.0	0.08	19898.4	4.35	0.0001254		keep
2021-10-01	61374.28	11.024746	21.66	0.36	202989.7	0.08	20061.8	4.35	0.0081369		keep
2021-11-01	57834.36	10.965338	22.34	0.30	220616.4	0.08	20273.9	4.35	0.0025818		keep
2021-12-01	47128.47	10.760633	24.27	0.29	238243.1	0.08	20423.6	4.35	0.0856546		delete

Tabla 3. Distancia de Cook

Código R

```
> library(readxl)           #para leer en el conjunto de datos del archivo de Excel
> library(tidyverse)       #para las funciones requeridas para transformar
datos (derivados de los paquetes dplyr y ggplot)
> library(caret)           #para conjuntos de entrenamiento y de test
> library(broom)           #para trabajar con la salida del modelo
> library(corrplot)        #to construct the required correlation matrix
> library(car)             #para construir la matriz de correlación requerida
> library(scatterplot3d)   #para trazado 3D
> library(olsrr)           #para comprobar residuos para distribución normal
> library(lmtest)          #para comprobar la homocedasticidad
> library(kableExtra)      #para mejorar el paquete kable
> library(ggpmisc)         #para determinar puntos críticos en la serie de tiempo
> #Elegir el directorio
> # Leer los datos
> Bitcoin <- read_excel("database.xlsx")
> #Compruebe el recuento de filas y los valores n/a o nulos.
> nrow(Bitcoin)

[1] 55
> #Cuenta los dtos faltantes por columna
> print(colSums(is.na(Bitcoin)))
Mes_Any          Precio BTC_Dificultad_Mineria
BTC_Rentabilidad      0          0          0
0

BTC_ATM_US          Fed_Tasa_US    M1_Oferta_Monetaria
Tasa_Inflacion_Global
0          0          0
0

> #Cuenta los datos nulos de la database
> is.null(Bitcoin)

[1] FALSE
>
> #La salida muestra 55 filas y ningún valor faltante o nulo.
>
> ##### Figura 2
> # Convertir la columna Mes_Any a formato Date
> Bitcoin$Mes_Any <- as.Date(Bitcoin$Mes_Any, format = "%b %Y")
> # Crear el gráfico
> ggplot(Bitcoin, aes(x = Mes_Any, y = Precio)) +
+   geom_line() +
+   stat_valleys(geom = "point", span = 11, color = "green", size = 4) +
+   stat_valleys(geom = "label", span = 11, color = "green", angle = 0,
+               hjust = -0.1, x.label.fmt = "%Y-%m-%d") +
+   stat_valleys(geom = "rug", span = 11, color = "green", sides = "b") +
+   stat_peaks(geom = "point", span = 11, color = "green", size = 4,
+             position = "identity", show.legend = FALSE) +
+   stat_peaks(geom = "label", span = 11, color = "green", angle = 0,
+             hjust = -0.1, x.label.fmt = "%Y-%m-%d",
+             position = position_nudge(x = 20), show.legend = FALSE) +
+   stat_peaks(geom = "rug", span = 11, color = "green", sides = "b",
+             position = "identity", show.legend = FALSE) +
+   annotate(geom = "point", x = as.Date("2021-10-01"),
```

```

+           y = 61374.28, color = "blue", size = 5) +
+   annotate(geom = "label", x = as.Date("2021-10-01"),
+           y = 61374.28, label = "2021-10", color = "blue",
+           angle = 0, hjust = -0.1, vjust = -0.5, size = 5) +
+   annotate(geom = "point", x = Bitcoin$Mes_Any[which.min(Bitcoin$Precio)],
+           y = min(Bitcoin$Precio), color = "red", size = 5) +
+   annotate(geom = "label", x = Bitcoin$Mes_Any[which.min(Bitcoin$Precio)],
+           y = min(Bitcoin$Precio), label =
format(Bitcoin$Mes_Any[which.min(Bitcoin$Precio)], "%Y-%m"),
+           color = "red", angle = 0, hjust = -0.1, vjust = -0.5, size = 5) +
+   labs(x = "Fecha", y = "Precio", title = "Gráfico de Precio de Bitcoin",
+        caption = "Fuente de datos: BTC.com") +
+   scale_x_date(date_labels = "%Y-%m", date_breaks = "1 year")+
+   theme(axis.title = element_text(size = 16),
+         axis.text = element_text(size = 14),
+         plot.title = element_text(size = 18, face = "bold"),
+         plot.caption = element_text(size = 14, hjust = 1))
> #Para garantizar el logro de la homocedasticidad (una suposición de prueba de modelo
> #requerida después de que construimos nuestro modelo), transforme el precio del Bitcoin
> #logarítmicamente.
> Bitcoin$Precio_Log = log(Bitcoin$Precio)
> #reordené las columnas
> Bitcoin <- Bitcoin %>% relocate("Precio_Log", .before = "BTC_Dificultad_Mineria")
> #Primer modelo para medir datos outliers
> mod <- lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad
+         + BTC_ATM_US + Fed_Tasa_US + M1_Oferta_Monetaria
+         + Tasa_Inflacion_Global, data = Bitcoin)
> #Creo la variable distancia de cook
> Bitcoin$cooks_d <- cooks.distance(mod)
> #definí valores atípicos basados con el criterio 4/n
> Bitcoin$outlier <- ifelse(Bitcoin$cooks_d < 4/nrow(Bitcoin), "keep","delete")
> #inspeccioné el conjunto de datos, el rango es 1:55 para todas las observaciones
> Bitcoin[1:55, ]
# A tibble: 55 × 11
Mes_Any      Precio Precio_Log BTC_Dificultad_Mineria BTC_Rentabilidad BTC_ATM_US
<date>      <dbl>   <dbl>         <dbl>             <dbl>         <dbl>
1 2017-06-01 2547.    7.84           0.71             1.21           766
2 2017-07-01 2719.    7.91           0.86             0.88           826
3 2017-08-01 4676.    8.45           0.89             1.62           853
4 2017-09-01 4208.    8.34           1.1              0.84           908
5 2017-10-01 6131.    8.72           1.45             1.11           948
6 2017-11-01 10543.   9.26           1.35             2.08           1052
7 2017-12-01 13062.   9.48           1.87             2.28           1125
8 2018-01-01 9914.    9.20           2.6              0.91           1216
9 2018-02-01 10903.   9.30           3.01             0.85           1319
10 2018-03-01 6970.    8.85           3.46             0.46           1475
#           45 more rows
#           5 more variables: Fed_Tasa_US <dbl>, M1_Oferta_Monetaria <dbl>,
#           Tasa_Inflacion_Global <dbl>, cooks_d <dbl>, outlier <chr>
#           Use `print(n = ...)` to see more rows
> df_Bitcoin_cooks_d <- Bitcoin[1:55, ]
> #tabla 3
> df_Bitcoin_cooks_d %>%
+   kbl() %>%
+   kable_styling() %>%
+   row_spec(which(df_Bitcoin_cooks_d$outlier == "delete"), bold = T,

```

```

+           color = "white", background = "red")
> # Crear la tabla formateada
> tabla_formateada <- kable(df_Bitcoin_cooksd) %>% kable_styling()
> # Guardar el contenido de la tabla en un archivo de texto
> cat(tabla_formateada, file = "tabla_formateada.txt")
> #Sumamos cuantos outlier hay
> outliers <- df_Bitcoin_cooksd$outlier == "delete"
> sum(outliers)
[1] 3
> #removí los outlier
> Bitcoin2 <- subset(Bitcoin, outlier != "delete")
> #confirmo que eliminé 3 datos
> nrow(Bitcoin2)
[1] 52
> #Eliminar columnas superfluas que resultaron de la fórmula de distancia de Cook
> Bitcoin3 <- Bitcoin2[c(-10,-11)]
> ##### TEST Y TRAIN #####
> #####
> #se crea la semilla para la muestra pseudoaleatoria
> set.seed(1)
> #usa el 70% de train y 30% test
> sample <- sample(c(TRUE, FALSE), nrow(Bitcoin3), replace=TRUE, prob=c(0.7,0.3))
> train <- Bitcoin3[sample, ]
> test  <- Bitcoin3[!sample, ]
> #Confirmo
> nrow(train)
[1] 35
> nrow(test)
[1] 17
> nrow(train)/nrow(Bitcoin3)
[1] 0.6730769
> nrow(test)/nrow(Bitcoin3)
[1] 0.3269231
> ##### MULTICOLINEALIDAD #####
> ##### VIF #####
> #Subset de variables independientes
> train2 <- train[c(-1,-2, -3)]
> par(mfrow=c(1,1))
> #ver matriz numerica
> matrix_datos <- cor(train2)
> #realizar la matriz de correlacion. Figura 1
> corrplot(matrix_datos , method = 'number')
> #calcular el VIF
> VIF<-vif(lm(lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad
+           + BTC_ATM_US + Fed_Tasa_US + M1_Oferta_Monetaria
+           + Tasa_Inflacion_Global, data = train)))
> #Tabla 1
> VIF %>%
+   kbl() %>%
+   kable_styling()
> # Crear la tabla formateada
> VIF <- kable(VIF) %>% kable_styling()
> # Guardar el contenido de la tabla en un archivo de texto
> cat(VIF, file = "tabla_VIF.txt")
> # Calcule el modelo de regresión lineal.
> model<-lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad

```

```

+           + BTC_ATM_US + Fed_Tasa_US + M1_Oferta_Monetaria
+           + Tasa_Inflacion_Global, data = train)
> summary(model)

```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad + BTC_ATM_US +
Fed_Tasa_US + M1_Oferta_Monetaria + Tasa_Inflacion_Global, data = train)
```

Residuals:

```

Min           1Q   Median           3Q      Max
-0.54677 -0.14387  0.03012  0.13608  0.42740

```

Coefficients:

```

Estimate Std. Error t value Pr(>|t|)
(Intercept)          4.714e+00  6.619e-01   7.122 9.49e-08 ***
  BTC_Dificultad_Mineria  7.520e-02  1.358e-02   5.538 6.38e-06 ***
  BTC_Rentabilidad       5.905e-01  1.512e-01   3.906  0.00054 ***
  BTC_ATM_US             6.591e-07  1.634e-06   0.403  0.68974
  Fed_Tasa_US            1.090e-01  1.052e-01   1.037  0.30864
  M1_Oferta_Monetaria    1.757e-05  1.795e-05   0.979  0.33619
  Tasa_Inflacion_Global  8.961e-01  1.993e-01   4.496  0.00011 ***
---

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2432 on 28 degrees of freedom
Multiple R-squared: 0.9256, Adjusted R-squared: 0.9096
F-statistic: 58.02 on 6 and 28 DF, p-value: 1.667e-14

```

> # Segundo modelo
> model<-lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad
+           + Fed_Tasa_US + M1_Oferta_Monetaria
+           + Tasa_Inflacion_Global, data = train)
> summary(model)

```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad +
Fed_Tasa_US + M1_Oferta_Monetaria + Tasa_Inflacion_Global,
data = train)
```

Residuals:

```

Min           1Q   Median           3Q      Max
-0.5430 -0.1556  0.0394  0.1362  0.4288

```

Coefficients:

```

Estimate Std. Error t value Pr(>|t|)
(Intercept)          4.531e+00  4.745e-01   9.549 1.86e-10 ***
  BTC_Dificultad_Mineria  7.566e-02  1.333e-02   5.674 3.91e-06 ***
  BTC_Rentabilidad       5.987e-01  1.476e-01   4.056  0.000344 ***
  Fed_Tasa_US            1.057e-01  1.033e-01   1.023  0.314770
  M1_Oferta_Monetaria    1.905e-05  1.732e-05   1.100  0.280297
  Tasa_Inflacion_Global  9.478e-01  1.504e-01   6.303 6.97e-07 ***
---

```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2396 on 29 degrees of freedom
 Multiple R-squared: 0.9251, Adjusted R-squared: 0.9122
 F-statistic: 71.66 on 5 and 29 DF, p-value: 2.007e-15

```
> #Tercer modelo
> model<-lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad
+           + M1_Oferta_Monetaria
+           + Tasa_Inflacion_Global, data = train)
> summary(model)
```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad + M1_Oferta_Monetaria +
Tasa_Inflacion_Global, data = train)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.51257	-0.13521	0.02273	0.14103	0.49358

Coefficients:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.598e+00	4.703e-01	9.777 7.71e-11 ***
BTC_Dificultad_Mineria	7.148e-02	1.270e-02	5.627 3.97e-06 ***
BTC_Rentabilidad	5.288e-01	1.310e-01	4.038 0.000344 ***
M1_Oferta_Monetaria	7.292e-06	1.296e-05	0.563 0.577871
Tasa_Inflacion_Global	1.012e+00	1.367e-01	7.406 2.98e-08 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2398 on 30 degrees of freedom
 Multiple R-squared: 0.9224, Adjusted R-squared: 0.9121
 F-statistic: 89.18 on 4 and 30 DF, p-value: 3.29e-16

```
> #Cuarto modelo
> model<-lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad
+           + Tasa_Inflacion_Global, data = train)
> summary(model)
```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad + Tasa_Inflacion_Global,
data = train)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.4913	-0.1487	0.0283	0.1471	0.4853

Coefficients:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.443317	0.376956	11.787 5.50e-13 ***
BTC_Dificultad_Mineria	0.076886	0.008222	9.351 1.55e-10 ***
BTC_Rentabilidad	0.557522	0.119266	4.675 5.45e-05 ***
Tasa_Inflacion_Global	1.051885	0.115775	9.086 3.01e-10 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2372 on 31 degrees of freedom
 Multiple R-squared: 0.9216, Adjusted R-squared: 0.914
 F-statistic: 121.5 on 3 and 31 DF, p-value: < 2.2e-16

```
> #Modelo definitivo
> model<-lm(Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad, data = train)
> summary(model)
```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad, data =
train)
```

Residuals:

```
Min          1Q  Median      3Q      Max
-0.62588 -0.42858  0.09613  0.35924  0.96722
```

Coefficients:

```
Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.69717    0.22158  34.738 < 2e-16 ***
BTC_Dificultad_Mineria  0.11588    0.01321   8.771 5.06e-10 ***
BTC_Rentabilidad       0.78244    0.21977   3.560 0.00118 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 0.4468 on 32 degrees of freedom
 Multiple R-squared: 0.7129, Adjusted R-squared: 0.6949
 F-statistic: 39.72 on 2 and 32 DF, p-value: 2.136e-09

```
> print(model)
```

Call:

```
lm(formula = Precio_Log ~ BTC_Dificultad_Mineria + BTC_Rentabilidad, data =
train)
```

Coefficients:

```
(Intercept) BTC_Dificultad_Mineria      BTC_Rentabilidad
7.6972      0.1159      0.7824
```

```
> ##### SUPUESTO MODELOS #####
```

```
> # Figura 3
> par(mfrow=c(1,1))
> plot(model)
```

```
<Enter> para ver el próximo gráfico:
> lmtest::bptest(model)
```

USUARIO
2023-09-01 11:29:45

En términos de ajuste y capacidad explicativa del modelo, el coeficiente de determinación R2 es de 0,7129

studentized Breusch-Pagan test data: model
 BP = 1.4359, df = 2, p-value = 0.4878

USUARIO
2023-09-01 11:27:59

"Dado que la hipótesis nula de la prueba de Breusch-Pagan es que hay heteroscedasticidad y el valor de p > 0.05 (p-value = 0.4878)"

```

> #Figura 4
> ols_plot_resid_hist(model)
> #Tabla 2
> ols_test_normality(model)

```

Test	Statistic	pvalue
Shapiro-Wilk	0.9379	0.0482
Kolmogorov-Smirnov	0.12	0.6507
Cramer-von Mises	4.0129	0.0000
Anderson-Darling	0.77	0.0411

```

> variables_restantes <- train[c(4,5)]
> par(mfrow=c(1,1))
> datamatrix <- cor(variables_restantes)
> #Figura 5
> corrplot(datamatrix, method = 'number')
> p <- predict(model, newdata = test)
> error <- test$Precio_Log - p
> RMSPE <- sqrt(mean(error^2))
> RMSPE
[1] 0.3822073
> c <- 7.69 + (0.12*23.17)+(0.78*0.19)
> #tomar la función de registro inverso, ya que transformamos los datos originales
> #logarítmicamente durante la fase de limpieza de datos
> exp(c)
[1] 40888.33
> media_muestra <- mean(train$Precio_Log)
> muestra_n <- length(train$Precio_Log)
> muestra_sd <- sd(train$Precio_Log)
> muestra_se <- muestra_sd/sqrt(muestra_n)
> alpha = 0.05
> grados_libertad = muestra_n - 1
> t.score = qt(p=alpha/2, df=grados_libertad,lower.tail=F)
> margen_error <- t.score * muestra_se
> int_bajo <- media_muestra - margen_error
> int_alto <- media_muestra+ margen_error
> print(c(exp(int_bajo),exp(int_alto)))
[1] 8197.598 14289.795

```