

Implementasi Hierarchical Clustering untuk Analisis FDMC Narrative Crypto Berbasis Web

M. Fathir Adha ^{a,1,*}, Hendrik Fery Herdiatmoko ^{a,2}

^a Program Studi Informatika, Fakultas Sains dan Teknologi
Universitas Katolik Musi Charitas
Jl. Bangau No.60, 9 Ilir, Kec. Ilir Tim. II, Kota Palembang, Sumatera Selatan 30114, Indonesia
¹ fathiradha7@gmail.com*; ² hendrik@ukmc.ac.id
* M. Fathir Adha

Submission:27/12/2025, Revision: 04/02/2026, Accepted : 13/04/2026

Abstract

This study implements the Hierarchical Clustering algorithm with Ward linkage and Euclidean distance methods to analyze 26 crypto narratives based on the Fully Diluted Market Cap (FDMC) metric. Using a hybrid method that integrates Waterfall, Cross-Industry Standard Process for Data Mining (CRISP-DM), and Knowledge Discovery in Databases (KDD), data was obtained from the CoinGecko API, manually clustered, and aggregated per narrative. Pre-processing involved logarithmic transformation (log-10) and Z-Score normalization to address power-law distributions and outliers, resulting in a more stable cluster structure. The clustering results mapped the market into five clusters: Bluechip (L1 with FDMC \$2.76T), Growth (PAY, MEME, CEX, DEX, DeFi totaling \$468.22B), Growth (AI, DePIN, DAO, L2, RWA, ORC, GameFi, XCH, DID, PRC, LST with \$192.91B), Speculative (NFT, MET, SocialFi, BTC Eco, W3I with \$17.55B), and Speculative (LPD, GambleFi, FTO, SEC with \$2.34B). The model was validated with a Silhouette Score of 0.650 and a Cophenetic Correlation Coefficient of 0.647, indicating cohesive and representative clusters. A web-based implementation using Django, D3.js, and Chart.js provides interactive visualizations and portfolio recommendations. Contributions include a novel fundamental valuation approach, an adaptive clustering model, and practical analytical tools for investors, with potential expansion to multidimensional metrics in the future.

Keywords: Hierarchical Clustering, FDMC, Crypto Narratives, Data Mining, Web Django.

Abstrak

Penelitian ini mengimplementasikan algoritma Hierarchical Clustering dengan metode Ward linkage dan Euclidean distance untuk menganalisis 26 narasi crypto berdasarkan metrik Fully Diluted Market Cap (FDMC). Melalui metode hybrid yang mengintegrasikan Waterfall, Cross-Industry Standard Process for Data Mining (CRISP-DM), dan Knowledge Discovery in Databases (KDD), data diperoleh dari CoinGecko API, dikelompokkan secara manual, dan diagregasi per narasi. Pra-pemrosesan melibatkan transformasi logaritmik (log-10) dan normalisasi Z-Score untuk mengatasi distribusi power-law dan outlier, menghasilkan struktur klaster yang lebih stabil. Hasil clustering memetakan pasar menjadi lima klaster: Bluechip (L1 dengan FDMC \$2.76T), Growth (PAY, MEME, CEX, DEX, DeFi dengan total \$468.22B), Growth (AI, DePIN, DAO, L2, RWA, ORC, GameFi, XCH, DID, PRC, LST dengan \$192.91B), Speculative (NFT, MET, SocialFi, BTC Eco, W3I dengan \$17.55B) dan Speculative (LPD, GambleFi, FTO, SEC dengan \$2.34B). Model divalidasi dengan Silhouette Score 0.650 dan Cophenetic Correlation Coefficient 0.647, menunjukkan klaster yang kohesif dan representatif. Implementasi berbasis web menggunakan Django, D3.js, dan Chart.js menyediakan visualisasi interaktif dan rekomendasi portofolio. Kontribusi meliputi pendekatan valuasi fundamental baru, model clustering adaptif, dan alat analitik praktis untuk investor, dengan potensi ekspansi ke metrik multidimensi di masa depan.

Kata kunci: Hierarchical Clustering, FDMC, Narasi Crypto, Data Mining, Web Django.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.



1. Pendahuluan

Dalam dekade terakhir, pasar crypto telah berevolusi dari eksperimen teknologi niche menjadi kelas aset global dengan dinamika kompleks. Kapitalisasi pasar total crypto pernah melampaui \$4T, didorong oleh adopsi institusional, ekspansi infrastruktur blockchain, dan diversifikasi kasus penggunaan [1]. Pertumbuhan ini tidak hanya tercermin dari kenaikan harga aset utama seperti Bitcoin atau Ethereum, tetapi juga dari munculnya ratusan aset baru yang terkelompok ke berbagai "narasi" seperti Decentralized Finance (DeFi), Game Finance (GameFi), Artificial Intelligence (AI), Real-World Assets (RWA), hingga Memecoin (MEME). Setiap narasi memiliki karakteristik ekonomi unik, membentuk ekosistem heterogen [2].

Heterogenitas ini menimbulkan tantangan bagi investor, analis, dan pembuat kebijakan [3]. Di pasar keuangan tradisional, kerangka seperti Global Industry Classification Standard (GICS) memudahkan valuasi relatif dan alokasi portofolio. Sebaliknya, pasar crypto belum memiliki sistem kategorisasi standar [4]. Pengelompokan narasi masih cair, dipengaruhi sentimen jangka pendek dan hype media, menyebabkan investor ritel terjebak euforia tanpa memahami posisi fundamental proyek. Investor sering kesulitan membedakan aset dengan monetary premium (seperti Layer 1) dari yang berbasis utilitas atau arus kas. Contohnya, selama bull market 2021, narasi seperti NFT melonjak karena hype, tapi kemudian koreksi tajam menunjukkan kurangnya fondasi valuasi jangka panjang.

Upaya akademik dan industri semakin mengadopsi Machine Learning, khususnya clustering, untuk segmentasi pasar crypto [5]. Studi sebelumnya fokus pada karakteristik deret waktu seperti volatilitas harga atau korelasi [6]. Misalnya, [4] menggabungkan K-Means dan Hierarchical Clustering dengan fitur kualitatif dari web-scraping, tetapi tetap bergantung pada volatilitas. Pendekatan lain menggunakan Hierarchical Clustering orde kedua untuk memetakan struktur korelasi [7], atau DBSCAN untuk deteksi anomali harga Bitcoin [8]. Studi mutakhir [9] mengelompokkan narasi berdasarkan volume perdagangan, meskipun volume fluktuatif dan dimanipulasi.

Keterbatasan umum adalah ketergantungan pada metrik harga historis yang reaktif, dipengaruhi sentimen jangka pendek, dan tidak mencerminkan valuasi fundamental jangka panjang [10]. Penelitian juga kurang mempertimbangkan granularitas narasi sebagai unit analisis utama di ekosistem crypto modern, serta mengabaikan distribusi power-law pada data kapitalisasi [11]. Narasi crypto masih dianggap sebagai 'masalah terbuka' [6], dengan tantangan data blockchain seperti volume masif dan tidak terstruktur memperburuk kompleksitas. Sebagai ilustrasi, meskipun pendekatan machine learning telah terbukti efektif dalam manajemen risiko finansial pasar crypto, pendekatan tersebut belum diadaptasi untuk membedah struktur narasi berbasis valuasi fundamental [12].

Berdasarkan tinjauan tersebut, terdapat tiga kesenjangan utama: (1) Kurangnya perspektif valuasi fundamental melalui Fully Diluted Market Cap (FDMC), metrik kunci tokenomics yang mengukur risiko inflasi dan potensi dilusi [12]; (2) Pengabaian distribusi power-law, menyebabkan kluster bias tanpa transformasi seperti log-transform; (3) Simplifikasi pemetaan pasar menjadi "Bitcoin vs Altcoin", menghilangkan hubungan antar-narasi.

Untuk mengisi kesenjangan ini, penelitian ini menggunakan total FDMC dari 26 narasi crypto sebagai fitur dasar dalam Hierarchical Clustering, mengalihkan fokus ke valuasi fundamental jangka panjang. Kami menerapkan log-transformasi dan normalisasi Z-Score untuk mengurangi distorsi distribusi miring, menghasilkan kluster stabil. Algoritma Agglomerative Hierarchical Clustering dengan Ward Linkage mengungkap hierarki pasar secara organik, diinterpretasikan sebagai kategori risiko (Bluechip, Growth, Speculative) berdasarkan statistik kluster. Kontribusi ini menyediakan kerangka pengelompokan narasi berbasis valuasi, berguna sebagai alat keputusan bagi investor dan peneliti. Secara praktis, hasil diimplementasikan dalam dashboard analitik web interaktif untuk membantu analisis dan alokasi portofolio, meningkatkan pemahaman dinamika pasar crypto.

2. Metode Penelitian

Penelitian ini dirancang dengan pendekatan kuantitatif eksploratif menggunakan metode hybrid yang mengintegrasikan alur Waterfall (SDLC), justifikasi bisnis CRISP-DM, dan pemrosesan teknis Knowledge Discovery in Databases (KDD) untuk memastikan struktur dan reproduktibilitas. Tahapan utamanya mencakup pengumpulan data, pra-pemrosesan dan transformasi, pemodelan clustering, serta evaluasi dan validasi.

2.1. Pengumpulan Data dan Pengelompokan Narasi

Data primer diperoleh secara terprogram melalui API publik CoinGecko, khususnya endpoint /coins/markets. CoinGecko dipilih karena independen dari exchange tertentu dan memiliki cakupan data luas. Sampel terdiri dari 1.009 aset crypto teratas berdasarkan kapitalisasi pasar global pada snapshot pengambilan data, mewakili >90% total valuasi pasar. Atribut yang dikumpulkan meliputi nama aset, simbol, FDMC, dan

tag kategori dari platform. Contoh data awal: Bitcoin (FDMC \$2T) dikelompokkan ke L1, sementara Dogecoin ke MEME.

Tantangan utama adalah inkonsistensi pelabelan, di mana satu aset bisa memiliki multiple tag (misalnya, "DeFi", "Binance Smart Chain Ecosystem", "Yield Farming"). Untuk analisis kluster, setiap aset direduksi menjadi satu Narasi Utama melalui pembersihan dan pengelompokan manual ketat: analisis kualitatif whitepaper, situs resmi proyek, dan cross-check dengan CoinGecko serta CoinMarketCap. Dalam proses penentuan Narasi Utama, kriteria utama yang digunakan adalah fungsi inti (core utility) dan peran dominan aset dalam ekosistem, bukan sekadar tag yang paling banyak muncul. Jika suatu aset memiliki lebih dari satu fungsi, maka pengelompokan diarahkan pada use-case yang paling merepresentasikan sumber nilai utamanya. Untuk meminimalkan subjektivitas, setiap keputusan pengelompokan dilakukan melalui pembacaan dokumentasi resmi proyek dan divalidasi silang dengan dua sumber eksternal independen (CoinGecko dan CoinMarketCap). Pendekatan ini bertujuan mereduksi noise pelabelan yang umum ditemukan pada platform agregator, sehingga narasi yang terbentuk lebih konsisten secara semantik. Proses ini memastikan akurasi, misalnya, mengelompokkan Uniswap sebagai DEX meski memiliki tag DeFi.

Hasilnya, aset dipetakan ke 26 narasi unik, seperti Layer 1 (L1), Layer 2 (L2), DeFi, GameFi, AI, RWA, MEME, DePIN, dll. Data aset individual diagregasi per narasi, dengan total FDMC dihitung sebagai $\sum FDMC_i$ untuk setiap narasi i . Agregasi FDMC per narasi dipilih untuk menangkap perspektif valuasi makro, di mana narasi diperlakukan sebagai unit analisis ekonomi yang merepresentasikan ekspektasi pasar terhadap kelas teknologi atau use-case tertentu, bukan sekadar performa aset individual. Pendekatan ini mempertahankan pengaruh proyek flagship sebagai representasi valuasi narasi, sambil memberikan estimasi valuasi jangka panjang realistis dengan mempertimbangkan total pasokan token maksimum.

2.2. Pra-pemrosesan Data dan Transformasi Matematis

Langkah ini mengatasi distribusi data keuangan crypto yang miring. Analisis deskriptif awal menunjukkan skewness positif ekstrem dan kurtosis tinggi, dengan narasi L1 mendominasi (\$2.76T) dibanding narasi niche (ratusan juta dolar), menciptakan disparitas magnitudo hingga 10^4 . Contoh: Skewness FDMC mentah >5 , menandakan distribusi tail-heavy. Untuk mitigasi, diterapkan transformasi logaritmik basis-10:

$$x_{\log} = \log_{10}(FDMC) \quad (1)$$

Dimana x adalah narasi. Yang memampatkan rentang nilai ekstrem, menstabilkan varians, dan membuat jarak scale-invariant. Ini mencegah "Dominasi Outlier L1" di ruang Euclidean, di mana L1 terisolasi sementara narasi lain terkompres. Transformasi juga menangani outlier dari kesalahan API atau data invalid, misalnya FDMC nol diganti dengan nilai minimum. Selanjutnya, data dinormalisasi dengan Z-Score:

$$Z = \frac{x_{\log} - \mu}{\sigma} \quad (2)$$

Dimana μ adalah rata-rata dan σ adalah deviasi standar dari seluruh nilai \log -FDMC. Meski analisis univariat, normalisasi ini memastikan stabilitas numerik dan interpretasi statistik konsisten. Dataset Z menjadi input final untuk clustering.

2.3. Algoritma Hierarchical Clustering

Penelitian ini menggunakan Agglomerative Hierarchical Clustering (AHC), sebuah metode bottom-up di mana setiap narasi awalnya merupakan sebuah singleton, lalu digabung secara iteratif berdasarkan kemiripan hingga membentuk satu kluster tunggal. AHC dipilih karena mampu menghasilkan dendrogram yang memberikan visualisasi hierarki secara intuitif [5]. Ukuran jarak yang digunakan adalah Euclidean Distance, yang dihitung menggunakan persamaan:

$$d(i, j) = \sqrt{\sum_{p=1}^n (i_p - j_p)^2} \quad (3)$$

Dimana i dan j mewakili dua narasi yang saling dibandingkan, serta p adalah dimensi variabel valuasi yang diukur. Untuk metode penggabungan (linkage), digunakan Ward's Minimum Variance. Metode ini bekerja dengan meminimalkan within-cluster variance pada setiap penggabungan dengan mencari pasangan kluster yang menghasilkan kenaikan terkecil pada nilai Error Sum of Squares (ESS), yang dirumuskan sebagai:

$$ESS = \sum_{x \in C} (x - \bar{x})^2 \quad (4)$$

Dimana x adalah titik data (narasi) yang tergabung di dalam kluster C , dan \bar{x} adalah nilai rata-rata (centroid) dari kluster tersebut. Metode Ward secara efektif menghasilkan kluster yang lebih kompak, seimbang, dan sangat relevan untuk segmentasi valuasi pasar crypto, sekaligus mencegah terjadinya efek chaining atau pembentukan kluster yang tidak bermakna secara ekonomi.

2.4. Evaluasi dan Validasi Model

Jumlah kluster optimal (k) ditentukan dengan kombinasi heuristik dan metrik statistik [13]:

1. Silhouette Score: Mengukur kohesi (kemiripan dalam kluster) vs. separasi (perbedaan antar-kluster), rentang -1 hingga +1. Skor >0.5 menandakan struktur baik.
2. Cophenetic Correlation Coefficient (CCC): Mengukur kesetiaan dendrogram terhadap jarak asli, rumus cophenetic correlation. Nilai mendekati 1 untuk validasi hierarki menunjukkan representasi akurat,

Penggunaan Silhouette Score dan Cophenetic Correlation Coefficient secara bersamaan dimaksudkan untuk menyeimbangkan dua aspek utama Hierarchical Clustering, yaitu kualitas pemisahan kluster dan kesetiaan struktur hierarki terhadap jarak asli data. Silhouette menilai seberapa baik setiap narasi terasosiasi dengan klusternya dibanding kluster lain, sementara CCC memastikan bahwa dendrogram tidak mengalami distorsi signifikan akibat proses penggabungan bertahap. Metrik ini dievaluasi secara iteratif; misalnya, $k=5$ dipilih karena memberikan skor silhouette tertinggi (0,650) dibanding $k=4$ (0,561) maupun $k=6$ (0,646), sementara nilai CCC relatif konstan pada semua k . Nilai Silhouette sebesar 0,650 pada $k=5$ mengindikasikan struktur kluster yang kuat dan terpisah dengan baik, sedangkan CCC sebesar 0,647 menunjukkan bahwa hierarki yang terbentuk cukup representatif terhadap matriks jarak awal. Meskipun $k=6$ menghasilkan skor yang mendekati, $k=5$ dipilih karena memberikan keseimbangan terbaik antara kualitas kluster dan interpretabilitas ekonomi struktur pasar.

Tabel 1: Metrik Evaluasi Clustering

K	Silhouette	CCC
2.000	0.575	0.647
3.000	0.582	0.647
4.000	0.561	0.647
5.000	0.650	0.647
6.000	0.646	0.647
7.000	0.623	0.647
8.000	0.585	0.647
9.000	0.581	0.647

2.5. Kategorisasi Kluster Berbasis Aturan

Setelah clustering, proses post-processing menetapkan label risiko. Total FDMC per kluster dihitung, lalu diterapkan aturan:

1. Bluechip: $FDMC > \$1T$
2. Growth: $\$10B < FDMC \leq \$1T$
3. Speculative: $FDMC \leq \$10B$

Kategori ini memberikan interpretasi kualitatif untuk rekomendasi dashboard, memisahkan hasil algoritma dari labeling manual untuk fleksibilitas.

2.6. Pengembangan Dashboard Web

Hasil hierarchical clustering diimplementasikan dalam sebuah dashboard web interaktif yang dirancang sebagai platform analitik untuk mengeksplorasi struktur pasar narasi crypto berbasis metrik FDMC. Backend dibangun menggunakan Django (Python) untuk pemrosesan data, normalisasi, dan penyajian hasil melalui API, sementara frontend memanfaatkan D3.js dan Chart.js untuk menampilkan dendrogram, treemap, scatter plot, pi chart, dan bar chart secara dinamis. Dashboard ini dilengkapi fitur insight otomatis yang merangkum karakteristik kluster berdasarkan dominasi FDMC dan kategori risiko (Bluechip, Growth, Speculative), serta modul contoh alokasi portofolio berbasis pendekatan Core-Satellite (misalnya 75% Bluechip, 15% Growth, 10% Speculative), sehingga hasil clustering tidak hanya divisualisasikan, tetapi juga diinterpretasikan secara praktis untuk membantu pengguna memahami dinamika pasar dan menyusun struktur portofolio berbasis kluster secara lebih terukur.

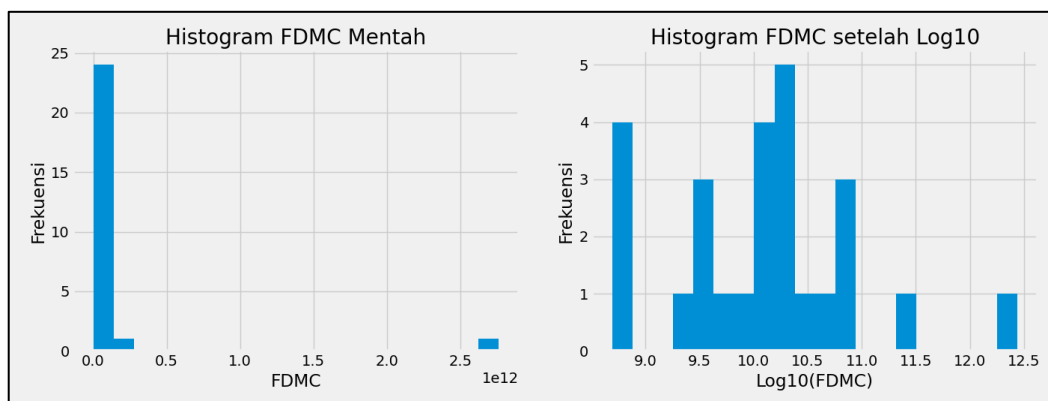
3. Hasil dan Pembahasan

Bagian ini menyajikan temuan empiris dari analisis data, mulai dari karakteristik statistik data mentah, dampak transformasi logaritmik, hingga interpretasi struktur kluster dan implikasinya bagi pemahaman pasar crypto.

3.1. Karakteristik Data dan Anomali "Tanpa Log"

Data awal dari 26 narasi menunjukkan disparitas valuasi ekstrem. Tabel statistik deskriptif pada data mentah mengonfirmasi dominasi narasi Layer 1 (L1) dengan FDMC \$2.76T (didorong Bitcoin, Ethereum, Solana), sementara Memecoin \$66.38B (2% dari L1) dan narasi niche seperti GambleFi ratusan juta dolar. Nilai skewness 5 dan kurtosis 25 menandakan distribusi ekstrem, di mana 80% valuasi terkonsentrasi pada 20% narasi (efek Pareto).

Eksperimen awal tanpa transformasi logaritmik menghasilkan kluster bias: algoritma mengisolasi Payments (PAY) dan L1 karena outlier seperti XRP (\$217B), menciptakan distorsi di ruang Euclidean. PAY terpisah dari narasi aplikasi lain meski fungsional mirip, menegaskan kebutuhan transformasi log untuk mengatasi distribusi power-law [11]. Tanpa itu, dendrogram gagal menangkap hierarki pasar antara infrastruktur dan aplikasi, dengan 90% narasi terkumpul dalam satu kluster padat.



Gambar 1. Perbandingan Distribusi FDMC

Tabel 2: Statistik Deskriptif Data FDMC

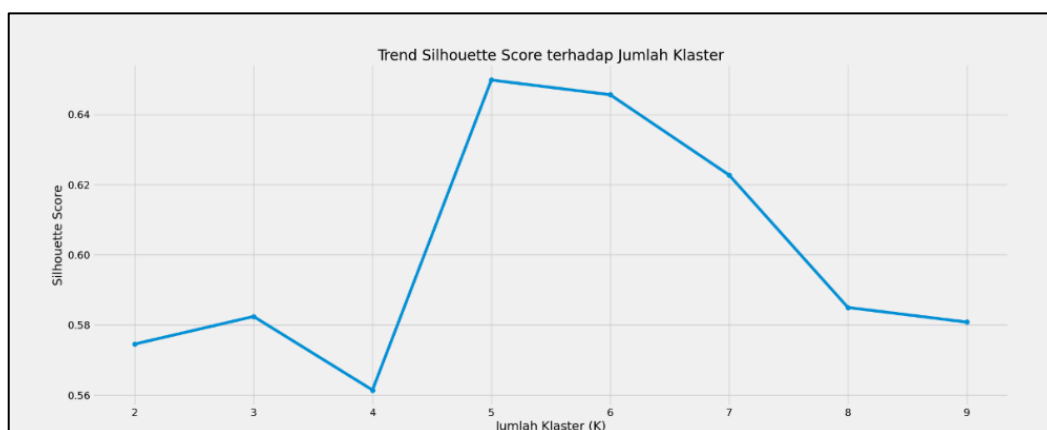
Klaster	FDMC
count	26.00
mean	132,296,951,208.81
std	537,685,716,131.83
min	493,939,926.00
25%	3,808,495,162.50
50%	14,073,124,108.50
75%	27,954,260,242.50
max	2,758,706,450,917.00
skewness	5.04
kurtosis	25.57

3.2. Struktur Kluster dengan Transformasi Logaritma

Setelah transformasi log-10 dan standarisasi Z-Score, struktur data lebih teratur dan outlier ekstrem diredam. Tren Silhouette Score (Gambar 2) memuncak pada k=5, mengindikasikan kualitas pemisahan kluster paling optimal, sehingga k=5 dipilih sebagai jumlah yang representatif.

Validasi: Silhouette Score 0.650 (indikasi kohesi tinggi dan separasi jelas); CCC 0.647 (representasi akurat jarak asli).

Pemilihan k=5 juga didukung aspek interpretabilitas ekonomi. Pada k lebih kecil (k<5), narasi dengan skala valuasi berbeda signifikan masih tergabung, berpotensi menyamarkan perbedaan risiko pasar. Sebaliknya, k≥6 memfragmentasi struktur menjadi kelompok kecil yang sulit diberi makna ekonomi konsisten. Transformasi logaritmik berperan menghasilkan dendrogram yang seimbang, mencegah narasi besar seperti L1 mendominasi penggabungan, sehingga kluster yang terbentuk merepresentasikan hierarki valuasi pasar secara akurat.



Gambar 2. Visualisasi Pergerakan Silhouette Score Terhadap Kluster

Tabel 3: Daftar Narasi dan FDMC

Narrative	FDMC	Log FDMC	Cluster
AI (Artificial Intelligence)	29,348,010,770.00	10.47	3
BTC Eco (Bitcoin Ecosystem)	3,021,501,856.00	9.48	2
CEX (Centralized Exchange)	62,934,302,653.00	10.80	4
DAO (Decentralized Autonomous Organization)	21,096,070,722.00	10.32	3
DEX (Decentralized Exchange)	59,169,767,608.00	10.77	4
DID (Decentralized Identity)	13,777,422,936.00	10.14	3
DeFi (Decentralized Finance)	46,473,718,252.00	10.67	4
DePIN (Decentralized Physical Infrastructure Network)	23,773,008,660.00	10.38	3
FTO (Fan Token)	496,030,235.00	8.70	1
GambleFi (Gamble Finance)	609,939,280.00	8.79	1
GameFi (Game Finance)	14,076,538,667.00	10.15	3
L1 (Layer 1)	2,758,706,450,917.00	12.44	5
L2 (Layer 2)	21,055,462,445.00	10.32	3
LPD (Launchpad)	735,365,020.00	8.87	1
LST (Liquid Staking)	9,081,236,629.00	9.96	3
MEME (Memecoin)	66,377,878,307.00	10.82	4
MET (Metaverse)	3,847,117,560.00	9.59	2
NFT (Non-Fungible Token)	4,954,319,317.00	9.69	2
ORC (Oracle)	17,778,558,935.00	10.25	3
PAY (Payments)	233,265,116,499.00	11.37	4
PRC (Privacy Coin)	10,695,911,341.00	10.03	3
RWA (Real World Assets)	18,153,491,229.00	10.26	3
SEC (Cybersecurity)	493,939,926.00	8.69	1
SocialFi (Social Finance)	3,795,621,030.00	9.58	2
W3I (Web3 Infrastructure)	1,934,241,085.00	9.29	2
XCH (Cross-Chain)	14,069,709,550.00	10.15	3

3.3. Interpretasi Profil Klaster: Hierarki Valuasi Pasar

Tabel 4. Hasil Pengelompokan Narasi Crypto Berdasarkan FDMC

Klaster	Narasi Dominan	Anggota Narasi	Total FDMC	Kelompok Risiko
Klaster 5	L1	L1	\$2.76T	Bluechip
Klaster 4	PAY	PAY, MEME, CEX, DEX, DeFi	\$468.22B	Growth
Klaster 3	AI	AI, DePIN, DAO, L2, RWA, ORC, GameFi, XCH, DID, PRC, LST	\$192.91B	Growth
Klaster 2	NFT	NFT, MET, SocialFi, BTC Eco, W3I	\$17.55B	Speculative
Klaster 1	LPD	LPD, GambleFi, FTO, SEC	\$2.34B	Speculative

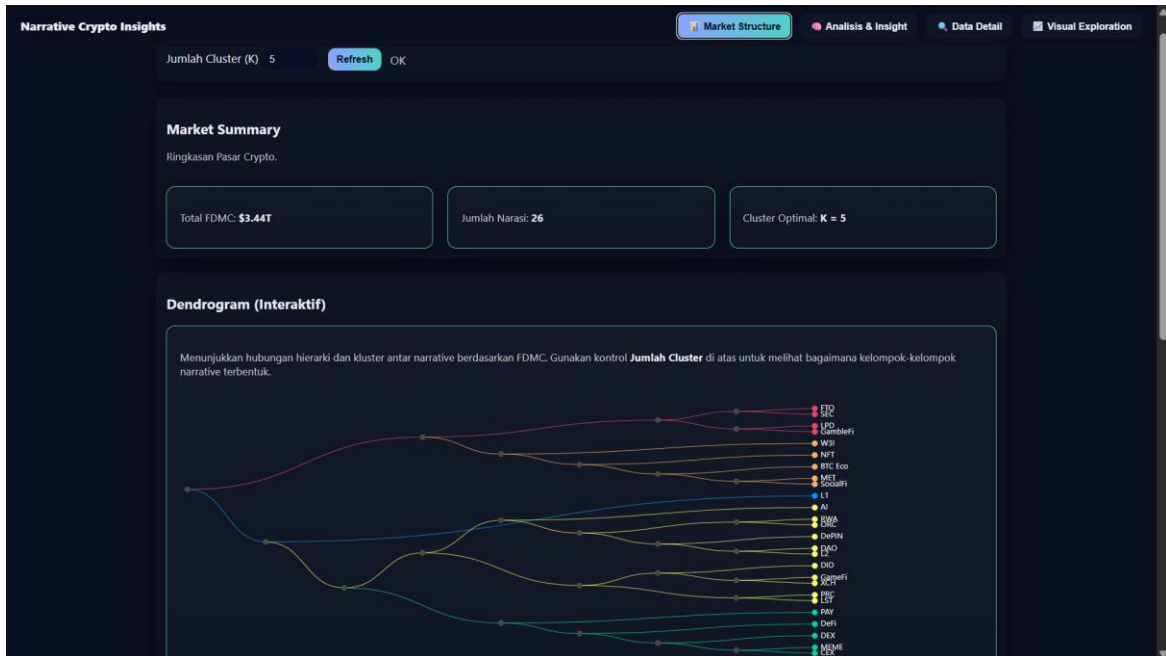
Analisis hierarchical clustering menghasilkan lima klaster yang merepresentasikan hierarki valuasi pasar dan tingkat risiko narasi crypto berdasarkan FDMC, sebagai berikut:

1. Klaster 5 – Bluechip (Very High Tier): Klaster tertinggi ini hanya terdiri dari satu narasi, yaitu Layer 1 (L1), dengan total FDMC sebesar \$2.76T (80.20% dari total pasar). Pada dendrogram, L1 terpisah paling awal dari narasi lain, menegaskan adanya dikotomi fundamental antara infrastruktur dasar dan aplikasi di atasnya. Temuan ini menunjukkan bahwa token L1 dipersepsikan pasar sebagai fondasi utama ekosistem crypto dengan tingkat kematangan dan stabilitas tertinggi, sehingga layak dikelompokkan sebagai Bluechip.
2. Klaster 4 – Growth (High Tier): Klaster ini dipimpin oleh narasi Payments (PAY) dan mencakup PAY, MEME, CEX, DEX, dan DeFi, dengan total FDMC sebesar \$468.22B (13.61%). Klaster ini merepresentasikan sektor aplikasi utama dengan valuasi besar dan peran sentral dalam aktivitas ekonomi crypto. Bergabungnya narasi utilitas tinggi seperti DeFi dan CEX dengan narasi berbasis atensi seperti MEME menunjukkan bahwa, meskipun berbeda secara karakter fundamental, pasar memberikan skala valuasi yang sebanding terhadap keduanya, sehingga seluruhnya diperlakukan sebagai sektor dengan potensi pertumbuhan tinggi di bawah dominasi L1.
3. Klaster 3 – Growth (Mid Tier): Klaster ini dipimpin oleh Artificial Intelligence (AI) dan terdiri dari 11 narasi, yaitu AI, DePIN, DAO, L2, RWA, ORC, GameFi, XCH, DID, PRC, dan LST dengan total FDMC \$192.91B (5.61%). Klaster ini mencerminkan sektor inovatif yang sedang berada pada fase pertumbuhan menengah. Narasi-narasi di dalamnya menunjukkan potensi perkembangan, namun secara valuasi belum mencapai skala klaster High Tier. Posisi Layer 2 (L2) dalam klaster ini juga menunjukkan bahwa meskipun bersifat infrastruktur, valuasi token L2 masih lebih dekat dengan karakteristik token aplikasi dibanding L1.
4. Klaster 2 – Speculative (Low Tier): Klaster ini dipimpin oleh Non-Fungible Token (NFT) dan mencakup NFT, MET, SocialFi, BTC Eco, dan W3I, dengan total FDMC sebesar \$17.55B (0.51%). Klaster ini merepresentasikan sektor dengan valuasi relatif kecil dan tingkat ketidakpastian tinggi. Masuknya BTC Eco ke dalam klaster ini menunjukkan bahwa meskipun berasosiasi dengan ekosistem Bitcoin, proyek-proyek di dalamnya secara agregat masih memiliki skala valuasi terbatas, sehingga dipandang pasar sebagai lapisan aplikasi yang belum matang dan karenanya lebih tepat dikategorikan sebagai Spekulatif.
5. Klaster 1 – Speculative (Very Low Tier): Klaster terendah dipimpin oleh Launchpad (LPD) dan mencakup LPD, GambleFi, FTO, dan SEC, dengan total FDMC hanya \$2.34B (0.06%). Klaster ini menunjukkan kelompok narasi dengan valuasi paling kecil dan risiko tertinggi di antara seluruh klaster. Posisi ini mencerminkan keterbatasan skala pasar dan tingkat kepercayaan yang masih rendah dibanding klaster lainnya.

Secara keseluruhan, segmentasi klaster menunjukkan pola yang konsisten dengan hierarki FDMC, L1 berada pada level Bluechip, dua klaster berikutnya membentuk kelompok Growth, sementara dua klaster terbawah terkonsentrasi sebagai Speculative. Hal ini menegaskan bahwa metode Hierarchical Clustering berhasil menangkap struktur kematangan dan tingkat risiko pasar crypto secara sistematis berdasarkan valuasi fundamental FDMC. Struktur lima klaster ini memperlihatkan adanya stratifikasi valuasi yang tajam dalam ekosistem crypto, di mana satu narasi infrastruktur utama (L1) menyerap mayoritas nilai pasar, sementara sebagian besar narasi lain tersebar pada tingkat valuasi menengah hingga rendah. Kondisi ini mencerminkan karakter pasar yang sangat terpusat pada fondasi jaringan, dengan sektor aplikasi membentuk lapisan pertumbuhan di atasnya. Dua klaster Growth berperan sebagai zona transisi antara stabilitas Bluechip dan risiko Speculative, menunjukkan bahwa pasar membedakan tingkat kematangan antar narasi aplikasi meskipun sama-sama berada di luar L1. Di sisi lain, klaster Speculative membentuk ekor distribusi (long-tail) dengan jumlah narasi cukup banyak namun kontribusi valuasi yang kecil, yang menandakan bahwa sebagian besar inovasi berada pada fase awal dan belum memperoleh kepercayaan pasar yang luas. Pola ini menegaskan bahwa risiko dan potensi imbal hasil dalam pasar crypto terdistribusi secara hierarkis seiring dengan skala valuasi fundamentalnya.

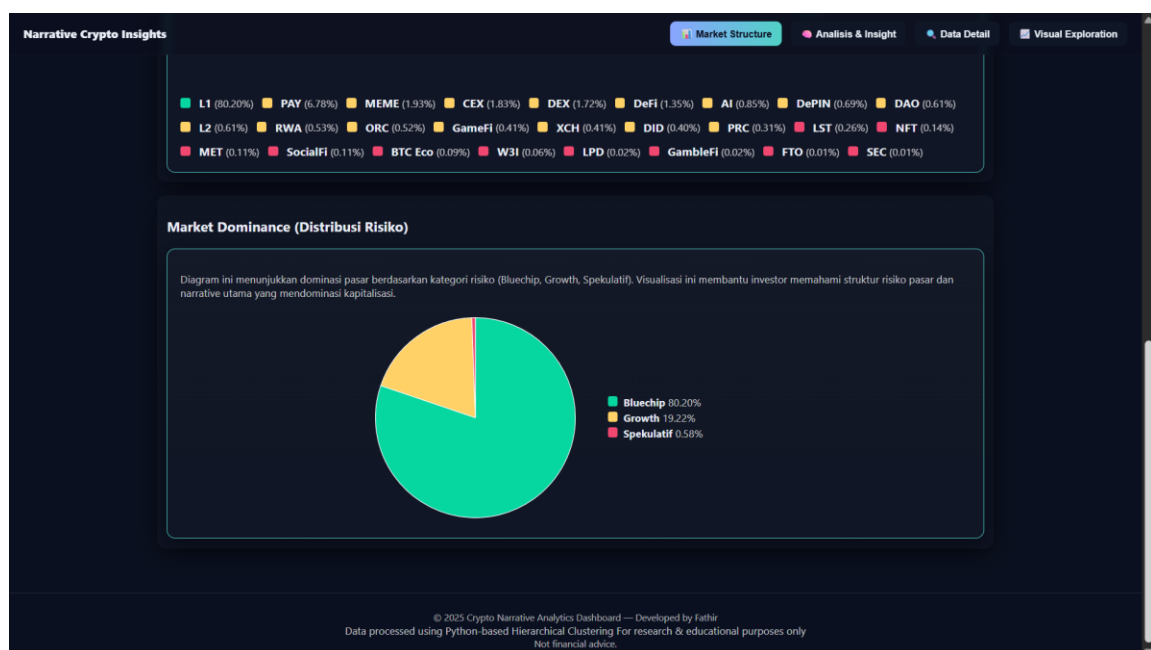
3.4. Visualisasi dan Implikasi Strategis

Visualisasi hasil clustering pada dashboard web interaktif berperan penting dalam membantu pengguna memahami struktur pasar crypto yang kompleks secara intuitif. Melalui integrasi dendrogram, grafik distribusi, serta fitur insight otomatis, hubungan antar narasi, hierarki valuasi, dan implikasi risiko dapat diamati secara langsung berdasarkan data FDMC yang digunakan dalam penelitian ini.



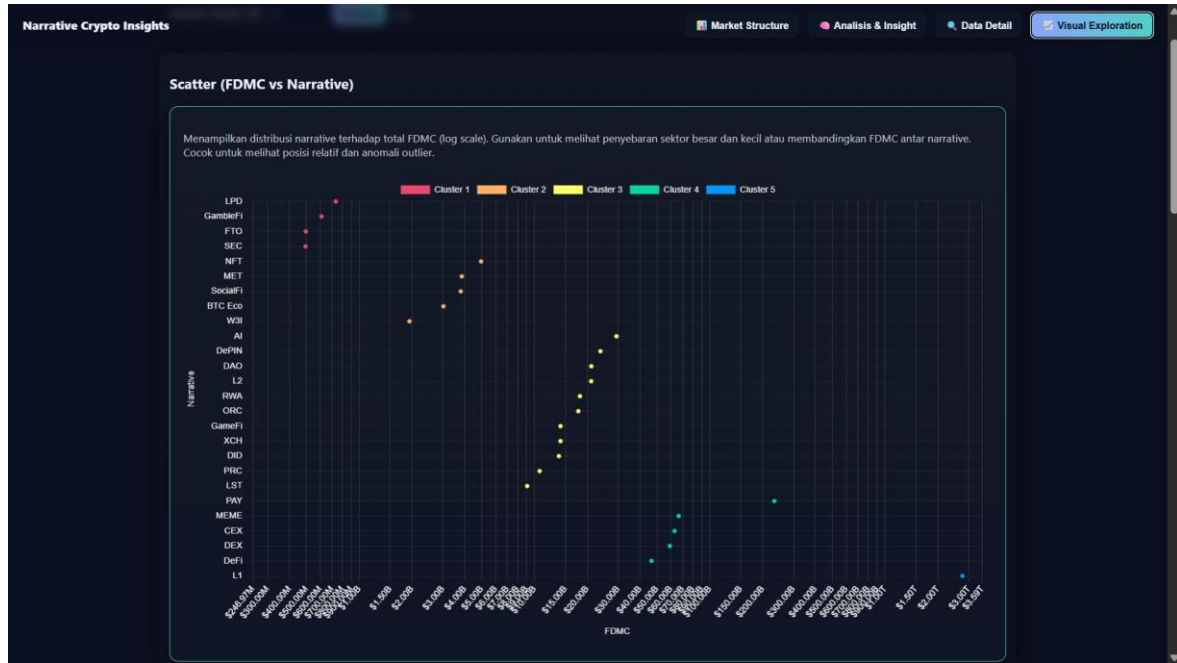
Gambar 3. Dendrogram Interaktif

Dendrogram menjadi representasi utama dari proses hierarchical clustering. Visual ini menunjukkan bagaimana narasi-narasi crypto digabungkan secara bertahap berdasarkan kemiripan FDMC. Pada tingkat awal terlihat pemisahan yang jelas antara Layer 1 (L1) dan narasi lainnya, yang menegaskan perbedaan fundamental antara sektor infrastruktur dasar dan sektor aplikasi. Selanjutnya, cabang-cabang dendrogram membentuk kelompok Growth dan Speculative, mencerminkan stratifikasi pasar berdasarkan skala valuasi dan tingkat kematangan. Struktur ini membantu pengguna memahami bahwa perbedaan utama di pasar tidak hanya terletak pada jenis narasi, tetapi juga pada level valuasi yang mendasarinya.



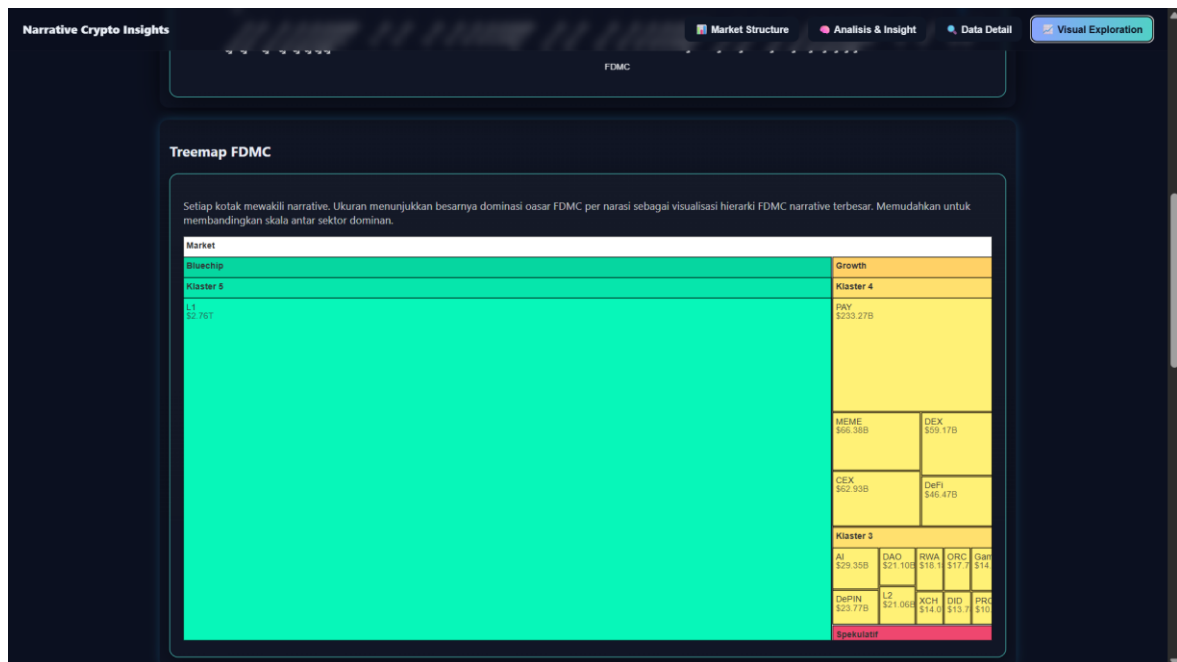
Gambar 4. Pie Chart

Pie chart menampilkan distribusi total FDMC berdasarkan kategori risiko kluster. Visual ini menunjukkan dominasi kategori Bluechip (L1) yang mencakup sekitar 80.20% dari total FDMC pasar, diikuti oleh kategori Growth, sementara kategori Speculative hanya mengambil porsi yang sangat kecil. Representasi ini memperjelas bahwa sebagian besar valuasi pasar crypto terkonsentrasi pada sedikit narasi berkapitalisasi sangat besar, sedangkan sebagian besar narasi lainnya berada pada ekor distribusi dengan nilai yang jauh lebih kecil.



Gambar 5. Scatter Plot

Scatter plot dengan skala logaritmik memetakan narasi terhadap nilai FDMC dan diwarnai berdasarkan kluster. Visual ini mengakomodasi rentang FDMC yang sangat lebar, dari kluster Speculative hingga Bluechip. Penyebaran titik memperlihatkan pemisahan antar kluster secara jelas serta keberadaan outlier bernilai sangat tinggi seperti L1, sehingga membantu mengidentifikasi posisi relatif setiap narasi dalam lanskap pasar secara menyeluruh.



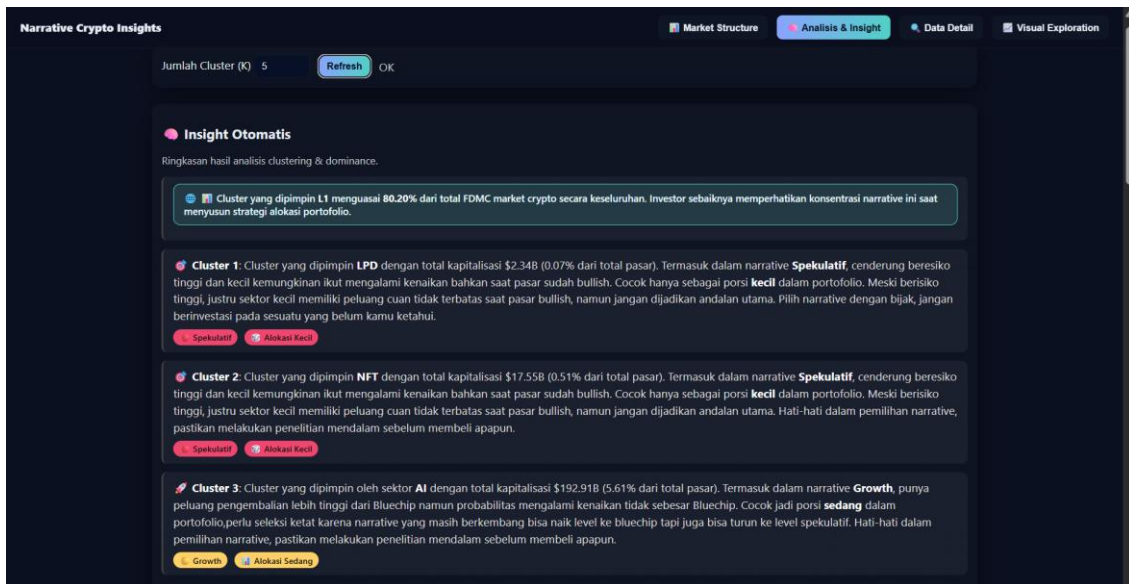
Gambar 6. Treemap

Treemap memetakan FDMC setiap narasi dalam konteks kategori klasternya. Luas setiap kotak merepresentasikan besarnya FDMC, sehingga memudahkan identifikasi narasi dominan di dalam masing-masing klaster. Klaster Bluechip didominasi oleh L1, klaster Growth terbagi ke dalam narasi seperti PAY, DeFi, CEX, DEX, MEME, serta AI, RWA, DePIN, dan L2 pada tier menengah, sedangkan klaster Speculative tampak sebagai kotak yang lebih kecil menegaskan skala valuasi yang jauh lebih rendah. Visual ini membantu pengguna melihat struktur internal tiap klaster secara lebih granular.



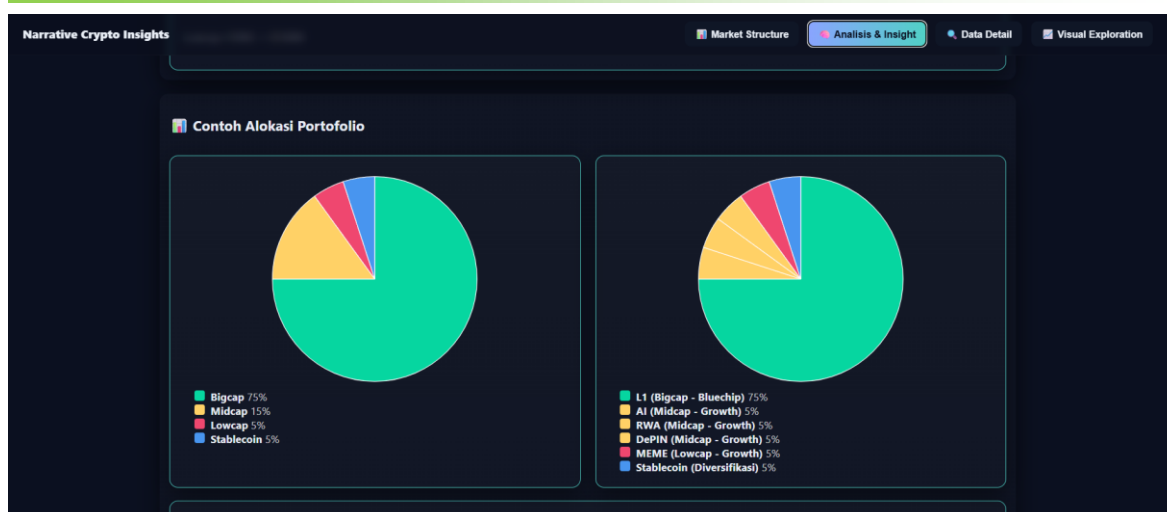
Gambar 7. Bar chart

Bar chart menampilkan perbandingan FDMC antar narasi utama dengan menghilangkan dominasi L1 dan PAY agar distribusi pada sektor aplikasi lebih terlihat jelas. Grafik ini menyoroti perbedaan skala antar narasi di luar Bluechip dan memperlihatkan pola long-tail pada distribusi valuasi pasar crypto, di mana hanya beberapa narasi Growth memiliki FDMC relatif besar dibanding mayoritas narasi lain di tier menengah dan bawah.



Gambar 8. Insight Otomatis

Dashboard menyediakan ringkasan insight otomatis yang merangkum karakteristik tiap klaster, seperti kategori risiko, dominasi FDMC, serta implikasi umum terhadap alokasi portofolio. Insight ini dihasilkan langsung dari hasil clustering dan berfungsi sebagai interpretasi awal yang membantu pengguna memahami makna data tanpa harus menafsirkan seluruh visual secara manual.



Gambar 9. Contoh Alokasi Portofolio

Sebagai ilustrasi, dashboard menampilkan contoh alokasi portofolio berbasis kategori kluster yang membagi porsi ke dalam Bluechip, Growth, dan Speculative. Visual ini tidak dimaksudkan sebagai rekomendasi investasi, melainkan sebagai demonstrasi bagaimana hasil clustering dapat digunakan untuk menyusun struktur portofolio berdasarkan tingkat risiko dan hierarki valuasi yang dihasilkan model. Implikasi ini selaras dengan pendekatan portofolio berbasis kluster dalam literatur, di mana teknik clustering dimanfaatkan untuk segmentasi aset dan pengelolaan risiko secara sistematis [14],[15]. Dalam konteks penelitian ini, hasil hierarchical clustering mengarah pada kerangka Core–Satellite, yaitu:

1. Core (Bluechip) sebagai fondasi portofolio pada kluster dengan valuasi tertinggi dan kematangan pasar paling mapan.
2. Satellite (Growth) sebagai diversifikasi pada kluster dengan potensi pertumbuhan menengah, dan
3. Satellite (Speculative) sebagai alokasi kecil dan terukur pada kluster berisiko tinggi dengan valuasi fundamental lebih rendah.

Pendekatan ini menegaskan peran clustering sebagai alat bantu dalam menyusun portofolio berbasis struktur pasar, bukan sekadar pemilihan aset individual, sebagaimana ditunjukkan dalam studi optimasi portofolio crypto berbasis kluster [14], [15].

Secara keseluruhan, rangkaian visualisasi ini memperkuat hasil hierarchical clustering dengan menyajikan struktur pasar crypto dalam bentuk yang mudah dipahami. Integrasi dendrogram, grafik distribusi, dan insight otomatis memungkinkan pengguna untuk:

1. mengidentifikasi dominasi dan ketimpangan valuasi pasar,
2. memahami posisi relatif setiap narasi dalam hierarki kluster, serta
3. mengeksplorasi implikasi risiko secara lebih sistematis.

Dengan demikian, dashboard berfungsi sebagai alat bantu analisis yang menjembatani hasil komputasi clustering dengan interpretasi praktis berbasis visual, sekaligus merefleksikan struktur pasar yang terbentuk dari data FDMC tanpa bergantung pada asumsi atau informasi eksternal di luar ruang lingkup penelitian ini.

4. Kesimpulan

Penelitian ini berhasil mengembangkan kerangka hierarchical clustering untuk memetakan struktur pasar crypto berdasarkan metrik valuasi fundamental FDMC. Penerapan transformasi logaritmik terbukti krusial dalam menangani distribusi data yang bersifat power-law, sehingga menghasilkan kluster yang lebih representatif dan stabil, dengan validasi yang baik melalui Silhouette Score 0.650 dan Cophenetic Correlation Coefficient 0.647. Hasil analisis menunjukkan bahwa pasar crypto tersusun secara hierarkis dalam tingkatan valuasi yang mencerminkan stratifikasi alami, dari narasi Bluechip hingga Speculative, yang dapat dirangkul ke dalam tiga kategori risiko utama: Bluechip, Growth, dan Speculative. Kontribusi utama penelitian ini meliputi penggunaan FDMC sebagai indikator valuasi fundamental, pengembangan model kluster untuk menangkap tingkat kematangan dan risiko pasar secara sistematis, serta implementasi dashboard analitik interaktif untuk memvisualisasikan hasil dan mendukung pemahaman struktur pasar. Dengan demikian, penelitian ini memberikan dasar metodologis dan praktis bagi analisis narasi crypto berbasis valuasi dan membuka peluang pengembangan lanjutan melalui penambahan metrik lain.

5. Daftar Pustaka

- [1] S. Dong, K. Abbas, M. Li, and J. Kamruzzaman, "Blockchain technology and application: an overview," *PeerJ Computer Science*, vol. 9, 2023. doi: 10.7717/peerj-cs.1705.
- [2] A. T. Aspembitova, L. Feng, and L. Y. Chew, "Behavioral structure of users in cryptocurrency market," *PLoS ONE*, vol. 16, no. 1, 2021. doi: 10.1371/journal.pone.0242600.
- [3] R. Majumder, "Is Crypto Tradeable? - A Perspective From Weak Form Efficiency," 2021. doi: 10.13140/RG.2.2.32933.42721.
- [4] B. Mahdavi-Damghani et al., "Cryptocurrency sectorisation through clustering and webscraping: Application to Systematic Trading," 2021. doi: 10.13140/RG.2.2.13136.84488.
- [5] T. Šťastný et al., "Clustering and Modelling of the Top 30 Cryptocurrency Prices Using Dynamic Time Warping and Machine Learning Methods," *Mathematics*, vol. 10, no. 19, 2022. doi: 10.3390/math10193672.
- [6] V. Azhmyakov, I. Shirokov, and L. A. G. Trujillo, "Advanced Statistical Analysis of the Predicted Volatility Levels in Crypto Markets," *Journal of Risk and Financial Management*, vol. 17, no. 7, 2024. doi: 10.3390/jrfm17070279.
- [7] H. Sadeqi, "A Second-Order Hierarchical Clustering of Cryptocurrencies," *Iranian Journal of Management Studies*, vol. 2022, no. 3, 2021.
- [8] A. Ş. Dokuz, M. Çelik, and A. Ecemiş, "DBSCAN Algoritması Kullanarak Bitcoin Fiyatlarında Anormallik Tespiti," *European Journal of Science and Technology*, pp. 436–443, 2020. doi: 10.31590/ejosat.araconf57.
- [9] M. F. Adha et al., "Analisis Hierarchical Clustering Pada Volume Trading Disetiap Narrative Crypto Menggunakan Data Dune Analytics Untuk Investasi Berbasis Data," *Jurnal Teknik Informatika Kaputama*, vol. 9, no. 1, 2025.
- [10] H. Sebastião and P. Godinho, "Forecasting and trading cryptocurrencies with machine learning under changing market conditions," *Financial Innovation*, vol. 7, no. 1, 2021. doi: 10.1186/s40854-020-00217-x.
- [11] M.-B. Lin, K. Khowaja, C. Y.-H. Chen, and W. K. Härdle, "Blockchain mechanism and distributional characteristics of cryptos," 2020. [Online]. Available: <http://arxiv.org/abs/2011.13240>
- [12] Z. Shahbazi and Y. C. Byun, "Machine Learning-Based Analysis of Cryptocurrency Market Financial Risk Management," *IEEE Access*, vol. 10, pp. 37848–37856, 2022. doi: 10.1109/ACCESS.2022.3162858.
- [13] L. Lorenzo and J. Arroyo, "Analysis of the cryptocurrency market using different prototype-based clustering techniques," *Financial Innovation*, vol. 8, no. 1, 2022. doi: 10.1186/s40854-021-00310-9.
- [14] T. Kocabıyık et al., "Cryptocurrency Portfolio Management: A Clustering-Based Association Approach," *Ekonomika*, vol. 103, no. 1, pp. 25–43, 2024. doi: 10.15388/Ekon.2024.103.1.2.
- [15] L. Lorenzo and J. Arroyo, "Online risk-based portfolio allocation on subsets of crypto assets applying a prototype-based clustering algorithm," *Financial Innovation*, vol. 9, no. 1, 2023. doi: 10.1186/s40854-022-00438-2.