REAL-TIME PRICES Methodology

Version 1.4

BLOCKSIZE produces the REAL-TIME PRICES Data Feed as part of its product BLOCKSIZE CONNECT, a collection of rates quoted in U.S. dollars, euros, and various other currencies for an expanding set of cryptocurrencies and other digital asset classes.

The REAL-TIME PRICES are designed to serve as a transparent and independent pricing source that promotes the functioning of efficient markets, reduces information asymmetries among participants, facilitates trading, and accelerates the adoption of cryptocurrencies as an asset class with the highest standards.

The REAL-TIME PRICES Data Feed is calculated using a robust and resilient methodology that is resistant to manipulation and adheres to international best practices for financial benchmarks.

This methodology is governed by the BLOCKSIZE Data Committee and is part of the BLOCKSIZE CONNECT Manifest (see Appendix A.1) that aims to ensure that the REAL-TIME PRICES Data Feeds serve as an accurate source of transparent and reliable pricing.

Contact email: connect@blocksize-capital.com

1 Calculation of Volume-Weighted Average Prices

The calculation algorithm of the REAL-TIME PRICES Data Feed as following:

1. Calculate the volume-weighted sum of prices $\sum_j p_j \times v_j$ denominated in units of the given asset from observable transactions that occurred for each of the accepted markets.
2. Calculate the aggregate volume for each of the accepted markets by adding the size of events across all accepted exchange markets, $\sum_j v_j$. The resulting figure is referred to as the volume weight.
3. Divide the volume-weighted sum of prices, $\sum_j p_j \times v_j$, by the total volume weight, $\sum_j v_j$, to obtain the volume weighted average price (VWAP) denoted by P_{VWAP}.

$$P_{VWAP} = \frac{\sum_j p_j \times v_j}{\sum_j v_j}$$

2 Data Contingency Rules

The following contingency rules are followed to address situations where data is delayed, missing, or unavailable due to periods of low liquidity such as extraordinary market circumstances or outside factors beyond the control of BLOCKSIZE.

1. If observable transactions from an accepted market are unable to be collected due to technical problems specific to the accepted market's exchange during the calculation of the REAL-TIME PRICES Data Feed (such as malformed data), the observable transactions are excluded from the calculation of the specific instance of the given Price Data Feed.
2. If no observable transactions from an accepted market exist during the current timeframe, the value of the REAL-TIME PRICES Data Feed will rely on the on various other accepted markets for its calculation.
3. If none of the accepted exchanges in Appendix A.3 report observable transactions, the REAL-TIME PRICES Data Feed will not report a calculated price. In the highly unlikely event that a user establishes a connection to a REAL-TIME PRICES Data Feed in a timeframe where none of the accepted exchanges (Appendix A.3) is reporting observable transactions, the price will be computed from the last observed transactions.

3 REAL-TIME PRICES Data Quality Assurance

All observable transactions from accepted markets are evaluated using algorithmic anomaly detection methods. If potential errors or anomalies in the transaction data are detected, it is not included in the calculation of the REAL-TIME PRICES Data Feed. The anomaly detection methods are applied to transaction data collected within liquidity-adjusted timeframes of a maximal duration of 120 minutes.

Complementary to the algorithmic anomaly detection approaches, the BLOCKSIZE Data Committee weighs in its expert judgment to maintain and improve the data quality of the REAL-TIME PRICES Data Feeds. The Data Committee may decide to include the transaction data from new sources that are vetted for data integrity and quality. Alternatively, the Data Committee may decide to exclude previously accepted markets that started to deliver erroneous or anomalous data. Any exercise of such expert judgment must be approved by staff members.

3.1 Overview: Data Processing and Anomaly Detection Methods

BLOCKSIZE uses mathematical statistics to automatically perform the task of ensuring high data quality for the volume weighted average price (VWAP). The anomaly detection and data processing use liquidity-adjusted timeframes to ensure the accuracy and contemporariness of confidence intervals as well as the computed VWAPs.

The approach to creating adaptive, liquidity-adjusted timeframes is specified in Subsection 3.2. . In Subsections 3.3 and 3.4, the Price-Based and Exchange-Based anomaly detection methods are introduced, respectively. These filter methods are based on the computation of statistical estimates for confidence intervals for price data and VWAPs. The VWAP-based method is used to filter out data sources / exchanges that deliver low-quality data and is referred to as “Exchange-Based Filter”. The price-based method establishes a filter for individual events with anomalous prices and is, hence, referred to as “Price-Based Filter”.
For the REAL-TIME PRICES Data Feeds, BLOCKSIZE considers transaction data in a liquidity-adjusted timeframe and applies first the Price-Based Filter and thereafter the Exchange-Based Filter for ensuring data quality.

3.2. Adaptive Liquidity-Adjusted Timeframes

To ensure that anomaly detection and computational methods can be applied reliably across all the pairs in BLOCKSIZE’s REAL-TIME-PRICES product, the timeframes considered are dynamically adapted to the number and volume of trade events that are observed for a given pair.

For this purpose, the timeframe for each pair, $\rho = [t - \tau, t]$, is dynamically, $\rho \equiv [t - \tau(t), t]$, attempting to fulfill the following conditions at all points in time:

1. There are $\mathrm{N} \geq 10$ events in the considered window, ρ

2. The aggregate volume of the events $\sum_j v_j$ in the window ρ is at least USD 1000 or an equivalent amount, if the quote currency is not USD
3. All included events in the window $\rho = [t - \tau(t), t]$ are less than 120 minutes old, i.e., $\tau(t) \equiv 120min. \forall\ t \geq 120min.$
4. There are at most 1000 trade events in the timeframe ρ

Conditions 1. and 2. are considered necessary the accuracy of statistical estimates used for the computation of the VWAP and the anomaly detection approaches described in the following Subsections 3.3. and 3.4.
Conditions 3. and 4. exclude data that is not sufficiently recent from the consideration for the REAL-TIME PRICES Data Feeds.

It is worthwhile to note that a timeframe $\rho = [t - \tau(t), t]$, due to the application of Conditions 3. and 4., may not fulfill the prerequisite Conditions 1. and 2. for the application of anomaly detection methods. This unlikely Edge case is discussed together with other edge cases in Subsection 3.5. .

3.3. Price-Based Filter Method

For the construction of the Price-Based Filter, the deviation of the price of each new trade event observation from the mean of the prices of all other non-anomalous trade events in the liquidity-adjusted timeframe for the computation of the VWAP is computed. This comparison is done using the average and the standard deviation of all prices in the VWAP window as follows.

Consider the trade event prices $\theta = \{p_1, p_2, \dots, p_N\}$ in the window $\rho = [t - \tau, t]$, i.e., $\theta = \{p_1, p_2, \dots, p_N\}$ with $t_i \in \rho$.

To decide if a new price p_{N+1} is anomalous or can be included, the standard deviation σ_θ and the average $\bar{p}_\theta$ of prices in θ are computed:

$$\theta = \{p_1, p_2, \dots, p_N\}$$

$$\bar{p}_\theta = \sum_{i=1,\dots,N} \frac{p_i}{N}$$

The confidence interval Ψ_θ for the inclusion of a new price p_{N+1} into the set θ of prices that are considered for the computation of the VWAP is obtained by adding/subtracting a multiple ξ of the standard deviation σ_θ from the average $\bar{p}_\theta$ of prices that are already in θ:

$$\Psi_\theta = [\ \bar{p}_\theta - \xi\sigma_\theta, \bar{p}_\theta + \xi\sigma_\theta]$$

If the new price event p_{N+1} is within the confidence interval Ψ_θ, it is accepted as valid and otherwise discarded as anomalous.

The value of the constant ξ determines the size of price deviations to be considered anomalous; BLOCKSIZE REAL-TIME price feeds found $\xi = 3.5$ a good choice to maintain exceptional quality standards.

3.4 Exchange-Based Filter Method

For each data source in REAL-TIME PRICES, BLOCKSIZE studies the deviation of its adaptive VWAP timeframe to the adaptive VWAP timeframe of all other data sources for the target price feed). For each data source, the confidence interval for its classification as anomalous is determined by two times the standard deviation of the combined snapshot sample excluding the data source in question.

Consider n data sources E that provide us prices $p_e(t)$ and volumes $v_e(t)$ where t is time and e is an exchange. BLOCKSIZE calculates within timeframe T data sources $E p_e(t)$ and volumes $v_e(t)$ where t is time and e is an exchange. BLOCKSIZE calculates within the adaptive timeframes collectively denoted as T

$$P^e_{VWAP} = \frac{\sum_{t \in T, e \in E} p_e(t) v_e(t)}{\sum_{t,e} v_e(t)}$$

the following VWAP that excludes exchange e' and the standard deviation of VWAP that excludes the e' timeframe T and the standard deviation of VWAP that excludes e' within timeframe T :

$$\bar{P}^{e'}_{VWAP} = \frac{\sum_{t \in T, e \in E, e \neq e'} p_e(t) v_e(t)}{\sum_{t \in T, e \in E, e \neq e'} v_e(t)}$$

$$\bar{\sigma}^{e'}_{VWAP} = \sqrt{\frac{\sum_{t, e \in E, e \neq e'} p_e^2(t) v_e(t)}{\sum_{t, e \in E, e \neq e'} v_e(t)} - \left(\bar{P}^{e'}_{VWAP}\right)^2}$$

Here, and in the following we use a bar $\bar{\cdot}$ to denote that an exchange has been excluded.

Then BLOCKSIZE calculates the distance between the VWAP of exchange E and the VWAP that excludes exchange e

$$d_e = P^e_{VWAP} - \bar{P}^e_{VWAP}$$

Based on the threshold t BLOCKSIZE performs the test τ_e

$$\tau_e = \begin{cases} 1 & d_e \in [-t\bar{\sigma}^e_{VWAP}, +t\bar{\sigma}^e_{VWAP}] \\ -1 & \text{otherwise} \end{cases}$$

If the test τ_e is positive, the trade data fits in the interval

$$[-t\bar{\sigma}^e_{VWAP}, +t\bar{\sigma}^e_{VWAP}]$$

and the exchange e is not an outlier. On the other hand, if τ_e is negative, the data is outside of the interval

$$[-t\bar{\sigma}^e_{VWAP}, +t\bar{\sigma}^e_{VWAP}]$$

, BLOCKSIZE considers the exchange e as an outlier. The data of such outlier exchanges is excluded from the provided data feed in alignment with the data contingency rules in Section 2. is positive, the trade data fits in the interval

$$[-t\bar{\sigma}^e_{VWAP}, +t\bar{\sigma}^e_{VWAP}]$$

and the exchange$e$ is not an outlier. On the other hand, ifτ_e

$$[-t\bar{\sigma}^e_{VWAP}, +t\bar{\sigma}^e_{VWAP}]$$

BLOCKSIZE considers the exchange$e$

3.5 Edge Cases and Applicability of Anomaly Detection Methods

The Price-Based Filter and the Exchange-Based Filter methods are applicable only when the confidence intervals they use for filtering anomalies are accurate. BLOCKSIZE empirically found that fulfilling Condition 1. -- $N \geq 10$ observed events -- and Condition 2. -- more than USD 1000 or equivalent in observed trading volume (cf. Subsection 3.2.) renders the performance of the Price-Based and Exchange-Based Anomaly Detection approaches in Subsections 3.3. and 3.4. .

In the rare case that Conditions 1. and 2. are not met, market conditions are considered anomalous as reliable confidence intervals for the validity of trade event observations cannot be established. In such anomalous situations, the Price-Based and Exchange-Based Filters are not applied and all available non-erroneous trade event observations are included for the computation of the VWAP.

In the rare event that the Exchange-Based Filter classifies multiple data sources within one timeframe as outliers, the filter is not applied and all non-erroneous trade event observations that were not excluded

by the Price-Based Filter approach are used for the calculation of the VWAP. The assumption is that abnormal market conditions can lead to a wider spread of the price distributions across different exchanges, breaking the threshold set to $t = 2\sigma$.

The Exchange-Based Filter for the detection of anomalies is applicable only for the case that data has been reported by three or more exchanges. In cases where fewer than three exchanges report trade data to BLOCKSIZE CONNECT, the total VWAP is computed from all reported trades not classified as anomalous by the Price-Based Filter.

Following rare abrupt price jumps to prices outside the confidence interval of the Price-Based Filter can trap the approach into excluding all following trade event observations as anomalous. For rectifying this flaw, where the method fails to update its confidence interval, a detector for price jumps was added to the Price-Based Filter approach. This detector checks if $N \geq 4$ outliers are detected in sequence above or in sequence below the confidence interval with $\geq 500\ USD$ or equivalent of aggregated volume. If these price jump conditions are met, the detector triggers a re-initialization of the confidence interval of the Price-Based Filter approach. The confidence interval is recomputed including the sequence of $N \geq 4$ trade events representing the sudden price action that have erroneously been classified as outliers. The re-initilization maintains the construction rules for the Adaptive Liquidity-Adjusted timeframe ρ as specified in Subsection 3.2. .

Appendix: Current Documents

The REAL-TIME PRICES Methodology described in here is part of the BLOCKSIZE CONNECT Manifest, which is describing how BLOCKSIZE is dealing with aspects of its BLOCKSIZE CONNECT suite of data subscriptions.

A.1 Current version of the BLOCKSIZE CONNECT Manifest:
https://www.blocksize.info/blocksize-connect/manifest/

A.2 Current list of supported instruments:
https://www.blocksize.info/blocksize-connect/instruments-realtime/

A.3 Current list of supported markets:
https://www.blocksize.info/blocksize-connect/markets-overview/

A.4 Current version of the REAL-TIME PRICES Methodology:
https://www.blocksize.info/blocksize-connect/manifest/real-time-prices-methodology/

Date: 2023-11-21