

BAB III

METODOLOGI PENELITIAN

3.1 Tempat dan Waktu Penelitian

Penelitian ini dilakukan di Startup XYZ, sebuah perusahaan fintech yang bergerak di bidang proprietary trading dengan fokus utama pada pasar aset kripto, khususnya Bitcoin. Startup ini mengelola akun trading yang beroperasi menggunakan modal dari berbagai proprietary trading firm internasional, serta menerapkan pendekatan berbasis data (data-driven) dalam pengambilan keputusan trading.

Salah satu aspek penting dalam strategi trading mereka adalah pemanfaatan analisis sentimen media sosial, terutama dari platform Reddit, untuk memperkirakan arah pasar dan merespons dinamika risiko secara real-time.

Penelitian ini berlangsung selama periode Februari 2025 hingga Juli 2025, yang mencakup tahapan:

- Perancangan model analisis
- Pengumpulan dan pengolahan data
- Analisis statistik serta interpretasi hasil

Adapun data historis yang dianalisis mencakup rentang waktu Januari 2019 sampai Desember 2024, dengan cakupan:

- Sentimen media sosial dari Reddit
- Volume perdagangan Bitcoin
- Harga harian Bitcoin
- Dominasi Bitcoin terhadap total kapitalisasi pasar kripto

- Indikator ekonomi global seperti Consumer Price Index (CPI) dan Federal Funds Rate (FFR)

3.2 Desain dan Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif eksplanatori dengan desain kausal-moderasi (causal moderation design). Tujuan utamanya adalah untuk menguji pengaruh sentimen media sosial (X1) dan volume perdagangan Bitcoin (X2) terhadap pergerakan harga Bitcoin (Y), serta menilai sejauh mana pengaruh tersebut dimoderasi oleh dominasi Bitcoin (Z1) dan faktor ekonomi global (Z2).

3.2.1 Variabel Penelitian

- X1: Sentimen Media Sosial (Reddit)
- X2: Volume Perdagangan Bitcoin
- Y: Pergerakan Harga Bitcoin (gabungan ATR dan $\Delta\%$ harga)
- Z1: Dominasi Bitcoin sebagai variabel moderasi
- Z2: Faktor Ekonomi Global (CPI dan Fed Funds Rate) sebagai variabel moderasi

3.2.2 Jenis dan Sumber Data

Penelitian ini menggunakan data sekunder time series harian yang dikumpulkan dari sumber-sumber berikut:

- Reddit: Data sentimen dan volume posting diperoleh melalui situs [academictorrents](https://www.reddit.com).
- TradingView: Seluruh data pasar kripto, termasuk harga harian, volume perdagangan spot, dan dominasi Bitcoin.

- FRED, IMF, Yahoo Finance: Data makroekonomi seperti CPI dan Federal Funds Rate.

3.2.3 Teknik Analisis Data

Analisis dilakukan melalui tahapan berikut:

- Statistik deskriptif
- Uji asumsi klasik (normalitas, multikolinearitas, autokorelasi, heteroskedastisitas)
- Regresi Linier Berganda (OLS)
- Moderated Regression Analysis (MRA), dengan interaksi:
 - $X1 \times Z1$
 - $X1 \times Z2$

3.2.4 Teknik Pengolahan Sentimen

- Sentimen dianalisis menggunakan teknik Natural Language Processing (NLP) berbasis lexicon seperti VADER dan TextBlob
- Postingan Reddit yang mengandung kata kunci seperti “Bitcoin”, “BTC” diklasifikasikan ke dalam tiga kategori: positif, netral, atau negatif dalam bentuk skala angka.
- Output dari analisis ini berupa:
 - Skor sentimen harian (skala 0–1)
 - Volume postingan harian

Keduanya kemudian dinormalisasi menjadi z-score

3.2.5 Alasan Pemilihan Pendekatan

Pendekatan kausal-moderasi dipilih karena mampu menjelaskan:

- Bagaimana sentimen dan aktivitas pasar mempengaruhi harga Bitcoin secara statistik
- Bagaimana pengaruh tersebut dapat diperkuat atau dilemahkan oleh struktur dominasi pasar BTC (Z1) dan kondisi makro ekonomi (Z2)

3.3 Unit Analisis

Unit analisis dalam penelitian ini merujuk pada entitas data harian yang dianalisis untuk menguji hubungan antara variabel-variabel penelitian. Penelitian ini difokuskan khusus pada Bitcoin, tanpa mencakup aset kripto lain, guna menjaga konsistensi dan fokus analisis terhadap satu instrumen yang dominan di pasar kripto.

Unit analisis meliputi:

- Y – Pergerakan Harga Bitcoin: Representasi volatilitas harga harian, yang diukur melalui kombinasi dua indikator, Average True Range (ATR) dan persentase perubahan harga harian ($\Delta\%$).
- X1 – Sentimen Media Sosial: Skor sentimen harian dan volume posting yang diperoleh dari platform Reddit, berdasarkan postingan yang mengandung kata kunci “Bitcoin” dan “BTC”.
- X2 – Volume Perdagangan Bitcoin: Nilai dan jumlah transaksi Bitcoin harian yang diperoleh dari TradingView.
- Z1 – Dominasi Bitcoin: Rasio dominasi pasar Bitcoin terhadap total kapitalisasi pasar kripto per hari.
- Z2 – Faktor Ekonomi Global: Indikator ekonomi makro yang terdiri dari Consumer Price Index (CPI) dan Federal Funds Rate (suku bunga acuan The Fed).

3.3.1 Populasi

Populasi dalam penelitian ini mencakup seluruh data harian yang tersedia selama periode 1 Januari 2019 hingga 31 Desember 2024, meliputi:

- Data pasar kripto: Harga harian, volume transaksi, dan dominasi Bitcoin, semuanya diperoleh dari platform TradingView.
- Data sentimen Reddit: Semua posting yang mengandung kata kunci “Bitcoin” atau “BTC”, dikumpulkan melalui situs academictorrents.
- Data ekonomi makro: Nilai CPI dan Federal Funds Rate, diperoleh dari sumber seperti FRED dan IMF.

3.3.2 Sampel

Penelitian ini tidak menggunakan teknik sampling karena seluruh data harian selama 6 tahun penuh (2019–2024) digunakan secara keseluruhan. Dengan demikian, populasi = sampel.

Rincian cakupan datanya sebagai berikut:

- Y (Pergerakan Harga Bitcoin): Data harian close price digunakan untuk menghitung ATR dan $\Delta\%$ harga, yang kemudian dikonversi ke dalam bentuk z-score.
- X1 (Sentimen Media Sosial): Postingan harian dari Reddit yang mengandung kata “Bitcoin” dan “BTC” dianalisis menggunakan teknik NLP (VADER/TextBlob), untuk menghasilkan skor sentimen dan volume harian, lalu dinormalisasi.
- X2 (Volume Perdagangan Bitcoin): Volume spot harian Bitcoin, baik dalam bentuk jumlah transaksi maupun nilai USD, diperoleh dari TradingView.
- Z1 (Dominasi Bitcoin): Proporsi kapitalisasi pasar Bitcoin terhadap total pasar kripto (% dominasi harian).

- Z2 (Faktor Ekonomi Global): Data bulanan CPI dan Fed Funds Rate disesuaikan menjadi data harian melalui interpolasi linier agar kompatibel dengan struktur data time series lainnya.

3.4 Teknik Pengumpulan Data

Penelitian ini menggunakan data sekunder yang dikumpulkan secara historis dari berbagai platform digital terpercaya untuk periode 1 Januari 2019 hingga 31 Desember 2024. Semua data diproses dalam format time series harian untuk mendukung analisis regresi dan uji moderasi.

3.4.1 Data Harga dan Volume Perdagangan Bitcoin

- Sumber: TradingView
- Jenis Data:
 - Harga penutupan harian Bitcoin (daily close)
 - ATR (Average True Range)
 - Persentase perubahan harga harian ($\Delta\%$)
 - Volume spot harian (jumlah transaksi dan nilai USD)
- Format: CSV, kemudian diolah dalam Python dan Excel menjadi z-score

3.4.2 Data Sentimen Media Sosial

- Sumber: Reddit (via scraping academictorrents)
- Jenis Data:
 - Postingan harian yang mengandung kata kunci “Bitcoin” dan “BTC”
 - Teks posting diklasifikasikan sebagai positif, netral, atau negatif

- Metode Pengolahan:
 - Natural Language Processing (NLP) dengan VADER dan TextBlob
 - Output: skor sentimen harian (0–1) dan volume posting, keduanya dinormalisasi ke dalam z-score

3.4.3 Data Dominasi Bitcoin

- Sumber: TradingView
- Jenis Data:
 - Proporsi kapitalisasi pasar Bitcoin terhadap total pasar kripto (BTC Dominance %)
- Frekuensi: Harian

3.4.4 Data Faktor Ekonomi Global

- Sumber: FRED, IMF, dan Yahoo Finance
- Jenis Data:
 - Consumer Price Index (CPI) – dalam persentase tahunan (YoY)
 - Federal Funds Rate – suku bunga acuan The Fed (%)
- Frekuensi Asli: Bulanan
- Metode Penyesuaian:
 - Diinterpolasi secara linier menjadi data harian untuk keperluan regresi time series

3.5 Definisi Operasional Variabel

Pada bagian ini, setiap variabel yang terlibat dalam penelitian ini dijelaskan secara rinci agar dapat diukur dan dianalisis dengan jelas. Definisi operasional diperlukan untuk memberikan pemahaman yang konsisten mengenai bagaimana setiap variabel akan diukur dalam penelitian ini. Variabel yang terlibat dibagi menjadi tiga kategori utama: variabel bebas, variabel terikat, dan variabel moderasi.

Tabel 3.1 Definisi Operasional Variabel

No	Variabel	Definisi Operasional
1	Sentimen Media Sosial (X1)	Persepsi pasar terhadap Bitcoin berdasarkan aktivitas dan isi diskusi di platform media sosial. Pada penelitian ini, media sosial yang diteliti adalah platform Reddit. Skor sentimen diklasifikasikan ke dalam kategori positif, negatif, atau netral menggunakan teknik NLP (VADER/TextBlob), dengan rentang awal -1.00 - 1.00 yang kemudian dinormalisasi ke z-score. Volume sentimen dihitung dari jumlah posting harian yang memuat kata kunci seperti 'Bitcoin', 'BTC'.
2	Volume Perdagangan (X2)	Aktivitas perdagangan Bitcoin berdasarkan dua indikator yaitu total transaksi harian dan total volume perdagangan harian. Pada penelitian ini, menghitung total transaksi bitcoin harian dan total volume perdagangan bitcoin harian menggunakan USD. Sumber data diambil dari TradingView. Keduanya dinormalisasi ke z-score untuk dianalisis dalam model regresi.
3	Pergerakan Harga Bitcoin (Y)	Dinamika harga Bitcoin pada jangka waktu tertentu. Pada penelitian ini menggunakan dua pendekatan: (1) volatilitas harian menggunakan indikator ATR, dan (2) persentase perubahan harga harian. Kedua indikator diambil dari TradingView dan dinormalisasi ke dalam bentuk z-score.
4	Dominasi Bitcoin (Z1)	Proporsi kapitalisasi pasar Bitcoin terhadap total pasar kripto secara keseluruhan. Dinyatakan dalam persentase dan diperoleh dari TradingView. Rasio dominasi harian ini digunakan sebagai variabel moderasi dalam model interaksi.
5	Faktor Ekonomi Global (Z2)	Tekanan makroekonomi global yang berpotensi memoderasi pengaruh sentimen media sosial terhadap harga Bitcoin. Indikator yang digunakan meliputi CPI YoY (%) dan Fed Funds Rate (%), yang diambil dari FRED, IMF, dan Yahoo Finance. Data bulanan diinterpolasi ke harian dan dinormalisasi menjadi z-score.

Untuk memudahkan memahami variabel dan dimensi/indikator berikut sumber datanya, berikut adalah tabel dimensi variabel:

Tabel 3.2 Dimensi & Indikator variabel

Nama Variabel		Indikator Utama	Sumber Data	Teknik Pengukuran
X1	Sentimen Media Sosial	- Skor Sentimen (-1 – 1) dari Reddit - Volume Sentimen (jumlah post harian)	Reddit via academictorrents	Analisis NLP (VADER/TextBlob), scraping keyword, normalisasi z-score
X2	Volume Perdagangan	- Jml transaksi BTC harian - Persentase kenaikan/penurunan volume (daily delta)	TradingView	Data jumlah transaksi dan volume dari TradingView, normalisasi z-score
Y	Pergerakan Harga Kripto	- Range pergerakan harian (ATR) - Persentase perubahan harga harian	TradingView	ATR dan Δ Price dihitung harian, lalu dinormalisasi ke z-score
Z1	Dominasi Bitcoin	- Dominance Index = Market Cap BTC / Market Cap Total	TradingView	Rasio dominasi harian, dikonversi ke z-score
Z2	Faktor Ekonomi Global	- CPI YoY (%) - Federal Funds Rate (%)	FRED, IMF, Yahoo Finance	Data bulanan diinterpolasi ke harian, lalu distandarisasi ke z-score

3.5.1 Variabel Bebas (X)

3.5.1.1 X1: Sentimen Media Sosial (Reddit)

Variabel ini mengukur sentimen yang berkembang di kalangan komunitas kripto melalui platform Reddit, yang dikenal sebagai salah satu kanal diskusi terbesar untuk topik kripto global.

Sentimen dihitung melalui dua komponen:

- **X1.1: Reddit Sentiment Score**
Skor sentimen harian dihitung menggunakan pendekatan Natural Language Processing (NLP) berbasis leksikal (misalnya VADER, TextBlob) dan transformasi skor menjadi skala standar (z-score). Analisis dilakukan terhadap konten post yang mengandung kata kunci terkait Bitcoin seperti “BTC”, “Bitcoin”.
- **X1.2: Reddit Volume**
Mengukur jumlah postingan harian di Reddit yang relevan dengan Bitcoin. Volume ini dianggap sebagai representasi intensitas diskusi publik. Volume juga ditransformasi dalam bentuk z-score untuk keperluan regresi.

Dari dua indikator ini, dibentuk satu variabel gabungan X1_combined melalui metode rata-rata tertimbang atau agregasi statistik untuk merepresentasikan sentimen media sosial secara keseluruhan dalam satu nilai harian.

3.5.1.2 X2: Volume Perdagangan Bitcoin

Variabel ini mengukur aktivitas perdagangan harian Bitcoin sebagai proksi dari likuiditas dan partisipasi pasar.

Indikator yang digunakan:

- **X2.1: Jumlah Transaksi Bitcoin Harian**
Diambil dari berbagai sumber terutama platform Tradingview, mewakili seberapa aktif Bitcoin transaksi harian BTC dilakukan.

- X2.2: Total Volume USD Perdagangan Bitcoin Harian
Diambil dari kombinasi data bursa yang ada di Tradingview, kemudian dikonversi ke bentuk z-score.

Variabel X2_combined dibentuk dari agregasi dua indikator tersebut dan digunakan sebagai representasi umum dari volume aktivitas pasar Bitcoin.

3.5.2 Variabel Terikat (Y)

3.5.2.1 Y: Pergerakan Harga Bitcoin

Variabel dependen dalam penelitian ini adalah Y_combined, yaitu indikator gabungan yang merepresentasikan pergerakan harga Bitcoin. Variabel ini disusun dari dua dimensi teknikal yang saling melengkapi:

- Y.1: ATR (Average True Range) z-score
Mewakili volatilitas harga Bitcoin, dihitung menggunakan indikator ATR yang telah ditransformasi ke dalam z-score untuk menjaga skala yang konsisten dengan variabel lain.
- Y.2: Persentase Perubahan Harga Harian z-score
Mengukur return harian Bitcoin dalam bentuk persentase perubahan, kemudian dinormalisasi ke bentuk z-score.

Kedua indikator ini digabung menjadi variabel Y_combined yang digunakan dalam model regresi untuk merepresentasikan intensitas pergerakan harga Bitcoin secara teknikal dan statistik.

3.5.3 Variabel Moderasi (Z)

3.5.3.1 Z1: Dominasi Bitcoin (BTC Dominance)

Mengukur proporsi kapitalisasi pasar Bitcoin terhadap total kapitalisasi pasar aset kripto secara keseluruhan. Variabel ini digunakan sebagai variabel moderasi untuk mengevaluasi apakah dominasi pasar oleh Bitcoin memperkuat atau melemahkan pengaruh sentimen media sosial terhadap harga.

Data BTC Dominance diubah ke dalam bentuk z-score harian (Z1_btc_dominance_zscore) agar kompatibel dalam model regresi.

3.5.3.2 Z2: Faktor Ekonomi Global

Variabel ini mencerminkan tekanan dan kondisi makro ekonomi global yang didasari atas rilis data ekonomi di US yang mungkin mempengaruhi pasar aset digital seperti Bitcoin. Dua indikator utama digunakan:

- Z2.1: US Consumer Price Index (CPI) Year-on-Year (%)
Mengukur tingkat inflasi tahunan di AS berdasarkan data dari FRED atau IMF. Nilai ini menunjukkan daya beli dan ekspektasi kebijakan moneter.
- Z2.2: US Federal Funds Rate (%)
Merupakan suku bunga acuan Bank Sentral AS (The Fed), yang mempengaruhi likuiditas dan preferensi investor terhadap aset berisiko.

Kedua data ini diselaraskan ke dalam frekuensi harian menggunakan interpolasi linier, lalu digabung menjadi satu variabel Z2_combined_zscore untuk kepentingan analisis regresi moderasi.

3.6 Teknik Analisis Data

Penelitian ini menggunakan pendekatan kuantitatif dengan teknik analisis data yang bertujuan untuk menguji pengaruh variabel bebas (X1 dan X2) terhadap variabel terikat (Y) serta peran variabel moderasi (Z1 dan Z2) dalam memperkuat atau memperlemah hubungan antara X1 dan Y.

Tahapan analisis data mencakup:

- Analisis deskriptif statistik, untuk menggambarkan distribusi, rata-rata, dan penyebaran masing-masing variabel.
- Uji kualitas data, meliputi uji validitas dan reliabilitas terhadap indikator pembentuk variabel gabungan.

- Uji asumsi klasik, termasuk uji normalitas, multikolinearitas, autokorelasi, dan heteroskedastisitas untuk memastikan validitas model regresi.
- Analisis regresi linear berganda, digunakan untuk mengukur pengaruh variabel bebas dan interaksi moderasi terhadap variabel terikat.
- Uji interaksi (moderasi) melalui pembentukan variabel interaksi $X1 \times Z1$ dan $X1 \times Z2$ untuk mengevaluasi konteks penguatan atau pelemahan pengaruh sentimen sosial media terhadap harga Bitcoin.

3.6.1 Skala Pengukuran dan Angka Penafsiran

Seluruh variabel dalam penelitian ini diukur dalam bentuk numerik berbasis skala interval atau rasio, dan selanjutnya dinormalisasi ke dalam bentuk z-score. Pendekatan ini dipilih untuk memastikan konsistensi skala antar variabel dan kompatibilitas dalam model regresi linier. Penjabaran masing-masing variabel dijelaskan sebagai berikut:

3.6.1.1 X1 Sentimen Media Sosial (Reddit)

Variabel ini mengukur persepsi publik terhadap Bitcoin berdasarkan konten harian di platform Reddit, yang merupakan salah satu pusat diskusi kripto terbesar dan paling aktif secara global. Dua indikator utama digunakan:

- X1.1 Reddit Sentiment Score

Skor sentimen diperoleh dari hasil analisis Natural Language Processing (NLP) terhadap isi postingan Reddit yang mengandung kata kunci seperti “Bitcoin”, “BTC”. Model NLP yang digunakan mencakup pendekatan leksikal yaitu VADE. Skor awal berada pada rentang -1 hingga 1, dengan interpretasi sebagai berikut:

Rentang Skor	Interpretasi
-1.00 – -0.60	Sangat Negatif
-0.59 – -0.21	Negatif
-0.20 – +0.20	Netral
+0.21 – +0.59	Positif
+0.60 – +1.00	Sangat Positif

Untuk keperluan model regresi, skor ini dinormalisasi ke dalam bentuk z-score (X1.1_zscore).

- **X1.2 Reddit Volume**
Merupakan jumlah postingan harian Reddit yang relevan dengan Bitcoin. Volume ini mencerminkan intensitas perbincangan publik dan diinterpretasikan sebagai ukuran popularitas atau tekanan sentimen. Nilainya dinormalisasi menjadi z-score (X1.2_zscore).
- **X1_combined**
Merupakan agregasi dari kedua indikator di atas, dihitung melalui metode rata-rata atau agregasi statistik, dan digunakan sebagai variabel bebas utama dalam model regresi.

3.6.1.2 X2 Volume Perdagangan Bitcoin

Variabel ini digunakan untuk menggambarkan tingkat aktivitas dan likuiditas pasar Bitcoin, yang berpotensi mempengaruhi harga aset tersebut. Terdiri dari dua indikator:

- **X2.1 Total Transaksi Bitcoin Harian**
Data diambil dari platform TradingView, yang menyediakan total transaksi harian Bitcoin pada pasar spot. Nilai ini mewakili volume aktivitas dalam satuan unit transaksi per hari dan diubah ke dalam z-score (X2.1_zscore).

- **X2.2 Volume Perdagangan Harian BTC dalam USD**
Diambil dari platform seperti TradingView, indikator ini menunjukkan nilai dolar AS yang berpindah setiap hari dalam perdagangan Bitcoin. Data ini juga dikonversi ke z-score (X2.2_zscore).
- **X2_combined**
Gabungan dari kedua indikator volume di atas, digunakan sebagai representasi terpadu aktivitas pasar Bitcoin dalam model regresi.

3.6.1.3 Y Pergerakan Harga Bitcoin

Sebagai variabel dependen dalam penelitian ini, Y_combined dirancang untuk merepresentasikan intensitas pergerakan harga Bitcoin, dengan mempertimbangkan baik arah maupun volatilitasnya. Terdiri dari dua indikator utama:

- **Y.1 ATR (Average True Range) z-score**
Mengukur volatilitas harga harian Bitcoin menggunakan indikator teknikal ATR, yang menunjukkan seberapa besar pergerakan harga harian rata-rata. Nilai ini dinormalisasi ke dalam z-score (Y.1_zscore) agar dapat digabung dengan indikator lainnya.
- **Y.2 – Persentase Perubahan Harga Harian z-score**
Mengukur return harian harga Bitcoin dengan rumus:

$$\text{Return}_t = \frac{P_t - P_{t-1}}{P_{t-1}} \times 100$$

Nilai return ini kemudian dinormalisasi ke dalam bentuk z-score (Y.2_zscore).

- **Y_combined**
Gabungan dari Y.1 dan Y.2 yang memberikan gambaran menyeluruh terhadap dinamika pergerakan harga Bitcoin, baik dari sisi fluktuasi maupun arah perubahan.

3.6.1.4 Z1 Dominasi Bitcoin

Variabel moderasi pertama adalah Dominasi Bitcoin, yang mengukur proporsi kapitalisasi pasar Bitcoin terhadap total pasar kripto secara global. Nilai ini diperoleh dari TradingView dan dinormalisasi sebagai z-score ($Z1_btc_dominance_zscore$).

Z1 digunakan sebagai variabel interaksi untuk mengetahui apakah kekuatan dominasi Bitcoin memperkuat atau justru melemahkan hubungan antara sentimen sosial media dan pergerakan harga Bitcoin.

3.6.1.5 Z2 Faktor Ekonomi Global

Variabel moderasi kedua terdiri dari dua indikator utama makro ekonomi, yang dipilih karena memiliki pengaruh besar terhadap pasar keuangan secara global, termasuk aset digital:

- **Z2.1 – CPI Year-on-Year (%)**
Mengukur tingkat inflasi tahunan di Amerika Serikat. Data diambil dari sumber seperti FRED (Federal Reserve Economic Data) dan dinormalisasi.
- **Z2.2 – Fed Funds Rate (%)**
Merupakan suku bunga acuan The Federal Reserve, yang menjadi faktor kunci dalam keputusan investor terhadap aset berisiko.

Kedua indikator tersebut diselaraskan secara temporal melalui interpolasi linier ke skala harian, kemudian digabung dan dinormalisasi sebagai $Z2_combined_zscore$, yang digunakan sebagai variabel moderasi dalam model interaksi.

3.6.2 Persamaan Regresi Linear Berganda

Regresi linear berganda digunakan untuk menguji pengaruh langsung dari variabel $X1$ (Sentimen Media Sosial) dan $X2$ (Volume Perdagangan) terhadap Y (Pergerakan Harga Kripto), serta untuk menganalisis peran moderasi dari $Z1$ (Dominasi Bitcoin) dan $Z2$ (Faktor Ekonomi Global).

Regresi linear berganda juga digunakan untuk melihat hubungan antara semua variabel. Persamaan regresi yang digunakan adalah sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 Z_1 + \beta_4 Z_2 + \varepsilon$$

Dimana:

- **Y** = Pergerakan harga Bitcoin (log return atau % return harian)
- **X1** = Skor sentimen media sosial (Reddit)
- **X2** = Volume perdagangan harian Bitcoin
- **Z1** = Dominasi Bitcoin (% market cap)
- **Z2** = Faktor ekonomi global (CPI, Fed Funds Rate)
- β = Koefisien regresi
- ε = Error

3.6.3 Uji Kualitas Data

Untuk memastikan kualitas data yang digunakan dalam penelitian ini, beberapa uji kualitas data perlu dilakukan. Kualitas data yang baik sangat penting untuk memastikan bahwa hasil penelitian dapat dipercaya dan valid. Uji kualitas data yang dilakukan mencakup validitas dan reliabilitas, yang bertujuan untuk memastikan bahwa instrumen yang digunakan dalam mengukur variabel, terutama sentimen media sosial, dapat memberikan hasil yang sahih dan konsisten.

3.6.3.1 Validitas

Uji validitas digunakan untuk memastikan bahwa instrumen pengukuran yang digunakan dalam penelitian ini benar-benar mengukur apa yang dimaksudkan untuk diukur. Validitas sangat penting untuk memastikan bahwa data yang dikumpulkan relevan dan sesuai dengan tujuan penelitian. Terdapat beberapa jenis validitas yang akan diuji dalam penelitian ini:

- **Validitas Konstruk:** Validitas konstruk mengacu pada sejauh mana instrumen pengukuran (dalam hal ini, alat yang digunakan untuk mengukur sentimen media sosial) dapat mengukur konsep yang dimaksud. Sebagai contoh, dalam penelitian ini, jika instrumen digunakan untuk mengukur sentimen media sosial, validitas konstruk akan memastikan bahwa instrumen tersebut benar-benar mengukur sentimen (positif, negatif, atau netral) dan bukan faktor lain yang tidak relevan, seperti jumlah posting atau frekuensi kata. Validitas konstruk dapat diuji melalui analisis faktor yang mengidentifikasi hubungan antara indikator-indikator yang diukur dan konstruk yang ingin dipelajari.
- **Validitas Isi:** Validitas isi berkaitan dengan sejauh mana instrumen pengukuran mencakup seluruh aspek dari konstruk yang diukur. Dalam hal ini, validitas isi akan menguji apakah instrumen untuk mengukur sentimen media sosial mencakup berbagai aspek atau dimensi sentimen yang relevan, seperti emosi yang muncul dalam postingan di media sosial (misalnya, kegembiraan, kemarahan, kebingungan), serta bagaimana faktor-faktor lain seperti konteks berita dan pengaruh media massa mempengaruhi persepsi pasar. Validitas isi bisa diuji dengan meminta pendapat para ahli atau panel untuk menilai kelengkapan dan relevansi item-item dalam instrumen pengukuran.

3.6.3.2 Reliabilitas

Uji reliabilitas digunakan untuk memastikan bahwa instrumen pengukuran yang digunakan dalam penelitian ini memberikan hasil yang konsisten ketika diulang. Reliabilitas mengukur konsistensi dan ketepatan hasil yang diperoleh dari instrumen pengukuran, serta sejauh mana instrumen tersebut memberikan hasil yang sama dalam kondisi yang serupa.

- **Cronbach's Alpha:** Uji reliabilitas akan dilakukan menggunakan Cronbach's alpha, yang digunakan untuk mengukur konsistensi internal dari instrumen pengukuran. Dalam konteks penelitian ini, Cronbach's alpha akan mengukur konsistensi item-item dalam instrumen yang digunakan untuk mengukur

sentimen media sosial. Nilai Cronbach's alpha berkisar antara 0 hingga 1, dengan nilai yang lebih tinggi menunjukkan konsistensi yang lebih baik. Umumnya, nilai Cronbach's alpha diatas 0,6 dianggap baik, menunjukkan bahwa instrumen memiliki reliabilitas yang cukup untuk digunakan dalam penelitian.

Jika instrumen untuk mengukur sentimen media sosial memiliki nilai Cronbach's alpha yang rendah, ini menunjukkan bahwa instrumen tersebut mungkin tidak konsisten dalam mengukur sentimen, sehingga perlu dilakukan perbaikan dalam alat pengukurannya.

3.6.3.3 Prosedur Pengujian

- Langkah pertama: Mengumpulkan data dari sumber media sosial yang relevan, seperti Reddit untuk menangkap sentimen pasar yang sedang berlangsung.
- Langkah kedua: Menggunakan instrumen pengukuran yang telah disusun untuk menilai sentimen dari data media sosial. Proses ini bisa mencakup teknik analisis teks, seperti Natural Language Processing (NLP), untuk mengidentifikasi emosi dan sentimen yang terkandung dalam postingan dan komentar.
- Langkah ketiga: Melakukan uji validitas konstruk dan validitas isi untuk memastikan bahwa instrumen dapat mengukur sentimen dengan tepat dan mencakup seluruh aspek yang relevan.
- Langkah keempat: Melakukan uji reliabilitas dengan menghitung Cronbach's alpha untuk memastikan bahwa instrumen memberikan hasil yang konsisten. Jika nilai Cronbach's alpha rendah, instrumen perlu diperbaiki atau diuji ulang.

3.6.4 Uji Asumsi Klasik

Sebelum melakukan analisis regresi, penting untuk memastikan bahwa model regresi yang digunakan memenuhi asumsi klasik yang diperlukan agar hasil analisis dapat dipercaya dan valid. Uji asumsi klasik ini bertujuan untuk mengevaluasi kualitas

dan kecocokan data yang digunakan dalam model regresi, sehingga hasil yang diperoleh dapat memberikan kesimpulan yang akurat.

3.6.4.1 Uji Normalitas

Uji normalitas digunakan untuk menguji apakah data yang digunakan dalam analisis regresi mengikuti distribusi normal. Asumsi normalitas penting untuk memastikan bahwa hasil uji statistik regresi valid, terutama untuk koefisien regresi yang diestimasi.

- Tujuan: Menguji apakah residual (error) dari model regresi terdistribusi secara normal.
- Metode: Uji normalitas dapat dilakukan dengan menggunakan beberapa metode, di antaranya:
 - Uji Kolmogorov-Smirnov: Uji ini menguji apakah sampel data berasal dari distribusi normal dengan membandingkan distribusi kumulatif sampel dengan distribusi normal yang diharapkan.
 - Uji Shapiro-Wilk: Uji ini lebih sensitif dibandingkan Kolmogorov-Smirnov dan sering digunakan untuk ukuran sampel kecil hingga sedang. Uji ini menguji hipotesis nol bahwa data berdistribusi normal.

Interpretasi:

- Jika nilai p dari uji normalitas lebih besar dari tingkat signifikansi (biasanya 0,05), maka data tidak berbeda signifikan dari distribusi normal, dan asumsi normalitas dapat diterima.
- Jika nilai p lebih kecil dari 0,05, maka data dianggap tidak berdistribusi normal, dan perlu dilakukan transformasi atau pemilihan metode lain.

Meskipun asumsi normalitas residual merupakan syarat penting dalam regresi linier klasik, dalam konteks data berukuran besar, pelanggaran terhadap asumsi ini tidak selalu menyebabkan estimasi menjadi tidak valid. Hal ini didasarkan pada prinsip *Central Limit Theorem (CLT)*, yang menyatakan bahwa dalam sampel besar, distribusi dari rata-rata akan mendekati normal, terlepas dari distribusi populasi aslinya. Oleh karena itu, selama ukuran sampel cukup besar, hasil estimasi koefisien regresi melalui metode OLS tetap dapat diandalkan dan pengujian hipotesis tetap valid (Gujarati & Porter, 2009; Wooldridge, 2016; Field, 2013; Tabachnick & Fidell, 2013).

Selain itu, jika ditemukan indikasi kuat ketidaksesuaian terhadap normalitas, pendekatan seperti robust standard error juga dapat digunakan untuk mengatasi potensi bias dalam estimasi varians (Lumley & Sheather, 2002).

3.6.4.2 Uji Multikolinieritas

Uji multikolinieritas dilakukan untuk menguji apakah terdapat korelasi yang tinggi antara variabel independen dalam model regresi. Ketika terdapat multikolinieritas, model regresi tidak dapat menghasilkan estimasi koefisien yang stabil, yang akan mengganggu interpretasi hubungan antar variabel.

- Tujuan: Untuk mengidentifikasi korelasi yang tinggi antar variabel independen yang dapat menyebabkan gangguan pada estimasi parameter regresi.
- Metode: Uji ini dilakukan dengan menghitung Variance Inflation Factor (VIF). VIF mengukur seberapa besar varians dari koefisien regresi meningkat akibat adanya korelasi antar variabel independen.
 - Formula: VIF untuk suatu variabel dihitung dengan cara menghitung R^2 dari regresi variabel tersebut terhadap semua variabel independen lainnya.
 - Interpretasi:

- Jika VIF lebih besar dari 10, ini menunjukkan adanya multikolinieritas yang tinggi antara variabel independen tersebut.
- Sebagai solusinya, dapat dilakukan penghapusan salah satu variabel yang sangat berkorelasi atau menggunakan teknik seperti analisis faktor untuk mengurangi multikolinieritas.

3.6.4.3 Uji Heteroskedastisitas

Uji heteroskedastisitas bertujuan untuk mengetahui apakah residual model regresi memiliki varians yang konstan (homoskedastisitas). Jika terjadi heteroskedastisitas (varian residual tidak konstan), maka hasil estimasi OLS tetap tidak bias, namun menjadi tidak efisien dan uji statistik (t dan F) menjadi tidak valid.

Metode Pengujian

Penelitian ini menggunakan Uji Glejser, yaitu dengan meregresikan nilai absolut residual terhadap variabel independen. Jika koefisien pada model Glejser signifikan secara statistik ($p\text{-value} < 0,05$), maka dapat disimpulkan adanya gejala heteroskedastisitas.

Interpretasi Hasil

- Jika $p\text{-value} > 0,05$, tidak ada indikasi heteroskedastisitas.
- Jika $p\text{-value} < 0,05$, terdapat heteroskedastisitas, sehingga disarankan untuk menggunakan robust standard error agar hasil estimasi tetap valid.

Tindakan Perbaikan

Jika ditemukan heteroskedastisitas, maka model akan disesuaikan dengan penggunaan robust standard errors (standard error yang tahan terhadap heteroskedastisitas) saat melakukan uji signifikansi koefisien.

3.6.4.4 Uji Autokorelasi

Uji autokorelasi digunakan untuk menguji apakah terdapat ketergantungan antara error term pada periode waktu yang berbeda dalam model regresi. Ketika terdapat autokorelasi, nilai error pada satu periode waktu dipengaruhi oleh nilai error pada periode waktu sebelumnya. Ini bisa menyebabkan estimasi koefisien yang bias dan kesalahan dalam pengujian hipotesis.

- Tujuan: Menguji apakah error residual dari model regresi pada satu periode waktu berhubungan dengan error pada periode waktu lainnya.
- Metode: Uji ini dilakukan dengan menggunakan Uji Durbin-Watson.
 - Uji Durbin-Watson menguji apakah residual memiliki korelasi dengan residual pada periode waktu sebelumnya.

Interpretasi:

- Nilai Durbin-Watson berkisar antara 0 hingga 4. Nilai sekitar 2 menunjukkan tidak adanya autokorelasi.
- Jika nilai Durbin-Watson lebih kecil dari 1 atau lebih besar dari 3, maka ada indikasi kuat adanya autokorelasi.
- Jika autokorelasi terdeteksi, perlu dilakukan perbaikan dengan model yang mempertimbangkan autokorelasi, seperti model ARIMA (Autoregressive Integrated Moving Average) atau menggunakan transformasi data.

3.6.5 Uji Hipotesis

Setelah model regresi linear berganda dibentuk berdasarkan data yang telah terkumpul, dilakukan pengujian hipotesis untuk mengetahui hubungan dan pengaruh antar variabel bebas, moderasi, dan variabel terikat dalam penelitian ini. Tujuannya adalah untuk menguji apakah Sentimen Media Sosial (X1) dan Volume Perdagangan (X2) memiliki pengaruh signifikan terhadap Pergerakan Harga Kripto (Y), dengan

mempertimbangkan peran moderasi dari Dominasi Bitcoin (Z1) dan Faktor Ekonomi Global (Z2).

Beberapa teknik statistik yang digunakan dalam pengujian hipotesis meliputi:

3.6.5.1 Uji t (Parsial)

Uji t digunakan untuk menguji pengaruh masing-masing variabel independen dan moderasi terhadap variabel dependen secara individual atau parsial.

Contoh pengujian:

- Apakah X1 (Sentimen Media Sosial) berpengaruh signifikan terhadap Y (Pergerakan Harga Kripto)?
- Apakah Z2 (Faktor Ekonomi Global) berperan dalam memoderasi hubungan antara X1 dan Y?

Interpretasi dilakukan dengan membandingkan nilai t-hitung dengan t-tabel dan melihat nilai signifikansi ($p\text{-value} < 0,05$).

3.6.5.2 Uji F (Simultan)

Uji F digunakan untuk melihat apakah seluruh variabel independen dan variabel moderasi secara simultan berpengaruh signifikan terhadap variabel terikat.

Contoh:

- Apakah X1, X2, Z1, dan Z2 secara bersama-sama mempengaruhi Y?

Jika nilai F-hitung $>$ F-tabel dan signifikansi $< 0,05$, maka model secara keseluruhan dianggap signifikan.

3.6.5.3 Koefisien Determinasi (R^2 dan Adjusted R^2)

Koefisien R^2 digunakan untuk mengukur seberapa besar proporsi variabel terikat (Y) yang dapat dijelaskan oleh variabel bebas dan moderasi.

- Nilai R^2 mendekati 1 menunjukkan model mampu menjelaskan variabilitas yang tinggi dari pergerakan harga kripto berdasarkan sentimen media sosial dan volume perdagangan.
- Adjusted R^2 digunakan untuk mempertimbangkan jumlah variabel prediktor dalam model, terutama penting saat melibatkan variabel moderasi.