# Market Data & Trade Analysis

Alan Matthew

March 13, 2025

## 1 Part A: Data Cleaning, EDA & Merging

In this part, I first ensured the data was clean. I then merged the trade and market data for each instrument and computed the `pnl` and markout `pnl` for each trade for analysis.

## 2 Part B: Market Data & Trades Visualization

After my EDA from **Part A**, I found an interesting one-hour period for each instrument where the trade `pnl` showed the most extreme spikes. I plotted both the market data and trades for this period. The one-hour window covers 30 minutes before and after the peak `pnl` timestamp.

## 3 Part C: Plotting the Aggregate Overall Margin Curves

For each instrument, I plotted the aggregate overall margin curve for all trades using the specified intervals in `horizon_ticks`. I first computed the `markout_pnl` for each trade at each horizon interval and divided the `markout_pnl` by the `trade_size` to compute **margin**, PNL normalized by volume.

## 4 Part D: Further Analysis

In the final part, I classified the strategy as a maker or taker for each instrument. I then analyzed the post-trade strategy characteristics and computed the maximum drawdown of the strategy for each instrument.

### 4.1 Maker or Taker Classification

The strategy is classified as a **Maker** for all instruments. I implemented a simple classification rule based on fill price to determine whether a trade was taking liquidity from the order book. The results showed that none of the trades were taking liquidity, indicating that the strategy must be a maker. One assumption I made in my calculation is that the bid and ask columns represent the best bid and offer (BBO). This classification is a best-effort heuristic, and for a more accurate determination, order submission and fill times would be required to assess how long after submission was an order filled, which would provide clearer insight into whether the strategy was making or taking.

# 5 Post-Trade Strategy Characteristics

In this section, I computed metrics which I believe are relevant for a market-making strategy, including trade size, PnL, inventory, holding time, trade frequency, quote aggressiveness, execution skew, and market impact. I will calculate maximum drawdown in the last section of the report. The results are summarized below:

| **Metric** | **BTAUSD** | **GMMAUSD** | **LMDAUSD** | **ZTAUSD** |
|---|---|---|---|---|
| Total Trades | 2,959 | 2,949 | 3,149 | 2,960 |
| Avg Trade Size | 5,489.02 | 5,488.64 | 5,493.49 | 5,487.50 |
| Total PnL | -827.41 | 4,041.50 | -862.30 | 4,149.15 |
| Avg PnL per Trade | -0.28 | 1.37 | -0.27 | 1.40 |
| Total Inventory | -264,000 | 110,000 | 683,000 | -655,000 |
| Avg Inventory | -89.22 | 37.30 | 216.89 | -221.28 |
| Avg Holding Time (ms) | 58,025.22 | 58,097.18 | 54,442.77 | 58,007.18 |
| Trade Frequency (trades/s) | 0.0172 | 0.0172 | 0.0184 | 0.0172 |
| Quote Agg. (Avg Dist. from Mid-Price) | 0.00005 | 0.00025 | 0.00005 | 0.00026 |
| Quote Agg. Over Tick Size | 5.18 | 5.07 | 5.10 | 5.17 |
| Execution Skew (Buy/Sell Ratio) | 0.99 | 0.98 | 1.07 | 0.96 |
| Avg Market Impact (15 sec. after fill) | 0.00000269 | 0.00004664 | -0.00000854 | -0.00005111 |
| Market Impact Over Tick Size | 0.269 | 0.933 | -0.854 | -1.022 |

Table 1: Post-Trade Strategy Characteristics

# 6 Maximum Drawdown Analysis

The table below summarizes the maximum drawdown and the timestamp when it occurred in milliseconds and UTC time.

| **Instrument** | **Max Drawdown** | **Timestamp (ms)** | **Timestamp (UTC)** |
|---|---|---|---|
| BTAUSD | 826.81 | 1683067370875 | 2023-05-02 22:42:50.875 |
| GMMAUSD | 1.30 | 1683010823350 | 2023-05-02 07:00:23.350 |
| LMDAUSD | 861.79 | 1683067273125 | 2023-05-02 22:41:13.125 |
| ZTAUSD | 0.90 | 1682951827050 | 2023-05-01 14:37:07.050 |

Table 2: Maximum Drawdown Analysis

# 7 Notes About the Jupyter Notebook

All Python code used to generate the analysis is provided in `main.ipynb`. The notebook contains more in-depth analysis and detailed explanations of both the results and code compared to this report. It is structured into sections that align with those in the report for easy reference.

**Note:** Before running the notebook, ensure that a Python environment is properly set up and linked to the kernel to allow for package installations and smooth execution.