Impact of Liquidity Pool Size on Trading Volume in BTC-ETH Pools

21 June 2023

Team #111			
Team Member	**GT Id**	**Background**	**Projects/Experience**
Matias Vizcaino	avizcaino3	Data Engineer	• BSc in Economics and Finance (2015) • Data platforms, AI Products / MLOps, Business Dashboards
Walter Jack Simmons	wsimmons35	Science Teacher	• BSc of Physics and Secondary Education (2016) • Predict AP test scores based on student performance
MingShen Wang	mwang709	Research and Development	• BSc of Electrical and Electronics Engineering (2021) • Rapid classification of electrical components
Vítor de Matos Castilho	vcastilho3	Instrument and Controls Engineer	• Post-graduate in Offshore Systems Engineering (2014) • B.Sc. Control and Automation Engineer (2008) • Offshore operations analysis

Background and Problem Statement

The decentralized finance (DeFi) ecosystem relies heavily on liquidity pools, which are pools of tokens locked in smart contracts to facilitate trading by providing liquidity. This project focuses on the BTC-ETH liquidity pools on decentralized exchanges, specifically Uniswap, and aims to explore the relationship between liquidity pool size and trading volume.

The **objective of this analysis is to investigate and quantify the impact of liquidity pool size on trading volume in BTC-ETH liquidity pools on decentralized exchanges, with a specific emphasis on Uniswap.** By examining the relationship between pool size and trading volume, we aim to gain insights into how changes in pool size can affect trading strategies and liquidity provision within the rapidly evolving DeFi landscape.

Research Questions

Primary Research Question (RQ):
What is the relationship between liquidity pool size and trading volume in BTC-ETH liquidity pools?

Supplemental Research Questions:

1. How does the size of the liquidity pool influence the slippage in BTC-ETH trading pairs?
2. How does BTC-ETH price volatility affect trading volume relative to liquidity pool size?
3. Are there specific periods or events that significantly influence the relationship between the size of the BTC-ETH liquidity pool and its trading volume?

Business Justification

DeFi is a rapidly expanding market with a total value locked of approximately 48.78 billion USD (April 23, 2023 defillama.com). Investigating the relationship between liquidity pool size and trading volume provides insights into market mechanics, benefiting businesses and stakeholders.

For liquidity providers (LPs), understanding this relationship helps optimize liquidity provision strategies. If larger pool sizes are associated with higher trading volumes, depositing in larger pools becomes advantageous, leading to increased fee returns and potential risk mitigation. By studying this relationship, we contribute to the efficient functioning of DeFi markets, optimizing strategies for participants, and influencing the allocation of capital within the ecosystem.

DATASET

Inspired by the *"DeFi modeling and forecasting trading volume" (2023)*[1] paper, we sourced and constructed trade information for at least 6 months. Initial extraction work has already been performed, and you can refer to our GitHub repository[2] for further details and access to the code. The dataset consists of data obtained from the following sources:

Figure 1: Screenshot of Uniswap data.

1. Uniswap's The Graph API[3]

Description: Provides transaction details, trading volumes, and block information from Uniswap v3 liquidity pools. Focus: Specifically, we focus on the Uniswap v3 WBTC-WETH liquidity pools for fee tiers 400 and 3000. We would like to also add WBTC-USDC and USDC-WETH pools to consider Network spillover effects. Available Attributes: Transaction IDs, timestamps, amounts, USD equivalents, and other related data.

2. Etherscan API[4] for Uniswap transaction hashes

Description: Used to extract corresponding transaction data from Etherscan based on transaction hashes. Data: Block hashes, block numbers, sender addresses, gas details, transaction hashes, and other relevant information.

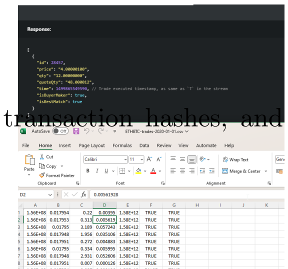

Figure 2: Screenshot of Binance data.

3. Binance[5] CEX Data for ETHBTC

Description: Daily zip trades downloaded using provided scripts from the Binance GitHub repository. Data: Detailed information about each trade executed on the Binance platform, including trade prices, quantities, timestamps, and buyer/seller characteristics.

Key Variables

Independent Variable: Liquidity pool trading volumes (amountUSD) over specific blocks.

Dependent Variables: The dependent variables are derived from a set of features and is categorized as follows:

1. Direct pool features: volatility, rate, number of trades, average trade size, total value locked (TVL).
2. Network spillover effects: trade flow imbalance.
3. CEX spillover effects: actual coin trade volume.
4. Price divergences between the 500 and 3000 pools.

Approach

1. Data Collection: Source data from multiple APIs (Uniswap, Binance) for liquidity pool sizes, trading volumes, and relevant variables. Associate data with Ethereum block numbers for consistent time measurement.
2. Data Preprocessing: Clean, format, and handle missing values. Ensure data consistency and address discrepancies arising from multiple sources.
3. Feature Engineering: Construct features to capture historical patterns, spillover effects, and price divergences. Utilize block-based time-series analysis framework.
4. Exploratory Data Analysis: Generate descriptive statistics, visualizations, and perform correlation analyses to uncover patterns, trends, and relationships between variables.
5. Model Selection and Development: Implement regression models (Linear Regression, Random Forest, Gradient Boosting) to investigate the pool size-trading volume relationship.
6. Model Evaluation and Optimization: Evaluate models using metrics (RMSE, MAE). Optimize selected model's hyperparameters using techniques like GridSearchCV or RandomizedSearchCV.
7. Final Analysis and Conclusions: Finalize analysis, draw conclusions, and communicate findings.

Throughout this process, we aim to maintain rigorous documentation of our methodologies and findings, which will allow us to ensure the transparency and replicability of our research.

[1] *"DeFi: modeling and forecasting trading volume on Uniswap v3 liquidity pools"* (2023), `https://ssrn.com/abstract=444535`.
[2] GitHub repository: `https://github.gatech.edu/MGT-6203-Summer-2023-Canvas/Team-111/tree/main/Code`
[3] Uniswap's The Graph API: `https://api.thegraph.com/subgraphs/name/uniswap/uniswap-v3`
[4] Etherscan API: `https://api.etherscan.io/api`
[5] Binance GitHub Repository: `https://github.com/binance/binance-public-data/blob/master/python/README.md`

Anticipated Conclusions, Hypothesis, and Business Impact

Through our analysis, we anticipate discovering a notable correlation between liquidity pool size and trading volume. Our hypothesis suggests that larger liquidity pools will demonstrate higher trading volumes, as they can facilitate larger transactions with minimal price impact.

The implications of our findings extend to various stakeholders within the ecosystem. For liquidity providers and traders, our research can provide insights to optimize their strategies, resulting in increased profits and minimized slippage. Additionally, DeFi platforms can leverage our analysis to design more efficient and appealing liquidity pools, fostering enhanced user engagement and driving overall platform growth.

Project Timeline

Week 1 Data Collection and Cleaning

- **Task 1.1** Data extraction from APIs (Uniswap and Binance) and consolidating the data.
- **Task 1.2** Initial data cleaning and preprocessing, which includes handling missing data and ensuring data consistency.

Week 2 Feature Engineering and Exploratory Data Analysis

- **Task 2.1** Constructing new features based on the existing data and guided by previous research.
- **Task 2.2** Conducting exploratory data analysis, generating descriptive statistics, and visualizations.

Week 3 Model Selection and Initial Evaluation

- **Task 3.1** Implementing various regression models like Linear Regression, Random Forest, and Gradient Boosting to investigate the relationship between pool size and trading volume.
- **Task 3.2** Conducting an initial evaluation of models based on metrics like RMSE and MAE.

Week 4 Model Optimization and Final Evaluation

- **Task 4.1** Optimizing the hyperparameters of the selected model using techniques such as GridSearchCV or RandomizedSearchCV to improve their predictive performance.
- **Task 4.2** Final evaluation of the models after hyperparameter optimization.

Week 5 Results Interpretation, Report Writing, and Finalization

- **Task 5.1** Final analysis, interpreting the results, and drawing conclusions.
- **Task 5.2** Writing and finalizing the report, which includes our findings and insights in relation to our initial research questions.
- **Task 5.3** Reviewing and refining the report to ensure its quality and accuracy.