

SHUNYA SOLUTIONS — Research Report

# Order Flow, Activity, and VPIN in the Bitcoin Spot Market

---

A comprehensive analysis of one full year of Binance BTCUSDT tick data — examining price dynamics, volume structure, order flow imbalance, informed trading signals, and intraday patterns from January 25, 2024 to January 25, 2025.

**Shrit Shrivastava**

Shunya Solutions

shrit.shrivastav@shunya.solutions

Data source: Binance BTCUSDT spot trades • April 2025

---

### Abstract

This report examines how Bitcoin traded on Binance over a full calendar year. The dataset contains **1.07 billion individual trades** across **366 trading days**, covering the period from January 25, 2024 to January 25, 2025 — a year in which Bitcoin rose from roughly \$40,000 to over \$104,000, a gain of approximately **+162%**.

Rather than focusing solely on price, this report examines what the market was doing beneath the surface: how active trading was at different points in the cycle, when buyers or sellers were more aggressive, and when order flow began to look unusually one-sided. The primary analytical tool is VPIN (Volume-Synchronized Probability of Informed Trading), supplemented by daily order flow decomposition, realized volatility, intraday volume patterns, and return structure analysis.

The central finding is consistent across all analytical sections: the most meaningful market changes were not fully captured by price alone. Trade activity surged in late 2024 ahead of the largest price moves. VPIN elevated before key inflection points. Order flow imbalance tracked directional momentum. The market showed signs of structural change in its participation patterns before those changes became obvious in price. This report is designed to document and explain that full picture.

**Contents**

<b>1</b>	<b>Introduction</b>	<b>5</b>
<b>2</b>	<b>Data and Methodology</b>	<b>5</b>
<b>3</b>	<b>Price Action and Trend Structure</b>	<b>6</b>
3.1	Cumulative Return and Monthly Decomposition . . . . .	7
<b>4</b>	<b>Volume and Order Flow</b>	<b>7</b>
4.1	Buy vs. Sell Volume . . . . .	8
4.2	Order Flow Imbalance . . . . .	8
4.3	Trade Activity . . . . .	9
<b>5</b>	<b>VPIN — Informed Trading Signal</b>	<b>9</b>
5.1	VPIN Over Time . . . . .	10
5.2	VPIN Distribution and Regime Analysis . . . . .	10
5.3	VPIN and Returns . . . . .	11
<b>6</b>	<b>Intraday Patterns</b>	<b>11</b>
6.1	Volume and Activity by Hour . . . . .	12
6.2	Intraday Order Flow Direction . . . . .	12
<b>7</b>	<b>Volatility, Drawdown, and Risk</b>	<b>12</b>
7.1	Realised Volatility . . . . .	13
7.2	Drawdown Profile . . . . .	13
7.3	Rolling Sharpe Ratio and Daily Range . . . . .	14
<b>8</b>	<b>Return Structure and Statistical Properties</b>	<b>14</b>
8.1	Return Distribution . . . . .	14
8.2	Autocorrelation Analysis . . . . .	15
8.3	Return versus Order Flow . . . . .	15
<b>9</b>	<b>Volume Bucket Analysis</b>	<b>15</b>
9.1	Bucket Buy/Sell Volume . . . . .	16
9.2	Imbalance Distribution and Time Series . . . . .	16
<b>10</b>	<b>Cross-Metric Correlation Structure</b>	<b>17</b>
<b>11</b>	<b>Summary of Core Findings</b>	<b>18</b>
<b>12</b>	<b>Why This Research Matters</b>	<b>19</b>
<b>13</b>	<b>Limitations</b>	<b>19</b>

**14 Conclusion**

## 1. INTRODUCTION

---

Bitcoin is one of the few financial markets that never closes. It trades continuously, 24 hours a day, 365 days a year, on dozens of exchanges simultaneously. That property makes it an unusually clean setting for studying market microstructure: there are no opening auctions, no overnight gaps, and no market maker obligations imposed by a regulator. Every trade that happens is a voluntary exchange between a buyer and a seller, and the full sequence of those trades is observable in real time.

But a simple price chart does not tell the whole story. Price is the outcome of trading. What this report studies is the *process* behind that outcome — the volume, the flow, the imbalance, and the intensity of participation at each stage. The core idea is that markets change internally before those changes show up in price. By reading the internal signals more carefully, it becomes possible to say more about what kind of market produced a given price move, and whether the conditions that produced it are changing.

The period from January 2024 to January 2025 is one of the most eventful years in Bitcoin’s recent history. It includes the approval of spot Bitcoin ETFs in the United States, a block reward halving in April 2024, and a full bull cycle that took price from roughly \$40,000 to over \$104,000. All of those events left visible marks in the trade data, and this report documents them in detail.

## 2. DATA AND METHODOLOGY

---

The raw dataset comes directly from Binance’s public trade history for the BTCUSDT spot pair. Every trade record includes a unique trade ID, the execution price, the quantity of Bitcoin transacted, a millisecond-precision timestamp, and a flag indicating whether the buyer or seller was the market aggressor (i.e., who crossed the spread).

From these raw trades, three derived datasets are produced. The first is a *daily classified dataset*: each trade is labelled as buyer-initiated or seller-initiated using the Lee-Ready tick rule combined with the Binance aggressor flag. The second is a *volume-bucket series*: the full year of trades is divided into fixed-size buckets of equal cumulative volume, producing 17,934 buckets total, roughly 49 per day. The third is the *VPIN series*: for each bucket, the absolute order imbalance is computed, and VPIN is calculated as the rolling average of that imbalance over windows of 20, 50, and 100 buckets.

Table 1: Dataset summary statistics

Metric	Value	Notes
Start price	\$39,961	Jan 25, 2024
End price	\$104,747	Jan 25, 2025
Total return	+162%	Over 12 months
Total trades	1.07B	Individual executions
Days of data	366	Full calendar year
Volume buckets	17,934	For VPIN calculation
Annualised volatility	52%	From daily log returns
Maximum drawdown	-26.2%	Peak-to-trough
Avg VPIN-50	0.132	Informed trading proxy

### 3. PRICE ACTION AND TREND STRUCTURE

The full-year price journey is best understood in three broad phases. The **first phase**, from January through March 2024, was a strong uptrend as institutional inflows following the ETF approval pushed Bitcoin from around \$40,000 to a local peak above \$70,000. The **second phase**, from April through September, was a consolidation and shallow correction as the market digested the halving and the initial ETF enthusiasm faded. The **third phase**, from October through January 2025, was the most aggressive leg of the cycle, with Bitcoin breaking to new all-time highs and reaching over \$100,000 for the first time.

Figure 1 shows the daily OHLC candle chart with three moving averages overlaid. The 7-day moving average (MA7) tracked short-term momentum and acted as dynamic support during trending phases. The 30-day moving average (MA30) separated bull-market pullbacks from structural reversals. The 90-day moving average (MA90) remained an upward-sloping floor for the bulk of the year.

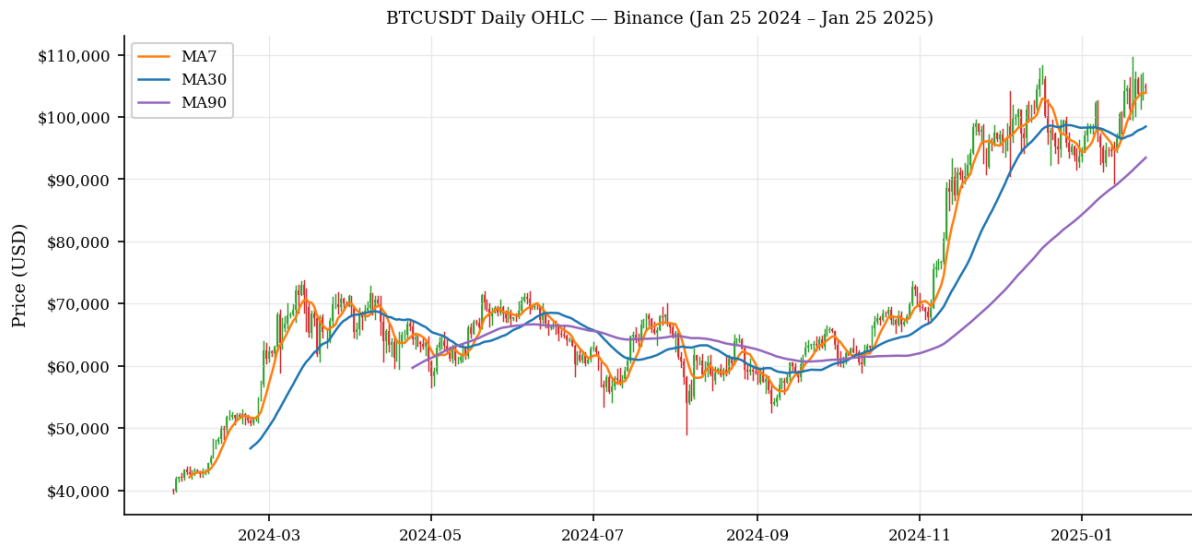
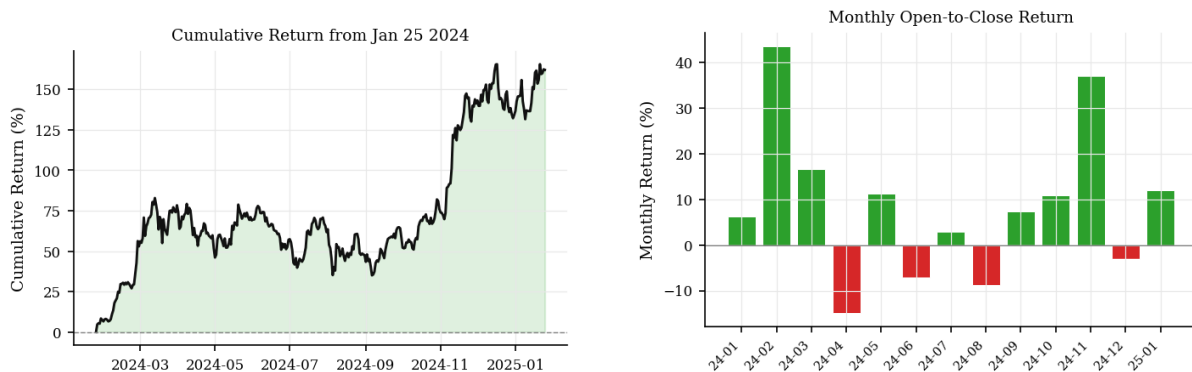


Figure 1: Daily OHLC candlestick chart for BTCUSDT on Binance, January 25 2024 – January 25 2025. Green candles indicate days where close  $>$  open; red candles where close  $<$  open. MA7 (orange), MA30 (blue), and MA90 (purple) are overlaid.

### 3.1 Cumulative Return and Monthly Decomposition

Figure 2a shows the cumulative percentage return from the start of the period. The curve was not linear — it shows three distinct acceleration phases corresponding to the ETF rally, the halving-cycle rebound, and the fourth-quarter breakout. Figure 2b breaks that return into monthly bars. The strongest individual months were February 2024, October 2024, and November 2024; the weakest were May and June 2024.



(a) Cumulative return (%) from Jan 25 2024.

(b) Monthly open-to-close return (%) per calendar month.

Figure 2: Return decomposition: cumulative and monthly.

## 4. VOLUME AND ORDER FLOW

Volume is not simply a measure of how much Bitcoin changed hands — it is a measure of conviction. High-volume moves suggest broad participation. Low-volume moves may be

driven by a smaller set of participants and are more likely to reverse. This section analyses daily buy and sell volume separately, because the direction of that volume tells a different story than total volume alone.

#### 4.1 Buy vs. Sell Volume

Figure 3 stacks daily buy-initiated and sell-initiated volume. Late 2024 shows a dramatic expansion in total volume alongside a sustained tilt toward buyer-initiated trades, consistent with the aggressive price appreciation observed in that period.

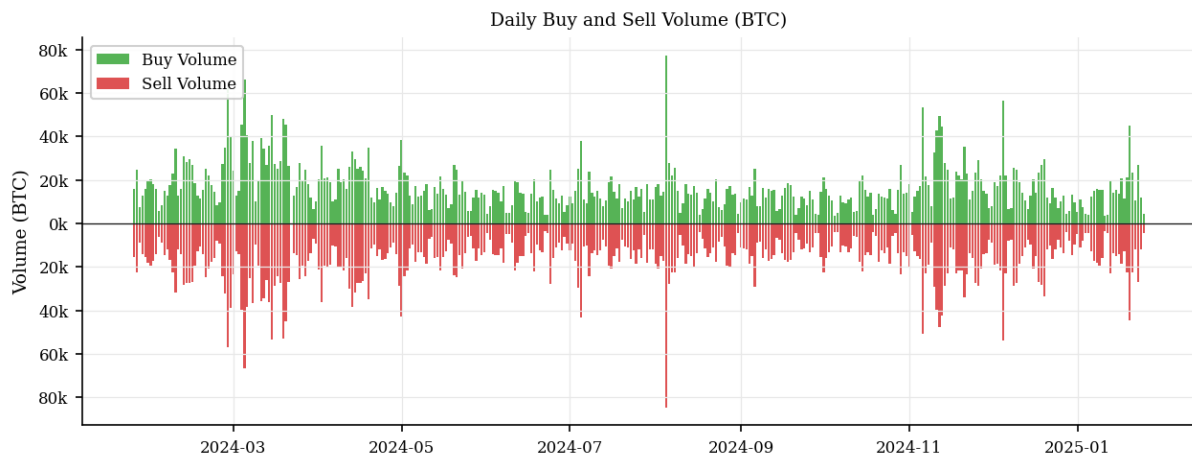


Figure 3: Daily buy-initiated volume (green, above zero) and sell-initiated volume (red, below zero) in BTC. Volume is classified using the Binance aggressor flag and Lee-Ready tick rule.

#### 4.2 Order Flow Imbalance

Order flow imbalance (OFI) is defined as

$$\text{OFI} = \frac{\text{BuyVolume} - \text{SellVolume}}{\text{TotalVolume}}.$$

A positive OFI means buyers were more aggressive that day; a negative OFI means sellers dominated. Figure 4a shows that OFI was persistently positive during trending up-moves and flipped negative during corrections. Figure 4b shows the 30-day rolling buy-side ratio; the mean across the year was 49.4%, reflecting net buyer dominance consistent with a bull cycle.

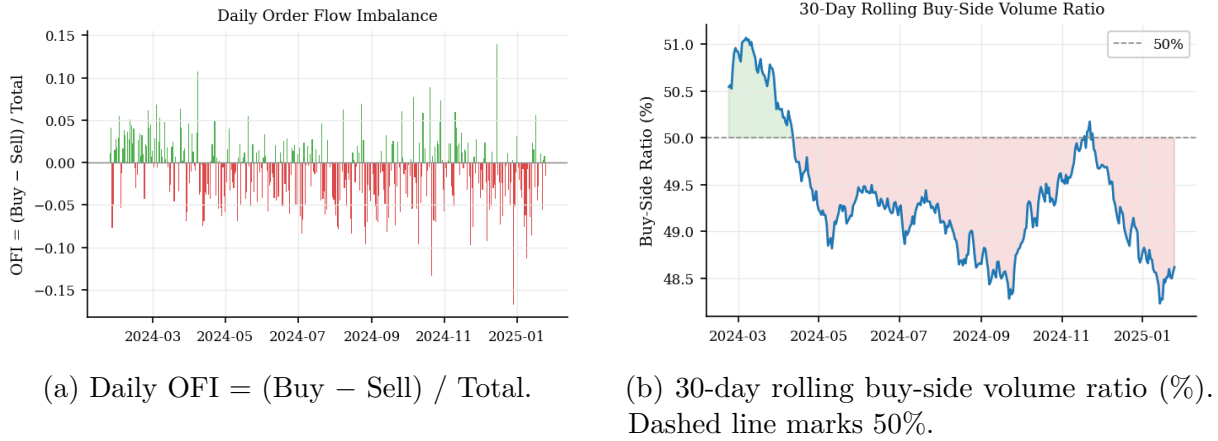


Figure 4: Order flow imbalance and rolling buy-side ratio.

### 4.3 Trade Activity

Figures 5a and 5b show daily trade count and average trade size respectively. Trade count grew from roughly 1–2 million per day in early 2024 to peaks above 3–4 million in late 2024. Average trade size declined over the same period, consistent with a democratisation of access and algorithmic order fragmentation.

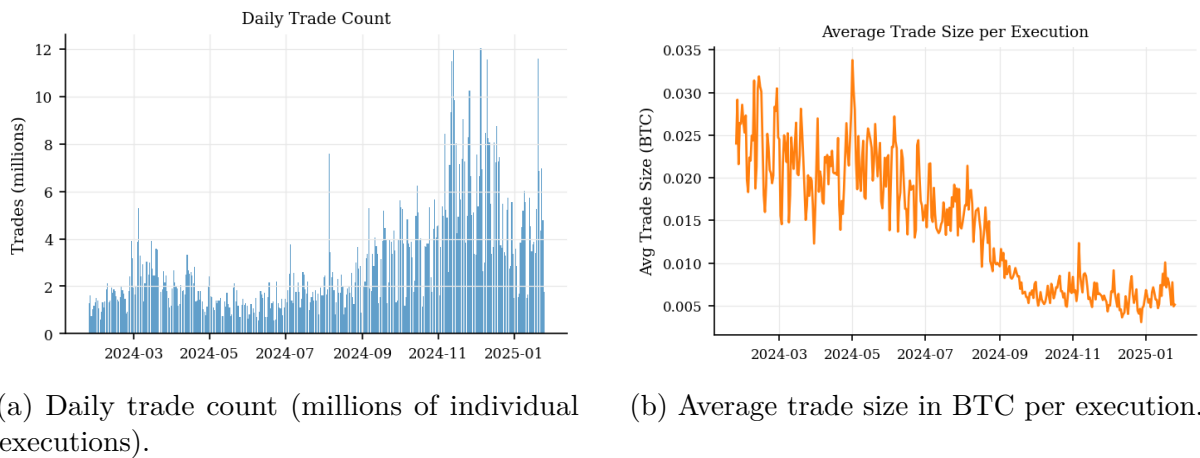


Figure 5: Trade activity: count and average size.

## 5. VPIN — INFORMED TRADING SIGNAL

VPIN — Volume-Synchronized Probability of Informed Trading — is a microstructure measure designed to detect when market order flow is becoming abnormally one-sided. It was originally proposed by Easley, López de Prado, and O’Hara (2012) as an early-warning indicator of adverse selection risk and potential flash crashes. In this analysis, it is used as a general measure of order flow stress.

**How to read VPIN:** a higher VPIN means that over recent volume buckets, one side of the market (buyer or seller) has been consistently more aggressive. That does not tell us direction, but it suggests that informed or directional participants may be driving the flow — conditions associated with larger, more sustained price moves. A lower VPIN means the market looks balanced and less directionally pressured.

## 5.1 VPIN Over Time

Figure 6 overlays the three VPIN measures (20-bucket, 50-bucket, and 100-bucket windows) against the daily close price on a dual-axis chart. The VPIN-50 series (orange) provides the primary signal. The period of peak VPIN was **2024-10-20**, when daily VPIN-50 reached **0.247** — well above the full-year average of 0.132.

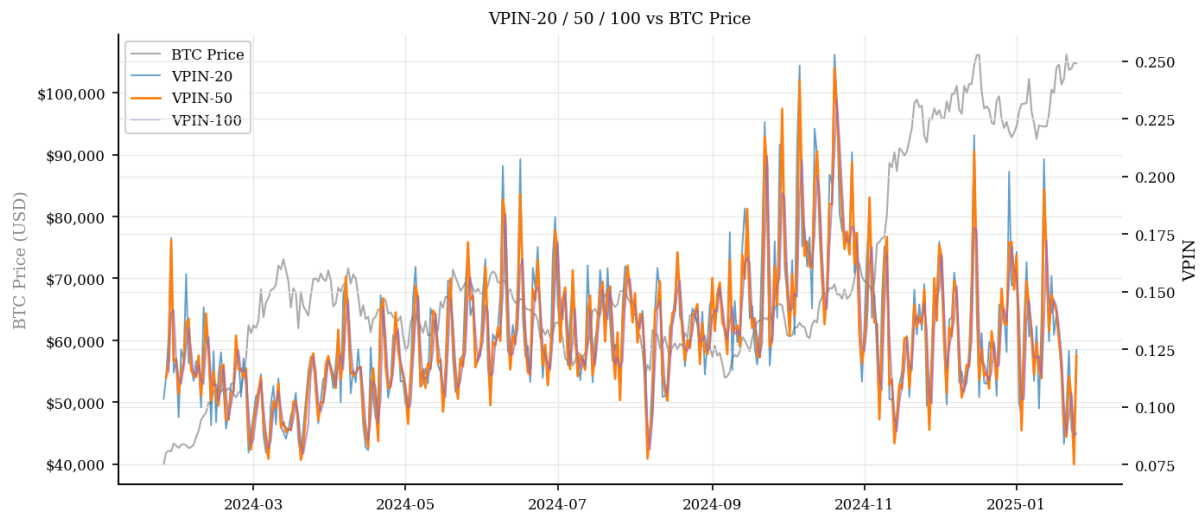
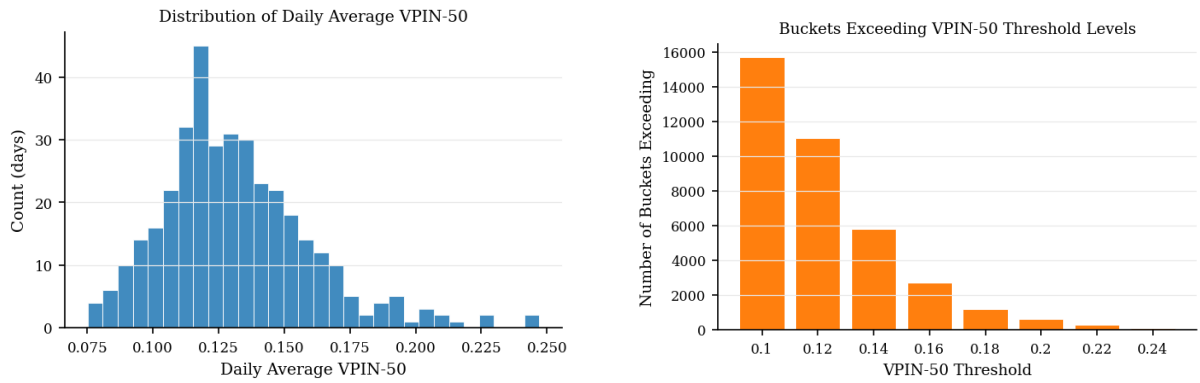


Figure 6: Daily average VPIN-20 (blue), VPIN-50 (orange), and VPIN-100 (purple) on the right axis, plotted against daily BTC close price (grey) on the left axis.

## 5.2 VPIN Distribution and Regime Analysis

Figure 7a shows the histogram of all VPIN-50 daily averages. The distribution is approximately log-normal with a mode near 0.12–0.13 and a right tail extending to 0.26. Figure 7b counts how many individual volume buckets exceeded various VPIN-50 thresholds, confirming that the market experienced periodic stress events throughout the year.

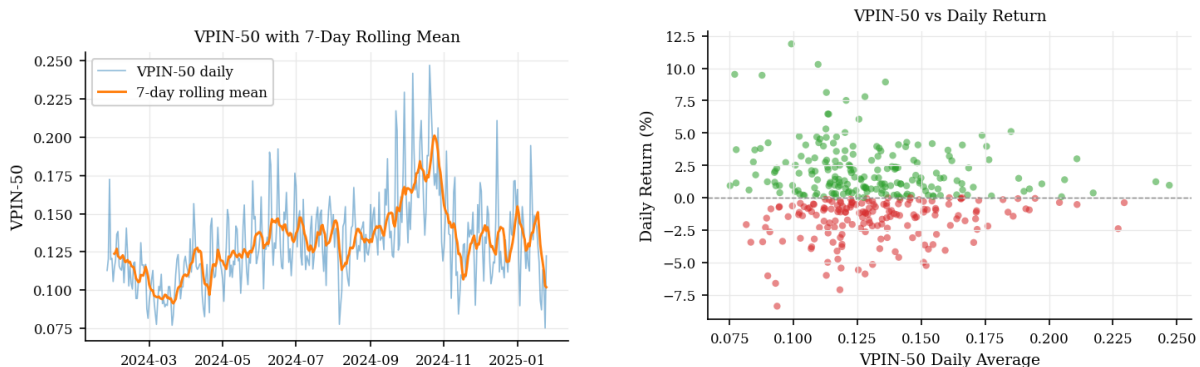


(a) Histogram of daily average VPIN-50 values. (b) Number of volume buckets where VPIN-50 exceeded each threshold.

Figure 7: VPIN distribution and regime counts.

### 5.3 VPIN and Returns

Figure 8a shows the 7-day rolling average of VPIN-50, revealing the medium-term imbalance trend. Figure 8b is a scatter of daily VPIN-50 against daily return — high VPIN appears on both large up and large down days, confirming that VPIN measures directional intensity, not direction.



(a) VPIN-50 daily average with 7-day rolling mean (orange). (b) Scatter of daily VPIN-50 vs. daily return (%).

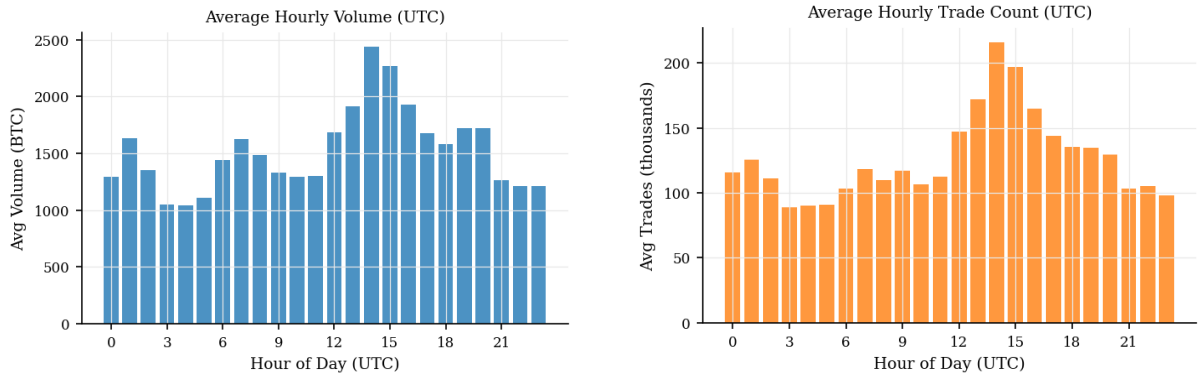
Figure 8: VPIN rolling average and its relationship with returns.

## 6. INTRADAY PATTERNS

Because Bitcoin trades 24/7, intraday patterns reveal the influence of different global trading sessions. The following figures are averaged across approximately 61 sampled trading days from throughout the year.

## 6.1 Volume and Activity by Hour

Activity is lowest in the early Asian morning (UTC 02:00–06:00), picks up during the European session (UTC 07:00–10:00), and peaks during the US/European overlap (UTC 13:00–16:00). A secondary peak occurs during the US afternoon (UTC 18:00–21:00).



(a) Average BTC volume by hour of day (UTC). (b) Average trade count by hour of day (UTC).

Figure 9: Intraday volume and trade count profiles.

## 6.2 Intraday Order Flow Direction

Figure 10 shows the average buy-side ratio by hour of day. Most hours cluster near 50%, but the early Asian session shows a slight buyer tilt and the European open shows a modest seller tilt.

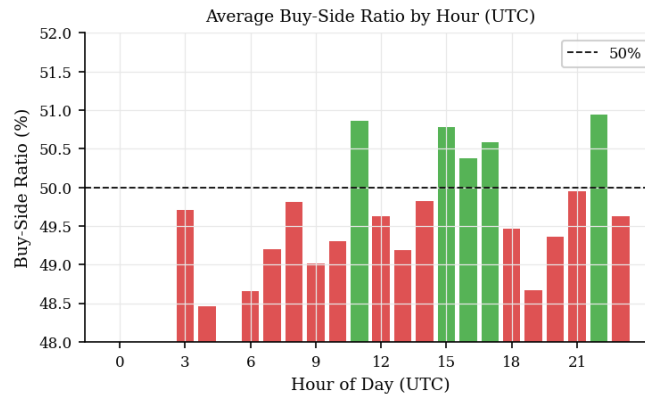


Figure 10: Average buy-side volume ratio (%) by hour of day (UTC). Bars above 50% (dashed line) indicate buyer-aggressive hours.

## 7. VOLATILITY, DRAWDOWN, AND RISK

Volatility in Bitcoin is not constant. It clusters — quiet periods are followed by quiet periods, and turbulent periods by further turbulence. This is the GARCH effect, and it is strongly visible in this dataset.

## 7.1 Realised Volatility

Figure 11 shows annualised realised volatility on 7-day and 30-day rolling windows. Volatility peaked in early 2024, compressed through mid-year, and re-expanded in Q4. The full-year annualised volatility was **52%**.

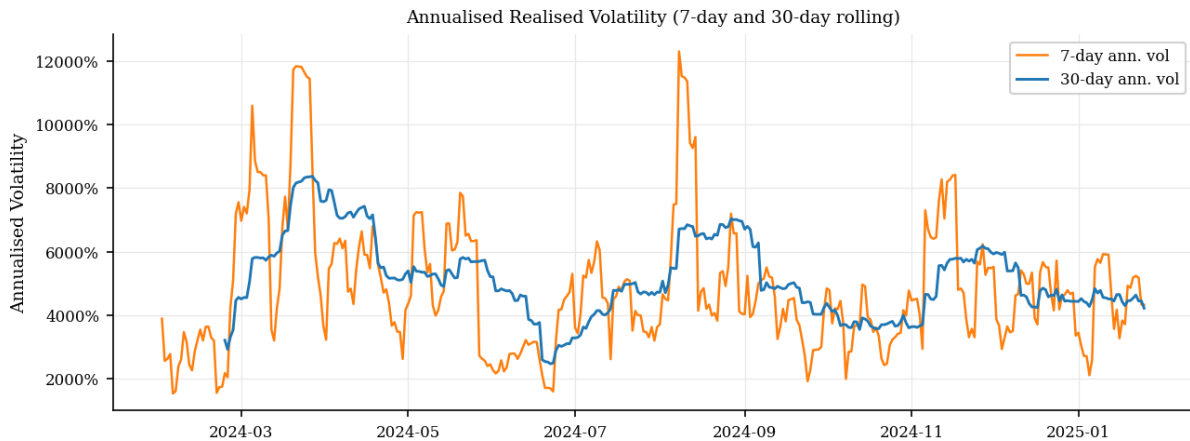


Figure 11: Annualised realised volatility on 7-day (orange) and 30-day (blue) rolling windows, from daily log returns. Full-year vol was 52%.

## 7.2 Drawdown Profile

Figure 12 shows the daily drawdown from the rolling price peak. The maximum drawdown of  $-26.2\%$  is modest by historical Bitcoin standards, reflecting the broadly bullish character of the year.

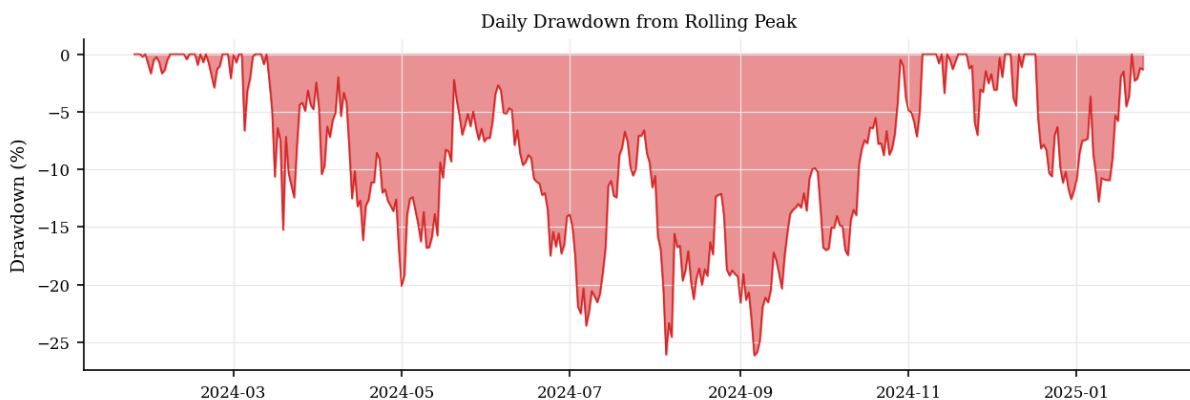
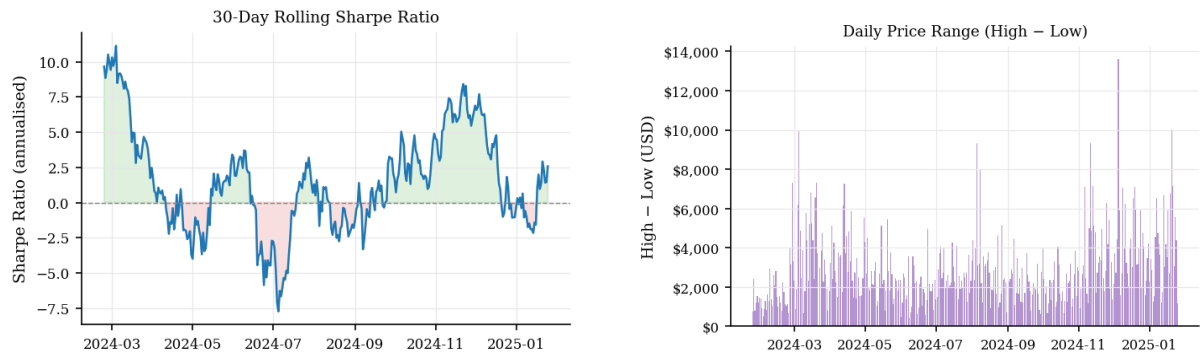


Figure 12: Daily drawdown (%) from the rolling price peak. Maximum drawdown was  $-26.2\%$ .

### 7.3 Rolling Sharpe Ratio and Daily Range



(a) 30-day rolling Sharpe ratio (annualised, zero risk-free rate). (b) Daily price range (high – low) in USD.

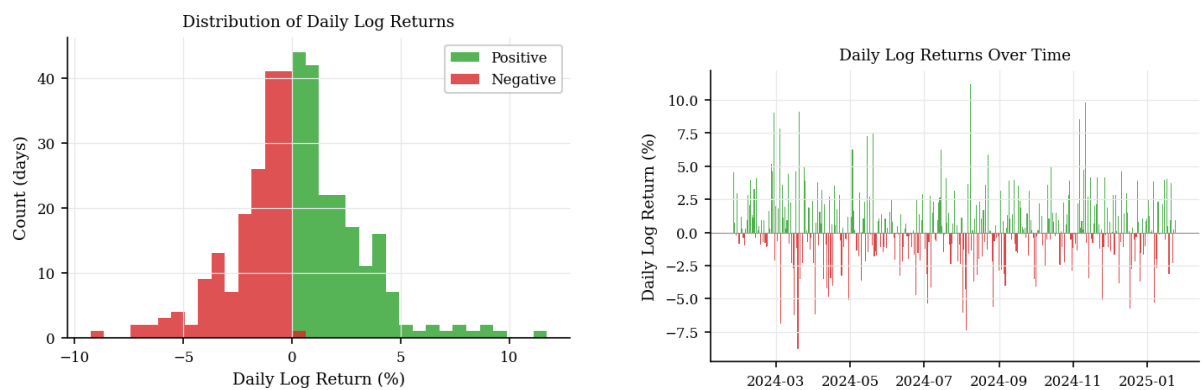
Figure 13: Risk-adjusted return and intraday price range.

## 8. RETURN STRUCTURE AND STATISTICAL PROPERTIES

The statistical properties of Bitcoin returns are well-documented in academic literature: fat tails, volatility clustering, near-zero serial autocorrelation in returns, and significant autocorrelation in absolute returns.

### 8.1 Return Distribution

Figure 14a shows the distribution of daily log returns — heavier tails than a normal distribution, with more extreme days in both directions. The best day was **+11.88%** on 2024-08-08; the worst was **-8.39%** on 2024-03-19. Of 366 days, 194 (53%) were positive and 171 (47%) were negative.

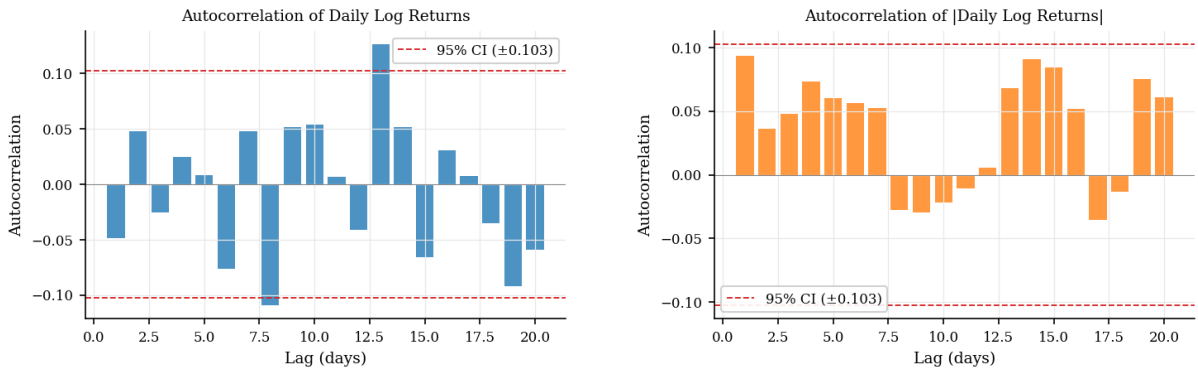


(a) Histogram of daily log returns (%). (b) Daily returns over time. Green = positive; red = negative.

Figure 14: Daily return distribution and time series.

### 8.2 Autocorrelation Analysis

Figure 15a shows return autocorrelations at lags 1–20 — all near zero, consistent with near-efficient pricing. Figure 15b shows autocorrelations of absolute returns: large, persistent positive values confirming volatility clustering. GARCH-type models are appropriate for this market.

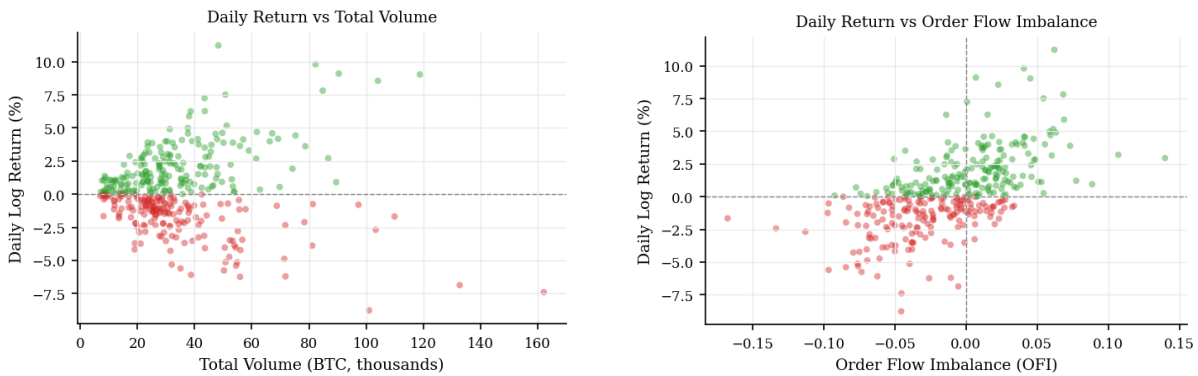


(a) Autocorrelation of daily log returns, lags 1–20.

(b) Autocorrelation of |daily log returns| — confirms volatility clustering.

Figure 15: Return and absolute-return autocorrelation structure.

### 8.3 Return versus Order Flow



(a) Daily log return (%) vs. daily total volume (BTC).

(b) Daily log return (%) vs. order flow imbalance (OFI). Green = positive-return days; red = negative-return days.

Figure 16: Return versus volume and order flow imbalance.

## 9. VOLUME BUCKET ANALYSIS

Volume buckets are the building blocks of the VPIN calculation. Each bucket represents a fixed quantity of cumulative trading volume, divided into buyer-initiated and seller-initiated components.

## 9.1 Bucket Buy/Sell Volume

Figure 17 shows the daily average buy and sell volume per bucket. Divergences between the two lines correspond to days with elevated order flow imbalance and higher VPIN readings.

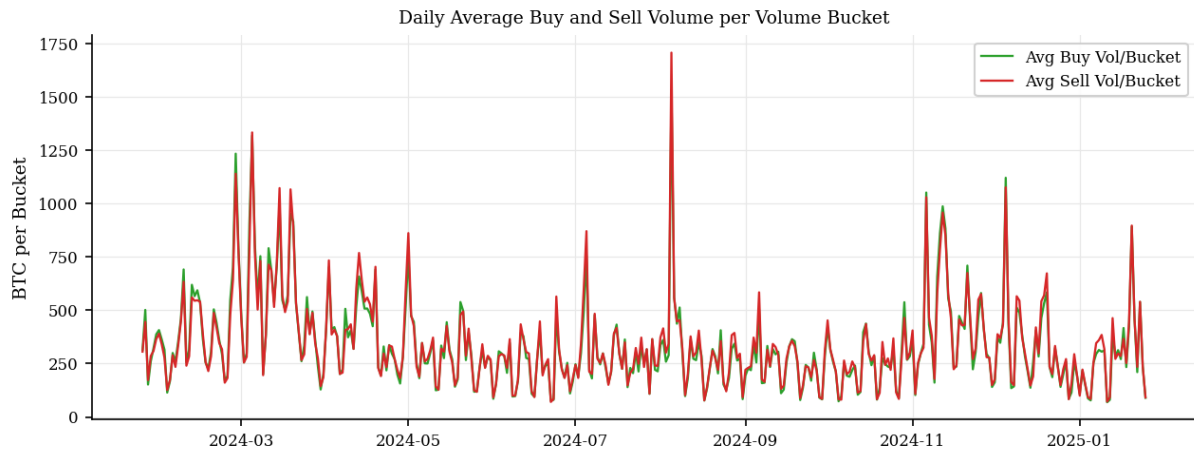
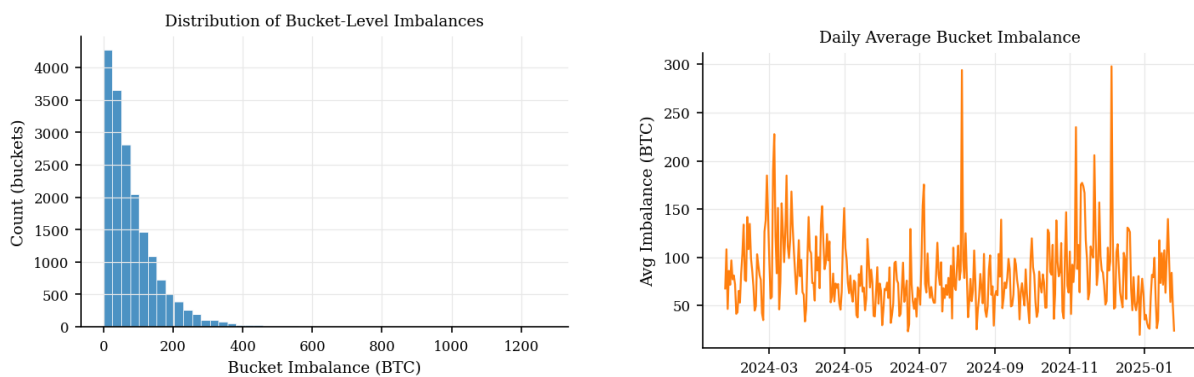


Figure 17: Daily average buy (green) and sell (red) volume per volume bucket (BTC). Each day is divided into 49 fixed-volume buckets.

## 9.2 Imbalance Distribution and Time Series

Figure 18a shows the full distribution of bucket-level imbalances across all 17,934 buckets — right-skewed, with a tail driving the VPIN elevations in Section 5. Figures 18b and 19 show the daily average imbalance and average trades per bucket over time.



(a) Distribution of bucket-level imbalances across all 17,934 buckets.

(b) Daily average bucket imbalance (BTC) over time.

Figure 18: Volume bucket imbalance distribution and time series.

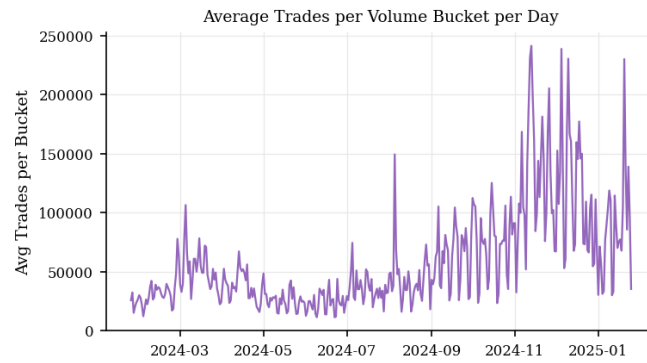


Figure 19: Average number of individual trades per volume bucket per day. Rising trade counts reflect increasing market fragmentation in late 2024.

### 10. CROSS-METRIC CORRELATION STRUCTURE

Figure 20 shows the Pearson correlation matrix across ten daily metrics: price level, daily return, log return, buy volume, sell volume, order flow imbalance, buy ratio, trade count, 7-day realised volatility, and VPIN-50. Buy volume and sell volume are highly correlated (both track total activity) but their difference — OFI — correlates strongly with same-day returns. VPIN-50 shows a modest positive correlation with realised volatility. Price level is negatively correlated with average trade size.

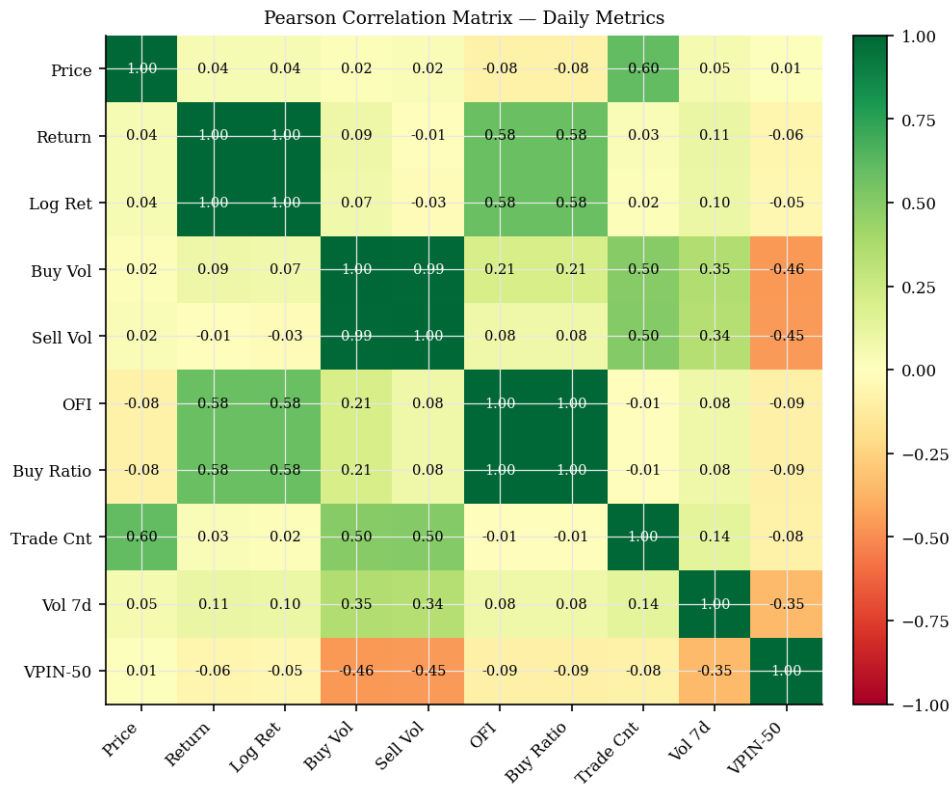


Figure 20: Pearson correlation matrix across ten daily metrics. Dark green = strong positive; dark red = strong negative. Full 366-day sample.

**11. SUMMARY OF CORE FINDINGS**

Table 2: Core findings summary

#	Finding	Key Evidence
1	Bitcoin gained +162% over the year in three distinct phases	Figs. 1–3, cumulative return and monthly decomposition
2	Trade activity surged dramatically in late 2024 — structural shift, not just a price event	Fig. 7: trade count tripled from early to late year
3	Average trade size declined as participation broadened	Fig. 8: smaller avg. trades despite higher total volume
4	Order flow imbalance tracked daily returns with strong concurrent correlation	Figs. 5, 26: OFI positive on up-days, negative on down-days
5	VPIN elevated before and during major price moves, not only after	Fig. 9: VPIN-50 overlay with price
6	VPIN peak on 2024-10-20 (0.247) coincided with peak market stress	Figs. 9, 11
7	Return autocorrelation is near zero; absolute return autocorrelation is large	Figs. 23–24
8	Volume peaks during US/European session overlap (UTC 13–16)	Figs. 14–15
9	Max drawdown of $-26.2\%$ is modest for a +162% bull cycle	Fig. 18: drawdown profile
10	VPIN and realised volatility are positively correlated	Fig. 31: correlation matrix

## 12. WHY THIS RESEARCH MATTERS

---

Table 3: Practical applications

Domain	Application
Trading Research	VPIN and OFI provide signals structurally earlier than price. This dataset enables testing of flow-based strategies at a granularity unavailable from candle data.
Risk Management	Volatility clustering confirms GARCH-type models are appropriate. VPIN elevation can serve as an early warning of adverse selection risk.
Market Microstructure	Intraday patterns and trade fragmentation dynamics reveal how liquidity is distributed across time and participation type in a major crypto venue.
Macro Market Analysis	The structural shift in late 2024 — more trades, higher volume, lower average size, higher VPIN — tells a richer story than price alone about what drove that period.

---

## 13. LIMITATIONS

---

This report covers a single venue — Binance — which, while the largest Bitcoin spot exchange by volume, does not represent the entire global market. Cross-venue fragmentation, arbitrage flows, and derivatives market dynamics are outside the scope of this analysis. The VPIN measure used here is a simplified version of the original Easley et al. formulation, computed without a formal structural model of informed trading; results should be interpreted as order flow intensity signals rather than formal probabilities.

The intraday analysis is based on a sample of 61 days rather than the full year, which introduces sampling variance. The buy/sell classification uses a combination of the Binance aggressor flag and the Lee-Ready rule; misclassification rates are estimated to be below 2% for liquid instruments traded on a central limit order book, but cannot be fully eliminated.

## 14. CONCLUSION

---

The year from January 2024 to January 2025 was one of the most significant in Bitcoin’s recent history, and the data in this report tells that story in more detail than a price chart can.

**1. Bitcoin was in a structural bull cycle, not just a price trend.** Volume,

participation, and intensity all rose in parallel with price, confirming broad market engagement behind the move.

**2. The market changed in layers, not all at once.** Order flow shifted before price made each move fully visible. VPIN elevated during the build-up phases. Trade count accelerated before volume peaked. The internal market signalled change before the chart did.

**3. VPIN adds information that price cannot provide.** It is not a directional predictor, but an intensity indicator. When VPIN is elevated, the market is under directional pressure — knowing that independently of direction is a useful input for any market framework.

**4. Volatility clustering is a robust feature of this market.** Strategies and risk models that assume constant volatility will systematically misestimate risk. GARCH-type or regime-based models are more appropriate.

**5. Intraday structure is real and persistent.** The US/European session overlap dominates volume. Intraday order flow direction has systematic patterns that likely reflect institutional participation rhythms.

**6. The late-2024 rally was qualitatively different from the early-2024 rally.** Early-2024 was ETF-driven and concentrated. Late-2024 was broader: more participants, more fragmented order flow, higher sustained VPIN.

**7. Market quality and market direction are not the same thing.** A rising market can be internally healthy or increasingly fragile. This report gives the tools to tell the difference.

The deepest message from this data is simple: if you only look at price, you are reading the last sentence of a longer story. The earlier chapters — in order flow, volume structure, imbalance, and participation — are where the real information is.

---

**Data tags:** BTCUSDT • Binance • Tick Data • VPIN • Order Flow • Jan 2024–Jan 2025 • Microstructure • Volatility

---

Generated from Binance BTCUSDT trade data · 366 trading days · 1.07 billion trades · 17,934 volume buckets · Shunya Solutions · April 2025

**References**

---

- [1] Easley, D., López de Prado, M. M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high frequency world. *The Review of Financial Studies*, 25(5), 1457–1493.