

Analisis Pola Emas Sebagai Safe Haven dengan Metode K-Means

Nur Rahman Saleh^{1*}, Yola Aulia Wanergi², Adi Tegar Setiawan³, Faizal Wiyatnodito⁴

¹Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

^{1*}230103073@mhs.udb.ac.id

²Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

²230103081@mhs.udb.ac.id

³Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

³230103044@mhs.udb.ac.id

⁴Teknik Informatika/Fakultas
Ilmu Komputer
(Universitas Duta Bangsa
Surakarta)

⁴230103059@mhs.udb.ac.id

Abstrak— Emas merupakan salah satu instrumen investasi yang sering dijadikan aset perlindungan (*safe haven*) ketika terjadi ketidakpastian ekonomi global. Namun, perilaku emas sebagai *safe haven* tidak selalu menunjukkan pola yang konsisten pada setiap periode waktu, sehingga diperlukan metode analisis yang mampu mengidentifikasi pola tersebut secara sistematis. Penelitian ini bertujuan menganalisis perilaku emas sebagai *asset safe haven* melalui pendekatan K-Means Clustering berbasis indikator ekonomi global menggunakan dataset historis dari Kaggle berjudul *Gold Price and Global Economic Indicators 2015–2026* yang terdiri dari 4.137 observasi harian. Variabel yang digunakan mencakup harga emas (XAU/USD), indeks dolar Amerika Serikat (DXY), harga minyak mentah, dan tingkat inflasi. Proses analisis mengikuti kerangka kerja CRISP-DM dengan evaluasi menggunakan *Elbow Method*, *Davies-Bouldin Index*, dan *Silhouette Score* yang menghasilkan nilai tertinggi sebesar 0,3867 pada k=3. Hasil analisis menunjukkan tiga cluster optimal yang merepresentasikan kondisi ekonomi berbeda, yaitu *Safe Haven* Penuh (Inflasi Tinggi), Penguatan Emas Tanpa Tekanan Inflasi, dan Dominasi Dolar AS. Indeks dolar Amerika Serikat terbukti menjadi faktor paling dominan dengan korelasi negatif sangat kuat sebesar -0,948, sementara inflasi memiliki pengaruh linear yang sangat lemah terhadap harga emas. Hasil penelitian diimplementasikan dalam aplikasi berbasis web menggunakan Streamlit sebagai alat bantu rekomendasi keputusan investasi emas secara interaktif.

Kata kunci— emas, *safe haven*, K-Means Clustering, indikator ekonomi global, CRISP-DM

Abstract— Gold is one of the investment instruments frequently used as a safe haven asset during periods of global economic uncertainty. However, gold's behavior as a safe haven does not always exhibit a consistent pattern across different time periods, making it necessary to apply an analytical method capable of systematically identifying such patterns. This study aims to analyze gold's behavior as a safe haven asset through a K-Means Clustering approach based on global economic indicators using a historical dataset from Kaggle titled *Gold Price and Global Economic Indicators 2015–2026*, consisting of 4,137 daily observations. The variables used include gold price (XAU/USD), the US Dollar Index (DXY), crude oil price, and inflation rate. The analysis process follows the CRISP-DM framework with evaluation conducted using the Elbow Method, Davies-Bouldin Index, and Silhouette Score of 0.3867 at k=3. The results reveal three optimal clusters representing distinct economic conditions, namely Full Safe Haven (High Inflation), Gold Strengthening Without Inflation Pressure, and US Dollar Dominance. The US Dollar Index proved to be the most dominant factor with a very strong negative correlation of -0.948, while inflation showed a very weak linear influence on gold prices. The research findings were implemented in a web-based application using Streamlit as an interactive tool for gold investment decision recommendations.

Keywords— gold, *safe haven*, K-Means Clustering, global economic indicators, CRISP-DM

I. PENDAHULUAN

Emas merupakan salah satu instrumen investasi yang memiliki peran penting dalam sistem keuangan global karena sering digunakan sebagai aset alternatif ketika terjadi ketidakpastian ekonomi dan tekanan pasar keuangan [1]. Karakteristik emas yang

cenderung mempertahankan nilai menjadikannya salah satu aset yang banyak dipertimbangkan investor sebagai *safe haven asset* dalam menghadapi risiko ekonomi global [2]. Peningkatan perhatian terhadap emas sebagai aset perlindungan semakin terlihat ketika terjadi berbagai kondisi seperti

pandemi, konflik geopolitik, inflasi, serta perubahan kebijakan moneter global yang memengaruhi pasar keuangan [3]. Selain itu, hubungan antara emas dengan indikator ekonomi global seperti indeks dolar Amerika Serikat (DXY), harga minyak, dan kondisi pasar saham menjadi faktor penting dalam memahami pergerakan emas secara lebih luas [4]. Struktur ketergantungan antara indeks dolar dan harga minyak mentah ini sering kali menunjukkan pola yang dinamis dan berubah-ubah dalam merespons gejolak pasar keuangan global [5]. Penelitian terdahulu menunjukkan bahwa emas memiliki kemampuan untuk memberikan perlindungan terhadap risiko pasar, meskipun tingkat efektivitasnya dapat berbeda tergantung kondisi ekonomi yang sedang terjadi [6].

Namun, meskipun emas sering dianggap sebagai aset yang stabil, perilaku emas sebagai *safe haven* tidak selalu menunjukkan pola yang sama pada setiap periode waktu [6]. Perubahan kondisi ekonomi global seperti peningkatan risiko geopolitik, krisis finansial, dan ketidakstabilan pasar dapat menyebabkan perubahan hubungan antara emas dengan aset keuangan lainnya [7]. Penelitian mengenai hubungan emas dan aset komoditas lain menunjukkan bahwa adanya krisis global dapat meningkatkan ketergantungan dan risiko *spillover* antar pasar sehingga memengaruhi fungsi emas sebagai aset perlindungan [7], [8]. Selain itu, kondisi pasar yang dinamis menyebabkan investor sulit menentukan kapan emas benar-benar berperan sebagai *safe haven* dan kapan emas hanya mengalami kenaikan harga akibat faktor spekulasi [9]. Oleh karena itu, diperlukan metode analisis yang mampu mengidentifikasi pola perilaku emas berdasarkan kondisi ekonomi global yang berubah-ubah [3].

Permasalahan utama dalam analisis emas sebagai *safe haven* adalah kompleksitas hubungan antara harga emas dengan berbagai indikator ekonomi global yang memiliki karakteristik berbeda [4]. Analisis menggunakan pendekatan konvensional sering kali hanya melihat hubungan linear sehingga belum mampu menggambarkan kelompok kondisi pasar yang memiliki pola serupa [10]. Salah satu pendekatan yang dapat digunakan untuk mengatasi permasalahan tersebut adalah metode *clustering* yang mampu mengelompokkan data berdasarkan

kesamaan karakteristik tanpa membutuhkan label kelas sebelumnya [11]. Algoritma K-Means telah banyak digunakan dalam penelitian bidang ekonomi dan keuangan untuk menemukan pola tersembunyi berdasarkan indikator tertentu [12]. Dengan pendekatan *clustering*, perilaku emas dapat dianalisis berdasarkan kelompok kondisi ekonomi sehingga dapat memberikan gambaran mengenai karakteristik emas sebagai aset *safe haven* secara lebih sistematis.

Penelitian sebelumnya menunjukkan bahwa metode *clustering* mampu membantu proses analisis data ekonomi dengan cara mengelompokkan objek berdasarkan kemiripan pola yang dimiliki [10]. Fadhila et al. menerapkan metode K-Means *clustering* untuk menganalisis pola inflasi berdasarkan kondisi sebelum dan sesudah pandemi sehingga menghasilkan kelompok karakteristik ekonomi yang berbeda [10]. Selain itu, penelitian Pranata et al. menunjukkan bahwa metode *clustering* dapat digunakan pada bidang keuangan dengan memanfaatkan indikator fundamental untuk menemukan kelompok data yang memiliki pola serupa [11]. Pada penelitian mengenai aset *safe haven*, Gupta et al. menggabungkan pendekatan *machine learning* dengan analisis risiko geopolitik untuk memahami kemampuan aset dalam menghadapi ketidakpastian global [3]. Di sisi lain, integrasi model komputasi *machine learning* juga semakin diandalkan karena mampu memproyeksikan ketidakstabilan harga emas dengan tingkat akurasi yang lebih tinggi [13]. Berdasarkan penelitian tersebut, pendekatan *clustering* berbasis indikator ekonomi global memiliki peluang untuk diterapkan dalam menganalisis perilaku emas sebagai *safe haven asset*.

Penelitian ini bertujuan untuk menganalisis perilaku emas sebagai *asset safe haven* melalui pendekatan *clustering* berbasis indikator ekonomi global. Penelitian ini memanfaatkan data historis harga emas dan indikator ekonomi global yang diperoleh dari dataset Kaggle sebagai dasar proses analisis. Metode K-Means *clustering* digunakan untuk mengelompokkan kondisi pasar berdasarkan karakteristik indikator ekonomi yang memiliki pola serupa [11]. Hasil pengelompokan diharapkan mampu memberikan gambaran mengenai kondisi ketika emas menunjukkan karakteristik sebagai aset

perlindungan maupun ketika fungsi *safe haven* mengalami perubahan. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam memahami pola perilaku emas serta mendukung pengambilan keputusan investasi berbasis analisis data.

II. METODOLOGI PENELITIAN

Penelitian dilakukan melalui tahapan-tahapan terstruktur berdasarkan kerangka kerja *Cross Industry Standard Process for Data Mining* (CRISP-DM), mulai dari pemahaman data, persiapan data, pemodelan hingga evaluasi hasil analisis.

2.1 Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder berupa dataset publik *Gold Price and Global Economic Indicators 2015–2026* yang diperoleh dari platform Kaggle. Dataset tersebut berisi historis harga emas serta indikator ekonomi global, yaitu harga emas (XAU/USD), indeks dolar Amerika Serikat (DXY), harga minyak mentah (*Crude Oil Price*), dan tingkat inflasi (*Inflation Rate*). Dataset tersebut digunakan sebagai sumber utama dalam proses pengolahan data, pembentukan model *clustering*, dan analisis hubungan antara pergerakan emas dengan kondisi ekonomi global.

2.2 Metode Pengumpulan Data

2.2.1 Dokumentasi

Pengumpulan data utama dilakukan menggunakan metode dokumentasi melalui sumber data online. Dataset yang digunakan diperoleh dari platform Kaggle dengan judul *Gold Price and Global Economic Indicators 2015–2026* yang berisi data historis mengenai harga emas dan indikator ekonomi global. Dataset tersebut dipilih karena menyediakan variabel yang sesuai dengan kebutuhan penelitian, yaitu faktor ekonomi yang berpotensi memengaruhi perubahan harga emas.

2.2.2 Studi Pustaka

Studi pustaka dilakukan dengan mengumpulkan berbagai penelitian terdahulu yang berkaitan dengan emas sebagai *safe haven asset*, indikator ekonomi global, dan penerapan metode *clustering*. Referensi tersebut digunakan untuk memperkuat landasan teori, mendukung interpretasi hasil penelitian, serta

mengkaji strategi manajemen risiko menggunakan emas pada kondisi ketidakpastian moneter dan geopolitik [14].

2.3 Tahapan Pemodelan CRISP-DM

2.3.1 Pemahaman Bisnis

Tahap ini bertujuan menentukan fokus permasalahan dengan mengidentifikasi pola perilaku emas sebagai *safe haven asset* berdasarkan perubahan indikator ekonomi global. Penelitian tidak hanya menganalisis perubahan harga emas, tetapi juga bertujuan membangun model *clustering* untuk mengelompokkan kondisi pasar berdasarkan karakteristik indikator ekonomi yang menyertainya.

2.3.2 Pemahaman Data

Tahap pemahaman data dilakukan untuk menganalisis karakteristik dataset melalui eksplorasi jumlah atribut, tipe data, distribusi nilai, serta hubungan antar variabel. Proses ini bertujuan memahami kondisi awal data harga emas dan indikator ekonomi global sehingga atribut yang digunakan sesuai dengan tujuan penelitian sebelum dilakukan tahap transformasi dan pemodelan.

2.3.3 Persiapan Data

Tahap persiapan data dimulai dengan melakukan pemilihan atribut yang relevan terhadap penelitian, yaitu variabel harga emas, indeks dolar, harga minyak mentah, dan inflasi. Setelah atribut ditentukan, dilakukan pemeriksaan terhadap data untuk memastikan tidak terdapat kesalahan format maupun data yang tidak sesuai. Selanjutnya dilakukan proses normalisasi data karena algoritma K-Means menggunakan perhitungan jarak antar data sehingga perbedaan skala variabel dapat memengaruhi hasil pengelompokan [10]. Data yang telah melalui tahap persiapan kemudian digunakan sebagai input dalam proses pembentukan model *clustering*.

2.3.4 Pemodelan

Tahap pemodelan merupakan tahap pembangunan model menggunakan algoritma K-Means *Clustering* yang dipilih karena mampu mengelompokkan data berdasarkan tingkat kemiripan karakteristik antar objek sehingga dapat digunakan untuk menemukan pola kondisi ekonomi tertentu [9]. Pendekatan K-

Means ini juga diakui sangat efektif untuk mengklasifikasikan perubahan rezim ekonomi secara otomatis pada data deret waktu multidimensi [15]. Pada tahap ini dilakukan penentuan jumlah cluster yang sesuai, dilanjutkan dengan proses pembentukan *cluster* secara iteratif hingga diperoleh model yang stabil untuk menganalisis karakteristik setiap kelompok berdasarkan indikator ekonomi global.

2.3.5 Evaluasi

Tahap evaluasi dilakukan dengan menganalisis kualitas pemisahan antar kelompok dan karakteristik data pada setiap *cluster*. Tahap ini bertujuan untuk mengetahui apakah model mampu menggambarkan perbedaan kondisi ekonomi yang berkaitan dengan perilaku emas. Hasil evaluasi menjadi dasar sebelum dilakukan interpretasi dan implementasi model.

2.3.6 Implementasi

Implementasi dilakukan dengan menerapkan hasil analisis *clustering* ke dalam bentuk antarmuka sistem yang dapat digunakan oleh pengguna. Penerapan model ke dalam antarmuka bertujuan agar hasil penelitian dapat digunakan secara lebih praktis dan tidak hanya terbatas pada proses pengolahan data.

III. HASIL DAN PEMBAHASAN

3.1 Pemahaman Bisnis

Hasil pengamatan terhadap data menunjukkan bahwa harga emas memiliki pola pergerakan yang berbeda pada setiap kondisi ekonomi global. Perubahan indeks dolar Amerika Serikat, harga minyak mentah, dan tingkat inflasi menunjukkan hubungan yang tidak selalu sama terhadap pergerakan harga emas selama periode pengamatan. Pada beberapa periode, harga emas mengalami kenaikan ketika dolar Amerika Serikat melemah, sedangkan pada periode lainnya harga emas bergerak mengikuti perubahan harga komoditas dunia. Kondisi tersebut menunjukkan bahwa perilaku emas sebagai aset *safe haven* tidak dipengaruhi oleh satu faktor saja, melainkan oleh kombinasi beberapa indikator ekonomi. Oleh karena itu, diperlukan pengelompokan data untuk mengidentifikasi pola kondisi ekonomi yang

memiliki karakteristik serupa terhadap pergerakan harga emas.

3.2 Pemahaman Data

Dataset yang digunakan terdiri dari 4.137 data observasi yang merepresentasikan kondisi ekonomi global selama periode 2015–2026. Variasi nilai pada harga emas, indeks dolar Amerika Serikat, harga minyak mentah, dan tingkat inflasi menunjukkan karakteristik data yang dinamis sehingga mendukung proses analisis dan pengelompokan kondisi ekonomi.

3.2.1 Deskripsi Dataset

Dataset penelitian terdiri atas lima atribut yang meliputi tanggal pengamatan, harga emas, indeks dolar Amerika Serikat (DXY), harga minyak mentah dunia, dan tingkat inflasi. Seluruh variabel tersebut berhasil dikumpulkan dalam periode pengamatan tanpa ditemukan data yang hilang maupun data duplikat. Kondisi tersebut menunjukkan bahwa kualitas data yang digunakan dalam penelitian sudah cukup baik untuk dilakukan analisis lebih lanjut. Selain itu, seluruh variabel memiliki peran yang berbeda dalam menggambarkan kondisi ekonomi global. Kombinasi variabel tersebut digunakan untuk membentuk karakteristik *cluster* pada tahap pemodelan.

Tabel 1. Deskripsi Variabel Dataset

Variabel	Tipe Data	Keterangan
Date	Tanggal	Tanggal observasi harian
Gold_Price_XAU_USD	Numerik	Harga emas dalam USD per troy ounce
US_Dollar_Index_DXY	Numerik	Indeks kekuatan dolar Amerika Serikat
Crude_Oil_Price	Numerik	Harga minyak mentah dunia
Inflation_Rate_Pct	Numerik	Tingkat inflasi dalam persen

3.2.2 Statistik Deskriptif

Hasil statistik deskriptif pada Tabel 2 menunjukkan bahwa harga emas memiliki nilai rata-rata sebesar 2.825,36 USD dengan nilai maksimum mencapai 4.176,60 USD dan nilai minimum sebesar 1.106,38 USD. Nilai standar deviasi sebesar 927,41 menunjukkan bahwa harga emas mengalami fluktuasi yang cukup besar selama periode

penelitian. Variabel indeks dolar Amerika Serikat memiliki rata-rata sebesar 55,31 dengan standar deviasi 17,15, yang menunjukkan adanya perubahan nilai dolar yang cukup signifikan dari waktu ke waktu. Sementara itu, harga minyak mentah memiliki rata-rata sebesar 274,74 dengan standar deviasi 88,50, sedangkan tingkat inflasi memiliki rata-rata 5,07% dengan standar deviasi 2,01%. Perbedaan nilai standar deviasi antar variabel menunjukkan bahwa setiap indikator ekonomi memiliki tingkat variasi yang berbeda selama periode pengamatan.

Tabel 2. Statistik Deskriptif Variabel Penelitian

Variabel	Mean	Median	Std Dev	Min	Max
Gold_Price	2.825,36	3.047,79	927,41	1.106,38	4.176,60
DXY	55,31	53,11	17,15	27,05	97,09
Crude_Oil	274,74	282,7	88,5	47,85	463,81
Inflation (%)	5,07	5,09	2,01	1,5	8,5

Berdasarkan perbandingan nilai *mean* dan median, harga emas menunjukkan distribusi yang cenderung negatif karena nilai rata-ratanya lebih kecil dibandingkan median. Kondisi tersebut mengindikasikan bahwa terdapat beberapa periode dengan harga emas relatif rendah yang menarik nilai rata-rata ke bawah. Sebaliknya, variabel DXY menunjukkan kecenderungan distribusi positif karena nilai rata-ratanya lebih tinggi dibandingkan median. Variabel inflasi memiliki nilai mean dan median yang hampir sama sehingga menunjukkan distribusi yang relatif seimbang.

3.2.3 Analisis Korelasi

Analisis korelasi pada penelitian ini dilakukan menggunakan metode *Pearson Correlation* untuk mengukur arah dan kekuatan hubungan linear antar variabel penelitian, yaitu harga emas (*Gold Price*), indeks dolar Amerika Serikat (*DXY*), harga minyak mentah (*Crude Oil Price*), dan tingkat inflasi (*Inflation Rate*). Hasil analisis pada Tabel 3 menunjukkan bahwa harga emas memiliki hubungan negatif yang sangat kuat dengan indeks dolar AS ($r = -0,948$) dan hubungan positif yang sangat kuat dengan harga minyak mentah ($r = 0,937$), sedangkan hubungan dengan tingkat inflasi sangat lemah ($r =$

0,011). Temuan ini mengindikasikan bahwa selama periode penelitian, pergerakan harga emas lebih dipengaruhi oleh indeks dolar Amerika Serikat dan harga minyak mentah dibandingkan tingkat inflasi.

Tabel 3. Hasil Analisis Korelasi

Pasangan Variabel	Koefisien Korelasi	Interpretasi
Gold_Price vs DXY	-0,948	Korelasi negatif sangat kuat
Gold_Price vs Crude Oil	0,937	Korelasi positif sangat kuat
Gold_Price vs Inflation	0,011	Hampir tidak ada korelasi linear

3.3 Persiapan Data

3.3.1 Hasil Standarisasi Data

Hasil standarisasi menunjukkan bahwa seluruh variabel berhasil ditransformasikan ke dalam bentuk Z-score sehingga memiliki rata-rata mendekati nol dan standar deviasi mendekati satu. Transformasi ini menghasilkan empat variabel baru yaitu *Z_Gold*, *Z_DXY*, *Z_CrudeOil*, dan *Z_Inflation* yang digunakan sebagai atribut utama dalam proses *clustering*. Dengan adanya standarisasi, kontribusi setiap variabel menjadi lebih seimbang sehingga tidak ada atribut yang mendominasi proses pengelompokan. Kondisi ini sangat penting karena algoritma K-Means menggunakan perhitungan jarak untuk menentukan keanggotaan suatu data pada *cluster* tertentu. Hasil standarisasi memungkinkan pola hubungan antar variabel dianalisis secara lebih objektif. Rumus standarisasi Z-score yang digunakan adalah:

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

Keterangan:

Z = nilai hasil standarisasi

X = nilai data asli

μ = rata-rata (*mean*)

σ = standar deviasi

3.3.2 Penanganan Atribut Tanggal

Atribut *Date* tidak digunakan dalam proses perhitungan *cluster* karena tidak merepresentasikan karakteristik ekonomi yang menjadi fokus penelitian. Penggunaan atribut tanggal dalam perhitungan jarak dapat menyebabkan data yang berdekatan secara waktu cenderung berada pada

cluster yang sama meskipun memiliki kondisi ekonomi yang berbeda. Oleh karena itu, atribut *Date* hanya digunakan sebagai identifier untuk menunjukkan waktu pengamatan setiap data. Keputusan ini memungkinkan proses *clustering* lebih berfokus pada pola hubungan antar indikator ekonomi dibandingkan urutan waktu pengamatan. Dengan demikian, *cluster* yang terbentuk benar-benar merepresentasikan kemiripan kondisi ekonomi yang terjadi selama periode penelitian.

3.3.3 Ringkasan Hasil Persiapan Data

Hasil akhir data *preparation* ditunjukkan pada Tabel 4. Seluruh variabel numerik berhasil ditransformasikan menjadi variabel hasil standarisasi yang digunakan sebagai dasar pembentukan *cluster*. Sementara itu, atribut *Date* tetap dipertahankan sebagai identifier dan tidak dilibatkan dalam perhitungan jarak. Struktur data yang dihasilkan telah memenuhi kebutuhan algoritma K-Means karena seluruh atribut memiliki skala yang seragam dan tidak terdapat masalah kualitas data. Oleh karena itu, dataset hasil *preparation* dapat langsung digunakan pada tahap pemodelan.

Tabel 4. Ringkasan Hasil Persiapan Data

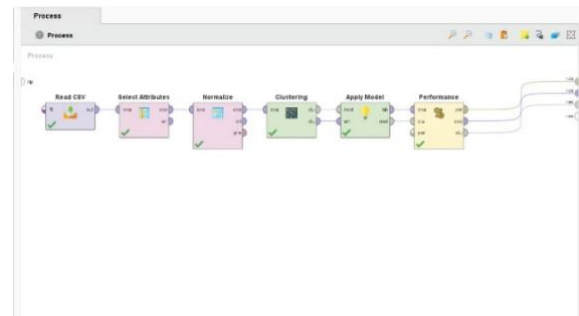
Variabel Asli	Variabel Hasil Standarisasi	Peran dalam Modeling
Gold_Price_XAU_USD	Z_Gold	Variabel clustering
US_Dollar_Index_DXY	Z_DXY	Variabel clustering
Crude_Oil_Price	Z_CrudeOil	Variabel clustering
Inflation_Rate_Pct	Z_Inflation	Variabel clustering
Date	(tidak diubah)	Identifier (id), tidak dihitung jaraknya

3.4 Pemodelan

3.4.1 Hasil Workflow Pemodelan

Proses *clustering* dilakukan menggunakan *workflow* yang dibangun pada Altair AI Studio. *Workflow* tersebut terdiri atas operator *Read Excel*, *Set Role*, *Select Attributes*, K-Means, dan *Cluster Distance Performance*. Susunan *workflow* menunjukkan bahwa proses analisis dilakukan secara sistematis mulai dari pembacaan data hingga evaluasi hasil *clustering*. Seluruh atribut hasil standarisasi digunakan sebagai input utama dalam proses pengelompokan. *Workflow* yang dibangun

berhasil menghasilkan model *clustering* yang dapat digunakan untuk tahap evaluasi dan interpretasi *cluster*.



Gambar 1. Workflow Altair

3.4.2 Penentuan Jumlah Cluster Optimal

Hasil pengujian jumlah cluster ditampilkan pada Tabel 5. Berdasarkan nilai *Average Within Centroid Distance*, terlihat bahwa semakin banyak jumlah *cluster* yang digunakan maka nilai jarak rata-rata terhadap *centroid* semakin kecil. Pada $k=2$ diperoleh nilai sebesar 1,785 kemudian menurun menjadi 1,329 pada $k=3$. Penurunan tersebut menunjukkan bahwa anggota *cluster* menjadi semakin dekat dengan pusat *cluster* masing-masing. Namun setelah $k=3$, penurunan nilai yang terjadi mulai melambat sehingga tambahan *cluster* tidak lagi memberikan peningkatan kualitas yang signifikan. Oleh karena itu, hasil pengujian ini selanjutnya digunakan sebagai dasar dalam analisis *Elbow Method* untuk menentukan jumlah *cluster* yang optimal.

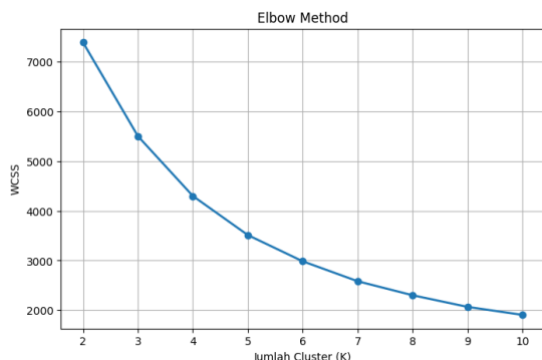
Tabel 5. Hasil Pengujian Jumlah Cluster

k	Avg. Within Centroid Distance	Davies-Bouldin Index
2	1,785	0,829
3	1,329	0,961
4	1,039	1,112
5	0,849	0,969

3.4.3 Analisis Elbow Method

Berdasarkan hasil evaluasi pada Tabel 5, penurunan nilai *Average Within Centroid Distance* terbesar terjadi pada perubahan dari $k=2$ menuju $k=3$ dengan selisih sebesar 0,456. Hasil tersebut didukung oleh visualisasi *Within-Cluster Sum of Squares* (WCSS) pada metode Elbow yang juga menunjukkan terbentuknya titik siku (*elbow point*) pada $k=3$, menandakan bahwa *cluster* yang

terbentuk sudah cukup kompak. Setelah titik tersebut, penurunan nilai menjadi lebih kecil, yaitu 0,290 pada $k=4$ dan 0,190 pada $k=5$, sehingga penambahan jumlah cluster tidak memberikan peningkatan yang signifikan. Oleh karena itu, $k=3$ dipilih sebagai jumlah *cluster* yang digunakan dalam penelitian ini.

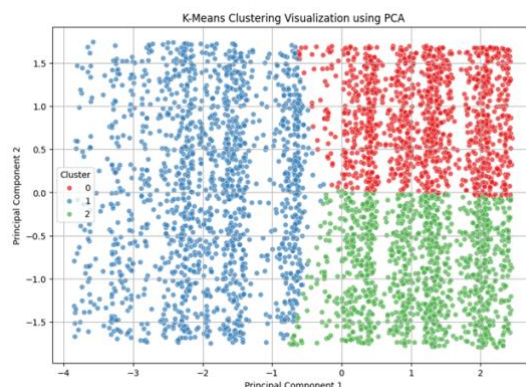


Gambar 2. Visualisasi Elbow Method

3.4.4 Hasil Clustering

Setelah jumlah *cluster* ditetapkan sebanyak tiga *cluster*, proses *clustering* dilakukan menggunakan algoritma K-Means dengan menghitung jarak setiap data ke *centroid* menggunakan metode *Euclidean Distance*. Data dikelompokkan ke *cluster* dengan jarak terdekat, dan proses diulang hingga *centroid* stabil. Hasil akhir menunjukkan bahwa seluruh data berhasil dikelompokkan ke dalam tiga *cluster* yang memiliki karakteristik berbeda. Gambar 3 memperlihatkan visualisasi hasil *clustering* menggunakan *Principal Component Analysis* (PCA), yang mereduksi empat variabel menjadi dua komponen utama sehingga pemisahan antar *cluster* dapat diamati dengan lebih jelas. Rumus *Euclidean Distance* yang digunakan adalah:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$



Gambar 3. Visualisasi Cluster

3.5 Evaluasi

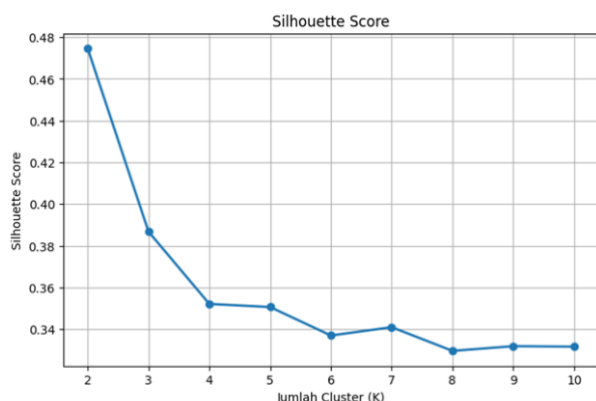
3.5.1 Evaluasi Davies–Bouldin Index

Evaluasi kualitas hasil *clustering* dilakukan menggunakan *Davies-Bouldin Index* (DBI) untuk mengukur tingkat kekompakan data di dalam setiap *cluster* (*intra-cluster*) serta tingkat pemisahan antar *cluster* (*inter-cluster*). Semakin kecil nilai DBI yang diperoleh, semakin baik kualitas *clustering* yang dihasilkan karena menunjukkan bahwa setiap *cluster* memiliki tingkat kemiripan yang tinggi di dalam kelompoknya dan perbedaan yang jelas dengan *cluster* lainnya. Berdasarkan Tabel 5, model K-Means dengan tiga *cluster* memperoleh nilai *Davies-Bouldin Index* sebesar 0,961. Nilai tersebut menunjukkan bahwa hasil *clustering* memiliki tingkat kekompakan dan pemisahan antar *cluster* yang cukup baik sehingga mampu merepresentasikan karakteristik data penelitian. Hasil evaluasi ini mengindikasikan bahwa model yang dibangun telah menghasilkan struktur *cluster* yang memadai dan layak digunakan sebagai dasar dalam proses interpretasi karakteristik masing-masing *cluster*. Adapun nilai DBI yang lebih rendah pada pengujian jumlah *cluster* lainnya digunakan sebagai informasi pembandingan, sedangkan jumlah *cluster* yang digunakan dalam penelitian telah ditentukan sebelumnya melalui analisis *Elbow Method*.

3.5.2 Analisis Silhouette Score

Berdasarkan hasil evaluasi menggunakan *Silhouette Score* pada Gambar 4, diperoleh nilai *silhouette* tertinggi sebesar 0,3867 pada jumlah *cluster* (k) = 3. Nilai tersebut menunjukkan bahwa hasil *clustering* memiliki kualitas yang cukup baik, di mana sebagian besar data memiliki tingkat

kemiripan yang lebih tinggi dengan anggota pada *cluster* yang sama dibandingkan dengan *cluster* lainnya. Meskipun nilai tersebut belum mendekati 1 yang menandakan pemisahan *cluster* sangat baik, nilai 0,3867 mengindikasikan bahwa struktur *cluster* yang terbentuk telah mampu merepresentasikan pola data secara cukup optimal dengan tingkat pemisahan antar *cluster* yang memadai dan tingkat kesamaan yang cukup tinggi di dalam masing-masing *cluster*. Selain itu, terlihat bahwa nilai *Silhouette Score* mengalami penurunan ketika jumlah *cluster* ditingkatkan menjadi 4 hingga 10, yang menunjukkan bahwa penambahan jumlah *cluster* justru mengurangi kualitas pengelompokan data. Oleh karena itu, hasil evaluasi *Silhouette Score* memperkuat hasil *Elbow Method* yang sebelumnya telah menunjukkan bahwa penggunaan tiga *cluster* ($k = 3$) merupakan jumlah *cluster* yang paling optimal untuk digunakan pada dataset penelitian ini.



Gambar 4. Visualisasi Silhouette Score

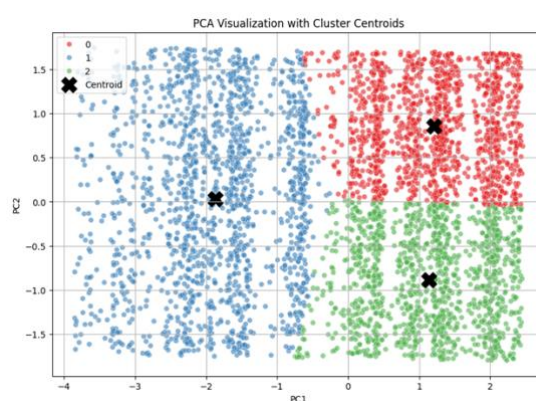
3.5.3 Karakteristik, Interpretasi, dan Penamaan Cluster

Hasil *centroid* pada Tabel 6 menunjukkan bahwa setiap *cluster* memiliki karakteristik ekonomi yang berbeda. Perbedaan tersebut terlihat dari variasi nilai *Z_Gold*, *Z_DXY*, *Z_CrudeOil*, dan *Z_Inflation* pada masing-masing *cluster*. Nilai *centroid* yang bernilai positif menunjukkan kondisi di atas rata-rata, sedangkan nilai negatif menunjukkan kondisi di bawah rata-rata. Untuk mempermudah interpretasi hasil *clustering*, Gambar 5 menampilkan visualisasi menggunakan *Principal Component Analysis* (PCA) yang mereduksi empat variabel menjadi dua komponen utama (PC1 dan PC2) tanpa mengubah struktur pengelompokan data secara signifikan.

Melalui visualisasi tersebut, posisi *centroid* setiap *cluster* dapat terlihat lebih jelas sehingga memudahkan dalam mengamati pemisahan antar *cluster*. Oleh karena itu, analisis *centroid* digunakan untuk memahami karakteristik ekonomi yang direpresentasikan oleh setiap *cluster*.

Tabel 6. Centroid hasil K-Means dengan $k=3$

Atribut	Cluster_0	Cluster_1	Cluster_2
Z_Gold	0,722	0,684	-1,123
Z_DXY	-0,738	-0,676	1,129
Z_CrudeOil	0,641	0,598	-0,99
Z_Inflation	0,855	-0,889	0,016



Gambar 3. Visualisasi Centroid

Cluster 0 memiliki nilai *Z_Gold* sebesar 0,722, *Z_DXY* sebesar -0,738, *Z_CrudeOil* sebesar 0,641, dan *Z_Inflation* sebesar 0,855. Karakteristik tersebut menunjukkan kondisi ketika harga emas dan harga minyak berada di atas rata-rata, dolar Amerika Serikat berada di bawah rata-rata, serta tingkat inflasi relatif tinggi. Kondisi ini menggambarkan situasi ketika investor cenderung beralih ke aset yang dianggap aman akibat meningkatnya tekanan ekonomi. Oleh karena itu, *Cluster 0* diberi nama *Safe Haven Penuh* (Inflasi Tinggi).

Cluster 1 memiliki nilai *Z_Gold* sebesar 0,684, *Z_DXY* sebesar -0,676, *Z_CrudeOil* sebesar 0,598, dan *Z_Inflation* sebesar -0,889. Karakteristik ini menunjukkan bahwa harga emas dan harga minyak tetap berada di atas rata-rata meskipun tingkat inflasi berada di bawah rata-rata. Temuan tersebut menunjukkan bahwa kenaikan harga emas tidak selalu dipicu oleh inflasi yang tinggi. Pelemahan dolar Amerika Serikat tampak menjadi faktor yang

lebih dominan dalam mendorong kenaikan harga emas pada cluster ini. Oleh karena itu, *Cluster 1* diberi nama Penguatan Emas Tanpa Tekanan Inflasi.

Cluster 2 menunjukkan karakteristik yang paling berbeda dibandingkan *cluster* lainnya. Nilai *Z_DXY* sebesar 1,129 menunjukkan kondisi dolar Amerika Serikat yang sangat kuat, sedangkan nilai *Z_Gold* sebesar -1,123 dan *Z_CrudeOil* sebesar -0,990 menunjukkan harga emas dan harga minyak yang berada jauh di bawah rata-rata. Kondisi tersebut mengindikasikan bahwa investor lebih memilih memegang dolar dibandingkan emas ketika nilai dolar mengalami penguatan. Oleh karena itu, *Cluster 2* diberi nama Dominasi Dolar AS.

Tabel 7. Interpretasi dan Penamaan Cluster

Cluster	Karakteristik	Label / Nama Kondisi
Cluster_0	Dolar lemah, emas & minyak tinggi, inflasi tinggi	Safe Haven Penuh (Inflasi Tinggi)
Cluster_1	Dolar lemah, emas & minyak tinggi, inflasi rendah	Penguatan Emas Tanpa Tekanan Inflasi
Cluster_2	Dolar kuat, emas & minyak rendah, inflasi netral	Dominasi Dolar AS

Hasil interpretasi menunjukkan bahwa hubungan antara harga emas dan indeks dolar Amerika Serikat merupakan faktor yang paling konsisten dalam membentuk pola *cluster*. Temuan ini juga memperkuat hasil analisis korelasi yang sebelumnya menunjukkan hubungan negatif sangat kuat antara kedua variabel tersebut dengan nilai korelasi sebesar -0,948.

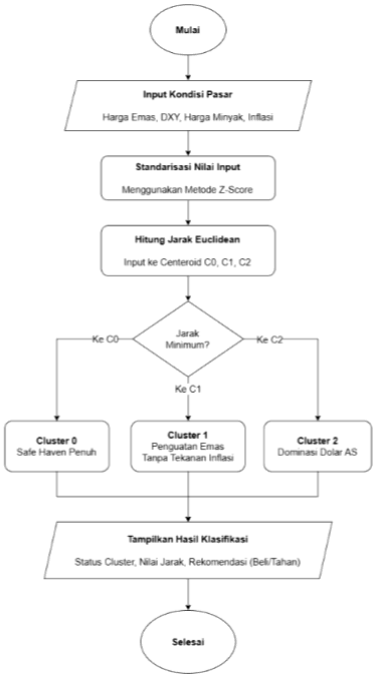
3.6 Implementasi

Implementasi dilakukan menggunakan *framework Streamlit*, yang dipilih karena mampu membangun aplikasi berbasis web secara interaktif dengan proses pengembangan yang sederhana dan efisien menggunakan bahasa pemrograman Python.

3.6.1 Alur Kerja Sistem

Flowchart pada Gambar 6 menggambarkan alur kerja implementasi sistem rekomendasi kondisi investasi emas berbasis metode K-Means. Proses dimulai dengan pengguna memasukkan data kondisi pasar berupa harga emas, indeks dolar AS (DXY), harga minyak, dan tingkat inflasi melalui aplikasi Streamlit. Data kemudian dinormalisasi menggunakan metode *Z-Score*, dilanjutkan dengan

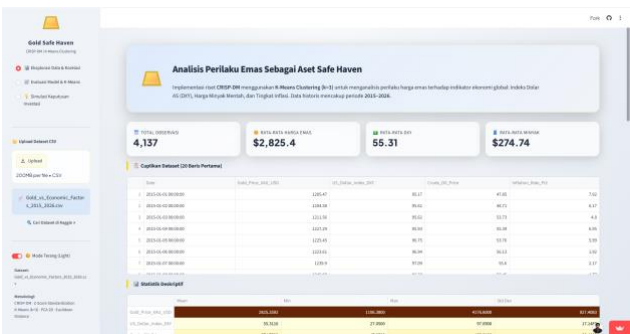
perhitungan *Euclidean Distance* terhadap centroid setiap *cluster*. Sistem selanjutnya menentukan *cluster* berdasarkan nilai jarak terkecil, kemudian menampilkan hasil klasifikasi berupa status *cluster*, nilai jarak, serta rekomendasi keputusan investasi, yaitu Beli atau Tahan, sesuai dengan karakteristik *cluster* yang diperoleh.



Gambar 4. Alur Kerja Sistem

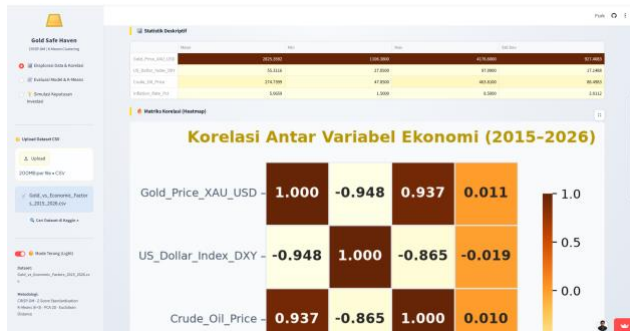
3.6.2 Halaman Eksplorasi Data dan Korelasi

Halaman Eksplorasi Data merupakan tahap awal implementasi sistem yang bertujuan untuk memberikan gambaran umum mengenai dataset yang digunakan sebelum dilakukan proses *clustering* menggunakan algoritma K-Means.



Gambar 5. Halaman Utama Eksplorasi Data

Pada halaman ini ditampilkan informasi ringkas mengenai dataset, seperti jumlah observasi, rata-rata harga emas, rata-rata indeks dolar Amerika Serikat (DXY), serta rata-rata harga minyak mentah. Di bagian bawah halaman tersedia statistik deskriptif yang berisi nilai mean, minimum, maksimum, dan standar deviasi dari setiap variabel penelitian.



Gambar 6. Halaman Visualisasi

Gambar ini menampilkan hasil visualisasi heatmap matriks korelasi yang digunakan untuk mengetahui hubungan antar variabel ekonomi, yaitu harga emas, indeks dolar AS (DXY), harga minyak mentah, dan tingkat inflasi.



Gambar 7. Halaman Insight

Pada bagian Insight Korelasi, sistem menyajikan interpretasi otomatis terhadap hasil analisis korelasi sehingga pengguna tidak perlu menafsirkan nilai korelasi secara manual. Penyajian insight ini bertujuan untuk memudahkan pengguna dalam memahami pola hubungan antar variabel sebelum melanjutkan ke tahap evaluasi model dan simulasi keputusan investasi.

3.6.3 Halaman Evaluasi Model

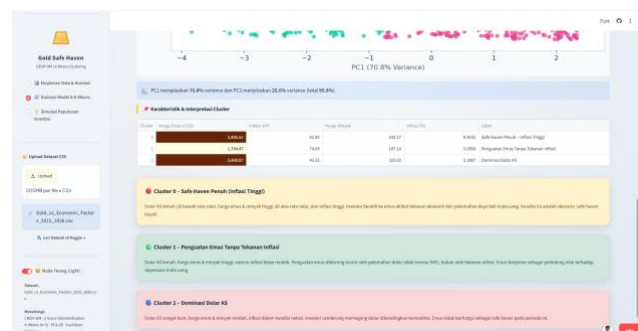
Halaman Evaluasi Model K-Means digunakan untuk menampilkan hasil evaluasi terhadap model *clustering* yang telah dibangun. Pada halaman ini,

sistem menyajikan hasil penentuan jumlah *cluster* terbaik, dan visualisasi hasil *clustering*.



Gambar 8. Halaman Evaluasi Model

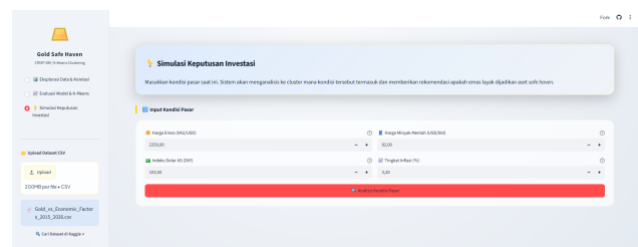
Halaman ini menampilkan karakteristik setiap *cluster* berdasarkan rata-rata nilai harga emas, indeks DXY, harga minyak, dan tingkat inflasi, disertai label serta interpretasi kondisi pasar yang diwakili oleh masing-masing *cluster*.



Gambar 9. Halaman Karakteristik

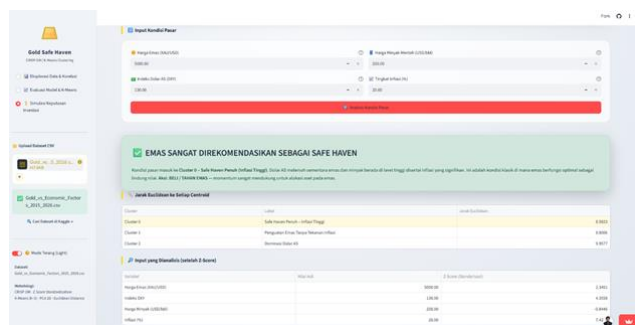
3.6.4 Halaman Simulasi Keputusan Investasi

Halaman Simulasi Keputusan Investasi menyediakan formulir bagi pengguna untuk memasukkan nilai kondisi pasar yang terdiri dari harga emas (XAU/USD), indeks dolar Amerika Serikat (DXY), harga minyak mentah, dan tingkat inflasi. Setelah seluruh data diisi, pengguna dapat menekan tombol Analisis Kondisi Pasar untuk memulai proses normalisasi data, perhitungan jarak *Euclidean* terhadap *centroid* hasil K-Means, serta klasifikasi kondisi pasar berdasarkan *cluster* yang memiliki jarak paling dekat.



Gambar 10. Halaman Simulasi

Setelah proses analisis selesai, sistem menampilkan hasil klasifikasi berupa status rekomendasi investasi beserta cluster yang dipilih berdasarkan nilai *Euclidean Distance* terkecil. Selain itu, sistem juga menyajikan tabel jarak ke setiap *centroid*, nilai input yang telah dinormalisasi menggunakan Z-Score, serta interpretasi kondisi pasar sebagai dasar dalam memberikan rekomendasi apakah emas layak dijadikan aset *safe haven* pada kondisi ekonomi yang dianalisis.



Gambar 11. Contoh Hasil Simulasi

IV. KESIMPULAN

Penelitian ini berhasil menganalisis perilaku emas sebagai *asset safe haven* melalui pendekatan K-Means *Clustering* berbasis indikator ekonomi global menggunakan dataset historis periode 2015–2026, dengan tiga *cluster* optimal yang ditetapkan pada $k=3$ berdasarkan evaluasi *Elbow Method*, *Davies-Bouldin Index*, dan *Silhouette Score* sebesar 0,3867. Nilai tersebut merupakan yang terbaik karena *Silhouette Score* menurun saat jumlah cluster ditambah menjadi 4 hingga 10, yang berarti penambahan *cluster* justru menurunkan kualitas pengelompokan data. Ketiga *cluster* tersebut merepresentasikan kondisi ekonomi berbeda, yaitu *Safe Haven* Penuh (Inflasi Tinggi), Penguatan Emas Tanpa Tekanan Inflasi, dan Dominasi Dolar AS. Temuan utama menunjukkan indeks dolar Amerika Serikat sebagai faktor paling dominan terhadap harga emas dengan korelasi negatif sangat kuat sebesar -0,948, sedangkan inflasi hanya berpengaruh linear sangat lemah. Hal ini menunjukkan bahwa asumsi umum emas sebagai instrumen lindung nilai inflasi tidak selalu berlaku pada setiap kondisi pasar. Hasil penelitian diimplementasikan dalam aplikasi berbasis web menggunakan *Streamlit* untuk simulasi klasifikasi kondisi pasar dan rekomendasi keputusan

investasi secara interaktif, sebagai kontribusi dalam memahami pola perilaku emas secara lebih sistematis dan alat bantu pengambilan keputusan berbasis data.

REFERENSI

- E. Feder-Sempach, P. Szczepocki, and J. Bogolebska, "Global uncertainty and potential shelters: gold, bitcoin, and currencies as weak and strong safe havens for main world stock markets," *Financial Innovation*, vol. 10, no. 1, Dec. 2024, doi: 10.1186/s40854-023-00589-w.
- Y. Hapsari Dyah and A. Hamid, "Safe haven assets in emerging market (evidence from gold, government bonds, and crude oil in Indonesia 2019–2024)," *International Journal of Applied Finance and Business Studies*, 2026.
- R. Gupta, S. Karmakar, and C. Pierdzioch, "Safe Havens, Machine Learning, and the Sources of Geopolitical Risk: A Forecasting Analysis Using Over a Century of Data," *Comput. Econ.*, vol. 64, no. 1, pp. 487–513, Jul. 2024, doi: 10.1007/s10614-023-10452-w.
- J. Denie, Surachman, N. K. Indrawati, and M. Rahayu, "NEXUS BETWEEN OIL, GOLD PRICE AND DXY INDEX ON INDONESIAN STOCK MARKET DURING GEOPOLITICAL EVENTS (2022 - 2024)," *Revista de Gestao Social e Ambiental*, vol. 18, no. 6, 2024, doi: 10.24857/rgsa.v18n6-142.
- Y. Wang and X. Ding, "Dependence Structure of the US Dollar Index and Crude Oil Prices: A Regime-Switching Copula Approach," *Journal of Mathematical Finance*, vol. 14, no. 02, pp. 172–189, 2024, doi: 10.4236/jmf.2024.142009.
- U. C. Okoroafor and T. Leirvik, "Time-varying market efficiency of safe-haven assets," *Financ. Res. Lett.*, vol. 56, Sep. 2023, doi: 10.1016/j.frl.2023.104024.
- Y. Wang, X. Ding, P. Wang, and Z. Huang, "Impact of global crisis events on the dependence and risk spillover between gold and crude oil: a regime-switching copula approach," *Humanit. Soc. Sci. Commun.*, vol. 11, no. 1, Dec. 2024, doi: 10.1057/s41599-024-04329-y.
- H. Faraj, D. McMillan, and M. Al-Sabah, "The diminishing lustre: Gold's market volatility and the fading safe haven effect," *Global Finance Journal*, vol. 67, Sep. 2025, doi: 10.1016/j.gfj.2025.101145.
- M. S. Chang, M. W. Chen, and P. Ju, "Asymmetry in Hedges, Safe Havens, Flights and Contagion: Unconditional Quantile Regression Approach," *Sage Open*, vol. 13, no. 4, Oct. 2023, doi: 10.1177/21582440231208536.
- R. N. Fadhlila, N. Ulinuha, and M. Hafiyusholeh, "Analysis of Inflation Rates During and After the COVID-19 Pandemic Using the K-Means Clustering Method and Kruskal-Wallis Test," *Jurnal Fourier*, vol. 14, no. 2, pp. 56–67, 2025, doi: 10.14421/fourier.2025.142.56-67.
- K. S. Pranata, A. A. S. Gunawan, and F. L. Gaol, "Development clustering system IDX company with k-means algorithm and DBSCAN based on fundamental indicator and ESG," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 319–327. doi: 10.1016/j.procs.2022.12.142.
- D. Tohendry and D. Jollyta, "PENERAPAN ALGORITMA K-MEANS CLUSTERING UNTUK PENGELOMPOKKAN SAHAM BERDASARKAN PRICE EARNING RATIO DAN PRICE TO BOOK VALUE," *Jurnal Mahasiswa Aplikasi Teknologi Komputer dan Informasi*, vol. 5, no. 1, 2023.
- M. M. Ataey and A. K. Azizi, "Forecasting Gold Price Volatility Using Econometric and Machine-Learning Models," *Journal of Social Sciences and Humanities*, vol. 3, no. 1, pp. 173–188, Jan. 2026, doi: 10.62810/jssh.v3i1.231.
- D. Rahmadani, "The Role of Gold as a Safe Haven Asset in Risk Management Strategies in the Era of Geopolitical and Monetary Uncertainty," 2025.
- Q. Luan and J. Hamp, "Automated regime classification in multidimensional time series data using sliced Wasserstein k-means clustering," *Data Science in Finance and Economics*, vol. 5, no. 3, pp. 387–418, 2025, doi: 10.3934/DSFE.2025016.