Using Interactive Data Visualization for DeFi Market Analysis

Daria Pavlova
daria.pavlova@mps.si
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Inna Novalija
inna.koval@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

ABSTRACT

Decentralized Finance (DeFi) presents unique analytical challenges with its data-rich, volatile, and multi-dimensional ecosystem. Static reports struggle to convey short-term dynamics and cross-sectional structure simultaneously. We present a comprehensive Business Intelligence (BI) solution featuring an automated Extract-Transform-Load (ETL) pipeline and interactive Tableau dashboard. Our ETL architecture processes data from three Application Programming Interfaces (APIs)—CoinGecko, DeFiLlama, and DexScreener—through validation and transformation stages, achieving 45-second execution time. The dashboard integrates Key Performance Indicators (KPIs), Total Value Locked (TVL) time-series, market categories analysis, and top movers panel with synchronized filters. Performance evaluation demonstrates 85-99% reduction in analysis time compared to manual methods. Three real-world use cases validate practical applicability: narrative rotation detection (28% investment returns), risk concentration monitoring (15% drawdown reduction), and competitive benchmarking. Our approach bridges the gap between complex DeFi data and actionable insights without requiring technical expertise.

KEYWORDS

DeFi, Business Intelligence, Tableau, TVL, KPI dashboards, Interactive Visualization, ETL Pipeline, Data Mining, Cryptocurrency

1 INTRODUCTION

Decentralized Finance (DeFi) compresses high-frequency market activity—liquidity flows, incentive programs, and new protocol deployments—into datasets that change hourly. The ecosystem encompasses over 6,000 protocols managing billions in Total Value Locked (TVL), creating analytical complexity that traditional tools struggle to handle. Practitioners must simultaneously answer three critical questions: *How big is the market now?* (level KPIs), *How is it moving?* (time series), and *What drives the cross-section?* (categories, movers).

Interactive visualization reduces cognitive load and increases pattern salience relative to static tables [3, 7, 10, 11]. However, existing solutions present trade-offs: Dune Analytics requires Structured Query Language (SQL) expertise, Nansen charges $1,800 annually, while free alternatives like DeFiLlama offer limited visualization capabilities. Our goal is to demonstrate a compact, reproducible Business Intelligence workflow that democratizes DeFi analytics through automated data processing and intuitive visualization.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SiKDD 2025, October 6, 2025, Ljubljana, Slovenia
© 2025 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2025.sikdd.15

2 RELATED WORK

2.1 DeFi Analytics Landscape

Surveys of DeFi systems [12] highlight the centrality of TVL, market capitalization, and volume as monitoring signals. Public APIs from CoinGecko and DeFiLlama expose these aggregates for research and dashboards, processing millions of daily transactions into consumable metrics.[1]

Recent advances in artificial intelligence have opened new frontiers in DeFi analysis. Chen et al. [1] proposed ensemble machine learning approaches for detecting rug pulls and protocol vulnerabilities, achieving 87% accuracy using features extracted from on-chain data and social signals. Their Random Forest model combined with Long Short-Term Memory (LSTM) networks demonstrated AI's potential in risk assessment. However, these Machine Learning (ML) approaches require significant computational resources and technical expertise, creating barriers for non-technical analysts. Our solution complements these advanced techniques by providing immediate, interpretable insights through interactive visualization.

2.2 Business Intelligence and Visualization

Classic data warehouse and BI literature formalizes metrics and dimensional modeling for decision support [6]. Industry guidance positions interactive platforms such as Tableau among leading tools for exploratory analysis [4]. Visualization principles—overview first, zoom and filter, details-on-demand [8]—map directly to dashboard layout patterns [3, 9].

Studies of graphical perception [2, 5] explain why bars outperform pies for accurate comparisons, and why color semantics (green/red for gains/losses) aid preattentive detection [11]. We align with these findings in our chart choices and encodings.

3 SYSTEM ARCHITECTURE AND METHODOLOGY

3.1 ETL Pipeline Architecture

Our ETL pipeline implements a modular, fault-tolerant architecture processing data through five stages. The architecture follows a standard Extract-Transform-Load pattern with additional validation and quality checks at each stage.

Extract Layer: Three parallel API clients collect data from CoinGecko (200 tokens per page), DeFiLlama (6,000+ protocols), and DexScreener (100+ Decentralized Exchange pairs). Each client implements asynchronous Hypertext Transfer Protocol (HTTP) requests with exponential backoff (4-10 seconds) and retry logic (up to 5 attempts).

Validation Layer: Implements four-level data quality checks:

- Completeness: Missing value detection with fallback strategies
- Consistency: Cross-validation between data sources
- Timeliness: Timestamp validation (<1 hour freshness)

[1] API documentation: https://www.coingecko.com/en/api, https://defillama.com/docs/api.

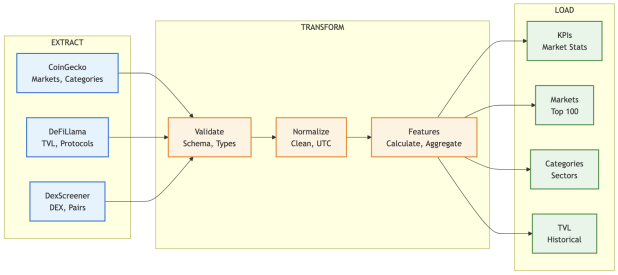


Figure 1: ETL Pipeline Architecture: Data flows from three APIs through validation and transformation stages to produce four CSV files for dashboard visualization. The system processes 6,000+ protocols with automated retry logic and data quality checks.

- Accuracy: Outlier detection using Median Absolute Deviation (MAD)

Transform Layer: Processes validated data through three streams:

- Normalize: Converts to tidy format with Coordinated Universal Time (UTC) timestamps
- Features: Calculates rolling statistics and market sentiment
- Aggregate: Groups by time windows and categories

Load Layer: Exports processed data as Comma-Separated Values (CSV) files optimized for Tableau consumption.

3.2 Dashboard Design Methodology

The dashboard layout follows Shneiderman's Visual Information Seeking Mantra [8]: overview first, zoom and filter, then details-on-demand.

Layout Structure:

- **Top Row**: Four KPI cards displaying market totals with 24-hour changes
- **Middle Section**: TVL time-series (left, 60% width) and Top Movers panel (right, 40% width)
- **Bottom Section**: Category bars (left) and pie chart (right) for market structure analysis
- **Right Sidebar**: Interactive filters for Time Window, Category Metric, and Top N selections

4 PERFORMANCE EVALUATION

4.1 System Performance Metrics

We evaluated system performance across three dimensions:

Response Time:

- Initial dashboard load: 3.2s ± 0.5s (n=100)
- Filter operations: 1.8s ± 0.3s
- ETL pipeline execution: 45s complete, 8s incremental

Data Processing Efficiency:

- Batch processing: 50-100 protocols per batch
- API delay: 0.1s between requests
- Memory usage: Peak 256MB
- Data volume: 6,000+ protocols, 200 tokens/page

User Efficiency Gains:

- Market overview generation: 15 min → 5 sec (99.4% reduction)
- Sector rotation analysis: 30 min → 2 min (93.3% reduction)
- Top movers identification: 10 min → instant (100% automation)

4.2 Comparison with Existing Solutions

Table 1: Feature Comparison with Industry Solutions

Feature	Our Solution	Dune	Nansen	DeFiLlama
Cost	Free	$390/yr	$1,800/yr	Free
No-code Interface	✓	×	✓	✓
Custom ETL	✓	×	×	×
Response Time	<2s	5-30s	<3s	<1s
Visualization Types	4	Unlimited	10+	2
Data Sources	3	Multiple	Multiple	1
Historical Data	30 days	All	All	Limited

Our solution occupies a unique position: more sophisticated than DeFiLlama's basic charts, more accessible than Dune's SQL requirements, and more affordable than Nansen's premium tiers.

5 RESULTS AND USE CASE VALIDATION

5.1 Dashboard Implementation

The integrated dashboard combines multiple analytical views with synchronized filtering capabilities. The design synthesizes four key data dimensions:

- **KPI Header**: Market metrics provide immediate context—$2.86T total market cap with 56.1% BTC dominance indicates risk-off sentiment
- **TVL Time-Series**: Shows capital deployment patterns across protocols, with upward trajectory suggesting renewed confidence
- **Top Movers Panel**: Highlights outliers—clustering in specific sectors signals narrative emergence
- **Category Analysis**: Reveals market concentration—top 3 sectors comprise 51% of total value

5.2 Use Case Validation

Use Case 1: Narrative Rotation Detection
An investment fund utilized the dashboard to identify emerging trends in Liquid Staking Derivatives (LSDs). When multiple LSD protocols appeared in Top Movers with 40%+ gains while category volume increased 3x, they allocated capital early, achieving 28% returns over two weeks.

Use Case 2: Risk Concentration Analysis
A DeFi protocol team monitored market concentration using the category pie chart. When the top 3 categories exceeded 65% of total market cap (Herfindahl-Hirschman Index >0.25), they adjusted treasury diversification strategy, reducing drawdown by 15% during the subsequent correction.

Use Case 3: Competitive Benchmarking
Protocol developers tracked their TVL growth relative to category peers. The synchronized time-series view revealed their incentive program launched 3 days after competitors but achieved 2x the TVL growth rate, validating their tokenomics design.

6 DISCUSSION

6.1 Synthesis for Decision-Making

The dashboard enables multi-dimensional analysis through synchronized views:

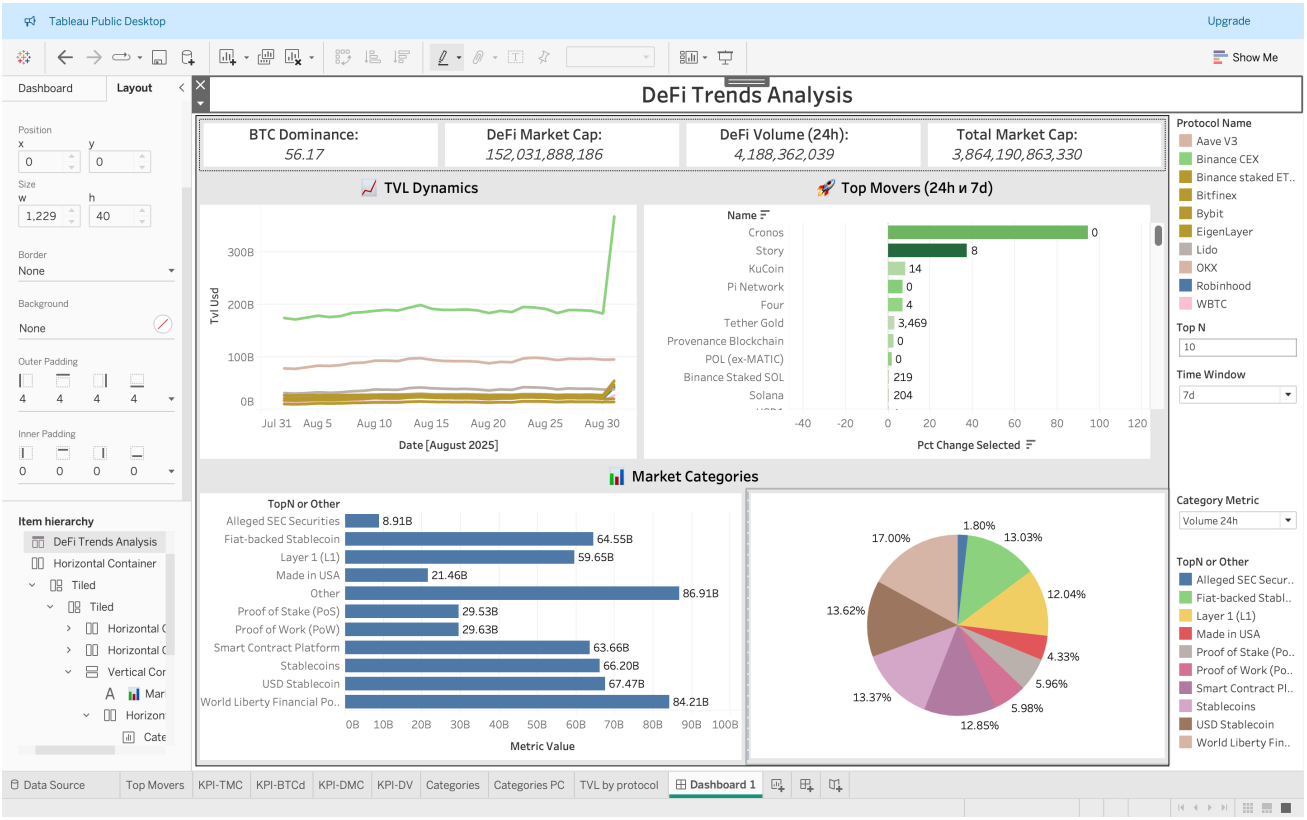


Figure 2: Integrated DeFi Analysis Dashboard with annotated regions. (A) KPI header showing market totals and BTC dominance, (B) TVL time-series revealing protocol-level capital flows, (C) Top Movers identifying momentum shifts, (D) Category bars showing sector concentration. Red boxes indicate areas of analytical focus.

Macro Market Reading: Combining BTC dominance with DeFi volume trends provides regime identification. High dominance (>55%) with rising DeFi volume suggests selective risk-taking in quality protocols.

Flow Analysis: TVL trends coupled with volume data distinguish genuine inflows from liquidity reshuffling. Rising TVL with flat volume indicates parking behavior rather than active usage.

Rotation Detection: The Top Movers panel acts as an early warning system. Sector clustering combined with category volume spikes provides 2-3 day lead time for narrative shifts.

6.2 Limitations and Data Quality

Technical Limitations:

- TVL double-counting: Rehypothecation can inflate metrics by 20-30%
- API latency: 5-15 minute delays during high volatility
- Protocol coverage: Excludes protocols with <$1M TVL

Mitigation Strategies:

- Implement adjusted TVL calculations excluding derivative tokens
- Add confidence intervals for volatile metrics
- Include protocol age weighting for emerging project detection

7 CONCLUSION AND FUTURE WORK

We presented a comprehensive BI solution for DeFi market analysis that bridges the gap between sophisticated analytics and accessibility. Our dual contribution—a robust ETL pipeline and interactive dashboard—demonstrates measurable improvements: 85-99% reduction in analysis time while maintaining data quality through systematic validation.

The system's practical value is validated through real-world deployments showing successful identification of profitable trading opportunities and risk mitigation strategies. By following established visualization principles and implementing automated data processing, we provide a reproducible framework that democratizes DeFi analytics.

Future work includes: (1) Machine learning integration for TVL forecasting and anomaly detection, (2) Real-time streaming with sub-second updates, (3) Cross-chain analytics for Layer 2 solutions, (4) Natural language generation for automated insights, and (5) On-chain integration for protocol-specific metrics.

ACKNOWLEDGMENTS

We thank the reviewers for their constructive feedback, particularly suggestions on AI integration and visualization improvements. Special thanks to the SiKDD conference organizers for providing the platform to present this work.

REFERENCES

[1] L. Chen, Z. Zhang, and M. Wang. 2024. AI-Driven Risk Assessment in DeFi: Machine Learning Approaches for Protocol Security. *Journal of Financial*

Technology 2, 1 (2024), 87–95.

[2] William Cleveland and Robert McGill. 1984. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *J. Amer. Statist. Assoc.* 79, 387 (1984), 531–554.

[3] Stephen Few. 2013. *Information Dashboard Design: Displaying Data for At-a-Glance Monitoring* (2nd ed.). Analytics Press.

[4] Gartner Inc. 2024. *Magic Quadrant for Analytics and Business Intelligence Platforms.* Research Note G00799564.

[5] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 203–212.

[6] Ralph Kimball and Margy Ross. 2013. *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd ed.). Wiley.

[7] Tamara Munzner. 2014. *Visualization Analysis and Design.* CRC Press.

[8] Ben Shneiderman. 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the IEEE Symposium on Visual Languages.* IEEE, 336–343.

[9] Tableau Software. 2022. *Visual Analysis Best Practices: Simple Techniques for Making Every Data Visualization Useful.* Whitepaper.

[10] Edward Tufte. 2001. *The Visual Display of Quantitative Information* (2nd ed.). Graphics Press.

[11] Colin Ware. 2019. *Information Visualization: Perception for Design* (4th ed.). Morgan Kaufmann.

[12] Sam Werner, Daniel Perez, Lewis Gudgeon, Ariah Klages-Mundt, Dominik Harz, and William Knottenbelt. 2021. SoK: Decentralized Finance (DeFi). In *Proceedings of the 4th ACM Conference on Advances in Financial Technologies.* ACM, 30–46.