CCCAGG Index Methodology

CryptoCompare

2021

Contact: research@cryptocompare.com

Published by CC Data Limited trading as CryptoCompare

data.cryptocompare.com

27 January 2021

Table of Contents

1. Version History

Version	Date	Details
1	2017-07-01	Initial version
2	2017-11-20	Added review process description
3	2018-02-26	Outlier methodology
4	2018-08-30	Amendment
5	2019-09-12	Outlier methodology update
6	2019-09-27	Constituent exchanges selection methodology
7	2019-12-16	Mathematical representation amendments
8	2020-07-20	Static and Dynamic Index Definitions
9	2020-09-18	Amendment
10	2020-10-26	Merge Constituent Methodology with CCCAGG Methodology
11	2021-01-28	Backtesting additions

2. Introduction

2.1 **Index description**

The Crypto Coin Comparison Aggregated Index ("CCCAGG") refers to the real-time index calculation methodology, the purpose of which is to show the best price estimation for cryptocurrency traders and investors to value their portfolio at any time. CCCAGG is CryptoCompare's proprietary index calculation methodology for digital assets, based on 24-hour volume weighted average calculation, time-penalty factor and outlier methodology. It aggregates transaction data of more than 250 exchanges, using a 24 hour volume weighted average. The CCCAGG is calculated for each cryptocurrency in each market it is trading in (example: CCCAGG BTC-USD). A detailed list of available cryptocurrencies is available on request.

Cryptocurrencies such as Bitcoin, Ethereum, Monero, etc. are traded at various markets against multiple currencies including fiat currencies (USD, JPY, GBP, etc) and other cryptocurrencies. Depending on the market type (exchange or OTC), liquidity level, trading volume, transaction fees, and many other factors, a coin can be traded at different prices across different markets, and therefore making it difficult to know the value of a coin at a certain time.

2.3 **Index properties**

Calculation agent	CC Data Limited
Dissemination	Real-time and historical
Day close	12:00 am GMT
Methodology	24 hour volume weighted average with time penalty adjustment
Calculation days	Every day of the week including business holidays
Markets	All cryptocurrency markets

3. Definitions

Term	Definition
API	Application Programming Interface
CCCAGG	Crypto Coin Comparison Aggregate Index
Calculation Date	Any day for which a CCCAGG is published
Currency Pair	the quotation of two different fiat currencies or cryptocurrencies, with the value of one currency being quoted against the other.
Top-tier exchanges	Exchanges with grade AA, A, B and BB based on the Exchange Benchmark results. Grades can be found on cryptocompare.com.
Liquid pairs	Pairs with both base and quote market trading above 1 million USD a day on average for the last 30 days aggregated across CCCAGG markets, or above 5 million USD a day aggregated across all markets.
Price difference	To assess the price difference of an exchange pair compared to the rest of the exchanges trading the same pair, one will calculate a median price of all markets for a certain day. The price difference is the difference between an exchange's price and median price in percentage terms. The assumption is that for a trading pair with multiple exchanges, most liquid exchanges will cluster around the median. This metric is preferred over a simple average, as it can detect outliers without skewing the metric for the whole sample.
Price impact	Price impact is a metric to assess how much of the price difference defined above will materialise when added to the

	CCCAGG index. If we assume CCCAGG price to be close to median, the price impact will be a volume weighted price difference as follows: This is an important metric as a higher price difference on a low-volume exchange will materialise less, in certain cases, than a lower price difference price feed on a high-volume exchange. This is due to the fact that CCCAGG is using 24 hour volume weighted average calculation.
Liquidity factor	Liquidity factor is calculated as the fraction of exchange average daily volume compared to the CCCAGG average daily volume. This number can be greater than 100%. This metric shows how much liquidity it will contribute to the CCCAGG index, which will be a key metric for selecting constituents for illiquid pairs.
Average daily volumes	Average daily volumes for each pair are calculated on daily historical data, all converted into USD at the day's CCCAGG USD rate.
24 hour volume	The total volume of a specific currency pair on a certain exchange
Time penalty factor	A factor used for penalising outdated prices
Exchange Benchmark	CryptoCompare's proprietary methodology for assessing exchange quality
Monthly Constituent Selection	A monthly review process to establish which exchanges to be included or excluded for CCCAGG indices based on eligibility or exclusion rules

4. Data collection

4.1 Data source

Transactional data (historical trades) is collected from each exchange via public REST API polled every 2-5 seconds or websocket endpoints. All collected data will be standardized internally, stored and backed up in servers.

Exchanges and markets are added on an ongoing basis based on research or user request. Exchanges that do not meet the technical requirements (available API for transactional data) cannot be added to the data collection. Unlike many data providers, who use snapshot data, CCCAGG approach of using transactional data enables auditability and replicability.

4.2 Data format

The collected data consists of:
- Trade ID: string or numerical
- Timestamp: Unix timestamp in seconds
- Price: numerical
- Amount: numerical
- Position: buy/sell

4.3 Data validation

Each trade is validated for the following:
- Each field has the correct data format
- Price and amount is positive
- Timestamp is not in the future
- Trades are not duplicated

4.4   Failure of data retrieval

In the event of failure to retrieve data from an exchange (due to service outage of the exchange API service), per design of the CCCAGG indices, the last price of the respective exchange will expire over time (its weighting will decrease to close to zero). As long as the currency pair is trading on other exchanges, the CCCAGG calculation is uninterrupted.

If the missed data is recoverable, CryptoCompare makes its best effort to retrospectively back fill the data for historical accuracy. This might result in recalculation of certain CCCAGG pairs, therefore CryptoCompare is publishing two sets of indices: the static index, which is immutable, and the dynamic index, which can be adjusted retrospectively. The default index price retrieved from the CryptoCompare API is the dynamic CCCAGG. The real-time index dissemination is also the dynamic index.

4.4.1 Static Index

Immutable CCCAGG index which does not account for missed trades.

4.4.2 Dynamic Index

CCCAGG index that might be subject to retrospective back filling in an event of failure to retrieve data on a timely manner.

5. Methodology

5.1 Input data

CCCAGG is calculated every time a new transaction is received. The following input data is needed from each transaction:

- Trade price
- Trade amount
- Trade timestamp
- Exchange where the transaction was executed

Input data sources are also reviewed via the qualitative review method, the Exchange Benchmark. Find further details in chapter 6.2.

5.2 Constituent exchanges

Constituent exchanges are selected based on the Constituent Selection and Review process. Details can be found in Chapter 6.

5.3 Index calculation

5.3.1 24 hour volume

CCCAGG uses a 24 hour volume weighted average to calculate prices. 24 hour volumes are calculated solely based on transactional data. This ensures CCCAGG gives greater weight to liquid market prices, and price impact of illiquid (and therefore more volatile) markets are negligible.
The 24 hour volume is calculated as the sum of the volume over the last 23 hour and the cumulative volume of the current hour.

5.3.2 Time penalty factor

A time penalty factor is added to ensure that exchanges that suspend trading have an expiring price impact. An example of a case where this methodology showed its advantage was the Bitfinex hack in 2016.

Bitfinex had one of the highest trading volumes in Bitcoin, therefore had a significant weight in most price indices. Therefore, when trading was suddenly suspended on Bitfinex, causing a crash on all other markets, most indices still showed a Bitcoin price close to the last price on Bitfinex, although markets have already moved on.

CCCAGG takes last trade time into account, therefore the last Bitfinex price expired with time and the index could move with the market.

5.3.3 Aggregation over trading currency

CCCAGG only takes direct trading pairs into consideration for calculation. For example CCCAGG BTC-USD only accepts trades from exchanges trading BTC-USD directly, therefore no conversion is needed for the aggregated index calculation.

The reason for this methodology is that a coin can trade on multiple currency markets with a significant price difference (premium or discount), therefore aggregating across all markets will result in an average price that is not useful for a trader or investor who holds a crypto position in a certain currency and will most likely trade in that currency.

5.3.4 Mathematical representation

For simplicity, we define the CCCAGG Index for a pre-specified currency pair ABC-XYZ, all notations below refer to this pair.

Variable	Explanation
i, j	Exchange
n	Number of all exchanges
t, s	Unix timestamp in seconds
P_t^i	Price at exchange i and time t
P_t^*	CCCAGG Index price at time t
W_t^i	Index weight of exchange i at time t
V_t^i	24 hour volume of exchange i at time t
x	Boolean, is excluded from pricing, x = {0,1}
γ_t^i	Time penalty factor
h_t	Last hour timestamp at time t
τ_t^i	Time since last trade in minutes on exchange i at time t
δ_t^i	Price deviation of trade on exchange i at time t
A	Price deviation threshold

At any time *t,* the Index price for a currency pair is the weighted average of prices on all exchanges.

$$P_t^* = \sum_i P_t^i W_t^i$$

A weight belonging to an exchange is the ratio of the 24 hour trading volume of a currency pair to the total volume of the pair on all exchanges.

$$W_t^i = \frac{V_t^i 1_{\{x=0\}} \gamma_t^i}{\sum_k V_t^k 1_{\{x=0\}} \gamma_t^k}$$

The 24 hour volume of an exchange at a certain time is defined as the sum of trading volume of the last 23 full hours and the cumulative volume of the current hour. In the formula below, 82800 is the number of seconds for 23 hours and h is the timestamp for the last hour (for example, if t is 14.34 then h is 14.00).

$$V_t^i = \sum_{s=h_{t-82800}}^{t} V_s^i$$

The exchange volume is adjusted with the time penalty factor, which decreases with the increasing time since the last trade.

$$\gamma_t^i = \begin{cases} 1 & \text{if } \tau_t^i \le 5 \\ 0.8 & \text{if } 5 < \tau_t^i \le 10 \\ 0.6 & \text{if } 10 < \tau_t^i \le 15 \\ 0.4 & \text{if } 15 < \tau_t^i \le 20 \\ 0.2 & \text{if } 20 < \tau_t^i \le 25 \\ 0.001 & \text{otherwise} \end{cases}$$

5.3.5 Outlier Detection

Aside from our qualitative based assessment on exchange suitability (discussed in Chapter 6), we also utilise real-time outlier detection mechanisms to exclude certain exchanges. This plays a crucial role in maintaining an accurate index, because any outliers can have a sizable impact on the CCCAGG value.

For a trade to be considered an outlier, it must deviate significantly from the previous index price. This kind of deviation could occur for a number of reasons - for example:

- Low liquidity on a particular instrument
- Erroneous data from an exchange
- Incorrect mapping of an instrument

The following outlier handling logic is implemented:

1. For one or two exchanges - no exclusion logic is applied. Everything is considered an inlier and included.
2. For three or more exchanges the following logic applies.
    a. Any trade that deviates by more than a certain threshold from the previously calculated index price (either up or down) causes the exchange to be labelled as an outlier and excluded temporarily.
    b. Excluded exchanges are automatically re-included in the Index if the Index average price moves to a point where the exchange price is no longer considered to be an outlier.
    c. Excluded exchanges result in exclusion of price weighting as well as volume.
    d. There can be a price change when an exchange is excluded due to recalculation of price with the smaller number of exchanges and different volume levels.
    e. Existing price decay and 24 hour window for inclusion in calculation are not impacted by this additional logic.

The deviation is calculated for each new trade in the following way:

$$
\delta_t^i = \begin{cases} \frac{1}{\frac{P_t^i}{P_{t-1}^*}} - 1 & \text{if } P_t^i < P_{t-1}^* \\ \frac{P_t^i}{P_{t-1}^*} - 1 & \text{if } P_t^i \geq P_{t-1}^* \end{cases}
$$

If δ_t^i exceeds a threshold, a trade is deemed an outlier. The threshold, A is currently set as A=3.

$$
x_t^j = \begin{cases} 0 & \text{if } \delta_t^j < A \\ 1 & \text{if } \delta_t^j \geq A \end{cases}
$$

This means that trades which exceed 400% or drop below 25% of the previous CCCAGG Index price will be deemed an outlier.

An edge case can exist where a split in the market occurs between two distinct groups of exchanges. If Group 1 exchanges trade at a significant price different to be included as outliers, only price and volume information from Group 2 will be included.

In this situation, if the number of exchanges in Group 1 (outliers) exceeds the number of exchanges in Group 2 (inliers), an inlier/outlier flip occurs with the previously included exchanges in Group 2 becoming excluded and the previously excluded Group 1 exchanges becoming included.

When an exchange is re-included - only current and future pricing data will reflect the re-inclusion. No back-filling or recalculation occurs for previously reported pricing in the Index.

5.4   Auditability and replicability

CCCAGG is auditable and replicable since its calculation is based on transaction data retrieved from exchanges via public API. Anyone who has access to this data can recreate the CCCAGG indices.

6. Constituent Selection Methodology

6.1 Introduction

CryptoCompare has integrated with a list of exchanges, but only a subset of them count towards the calculation of CCCAGG.

While CryptoCompare strives to include as many exchanges as possible after a testing period, exclusion generally happens in the following cases:

- Volatile prices compared to market average (OTC markets excluded)
- Exchange suspends trading activity
- Malfunctioning API (from constituent exchange)
- Exclusion to due risk factors identified in Exchange Benchmark (see more information in section 6.2)

The constituent selection methodology consists of a qualitative component, using the CryptoCompare Exchange Benchmark, and a quantitative component, the Monthly Constituent Selection process, looking at historical price movements and liquidity.

Review component	Method	Frequency
Exchange Benchmark	Qualitative	Semi-annual
Monthly Constituent Selection	Quantitative + using Exchange Benchmark results	Monthly

The Exchange Benchmark is conducted as a due diligence process on exchanges, it is updated semi-annually and its results are used in the Monthly Constituent Selection process.

The Monthly Constituent Selection and Review Process is conducted to decide on whether an exchange should be included or excluded from CCCAGG. Each exchange is tested for eligibility requirements to be included in the aggregation on a currency pair basis.

When assessing the eligibility of an exchange to be included/excluded in each CCCAGG index, CryptoCompare differentiates between the most liquid pairs (liquid pairs) and other pairs. This is due to the fact that, for liquid pairs, the main goal is to maintain price integrity, while for illiquid pairs, the goal is to give the best price discovery. The difference in incentives means that the eligibility rules also differ.

While the review will take the Benchmark grades into consideration, a top grade does not automatically determine inclusion. Historical prices are used to assess the price differences for each trading pair by calculating differences to median price. Moreover, further metrics will be used for the assessment, such as impact and liquidity factors.

Each case for exclusion is reviewed by the Technical Committee with sign off by the Oversight Function. This process occurs every calendar month, and on an ad-hoc basis when necessary. An excluded exchange can be re-included if the Technical Committee finds that the problem causing the market disturbance has been solved. Changes in constituent exchanges are communicated via a published report every 15th of the calendar month on data.crytocompare.com, and implemented on the 20th of the calendar month (adjusted if not a business day).

6.2 Qualitative Review: Exchange Benchmark

The CryptoCompare Exchange Benchmark seeks to bring clarity to the digital asset exchange sector by providing a framework for assessing risk, bringing transparency and accountability to a complex and rapidly evolving market. This is approached in several dimensions using a comprehensive data set, covering 160 exchanges across 8 categories of evaluation:
- Legal/Regulation
- Data Provision
- Security
- Team/Exchange
- Investment
- Trade Monitoring
- Market Quality

- Negative Events Penalty

The output of the Exchange Benchmark framework is a rating system, whereby each exchange receives a grade from AA (best quality) to E (lowest quality). This framework is used as a qualitative assessment for the eligibility review.

Benchmark results are updated semi-annually and published on cryptocompare.com. The full methodology can be found here: data.cryptocompare.com/research.

For the purposes of constituent selection, Top-Tier exchanges are those which receive an AA, A, BB or B in the Exchange Benchmark and Non Top-Tier exchanges are those receiving a C,D, or E. Exchanges in the Top-Tier meet our minimum threshold for acceptable risk.

E or non graded exchanges are automatically excluded from all CCCAGG.

The Exchange Benchmark is updated and published semi-annually.

The Exchange Benchmark results are used in the constituent selection process as described in 6.3.

6.3 Quantitative Review: Monthly Constituent Selection

6.3.1 Overview

Constituent selection is divided into two main components: review of liquid pairs and review of illiquid pairs. This is an important distinction, as the incentives of review for the two groups are different. For the liquid pairs, the review prioritizes price integrity, so top-tier exchanges are preferred with strict rules for inclusion. For other pairs, the incentive is to provide the best price discovery, therefore eligibility rules will be more relaxed.

This process is conducted once every calendar month, aggregating the data of the last 30 days.

The following metrics are calculated for each pair for each exchange before conducting the selection process:

Price difference: The price difference is the difference between an exchange's price and the median price of all exchanges in percentage terms.

Price impact: If we assume CCCAGG price to be close to median, the price impact will be a volume weighted price difference as follows:
Price_impact = price_difference * average_daily_volume_exchange / average_daily_volume_cccagg

Liquidity factor: Liquidity factor is calculated as the fraction of exchange average daily volume compared to the CCCAGG average daily volume.

Average daily volume: Average daily volumes for each pair are calculated on daily historical data, all converted into USD at the day's CCCAGG USD rate.

6.3.2 Liquid pairs

Liquid pairs are calculated based on average daily volumes for each base and quote asset that reach a certain threshold.

The following reviews are done on the pair-by-pair basis: top-tier exchanges to include, all exchanges to exclude and other exchanges to include.

6.3.2.1 Top-Tier Exchanges To Include

For inclusion review of top-tier exchanges that have not been included in certain markets, CryptoCompare uses 30-day average price difference metric and 30-day average price impact. Exchanges with low price difference and low price impact will be added to the index. Both of the following thresholds need to be reached for the inclusion as show the table below:

Metric	Threshold	Condition
30-day average price difference (absolute)	2%	Less
30-day average price impact (absolute)	10%	Less

6.3.2.2 Exchanges to Exclude

Exclusion tests are done with all exchanges including top-tier exchanges. This is important for liquid pairs as the goal is to maintain price integrity. The metrics used for the exclusion test are 30-day average price difference metric and 30-day average price impact. Exchanges with high price difference or high price impact will be excluded from the index as below:

Metric	Threshold	Condition
30-day average price difference (absolute)	10%	Greater
30-day average price impact (absolute)	50%	Greater

Exclusion thresholds are higher than inclusion thresholds as we need to account for periodical differences due to the overall market liquidity seasonalities.

6.3.2.3 Non Top-Tier Exchanges To Include

All other exchanges that are non top-tier are also reviewed for inclusion. The metrics used for this review are 30-day average price difference metric, liquidity factor, and daily average volume. First, exchange pairs need to meet a minimum trading activity threshold to avoid stale prices. Once that is met, the price difference needs to meet a certain level. It is also important that the liquidity ratio is reasonably high, adding price liquidity to the index. Exchanges graded D or below or non-graded will not be included. All of the following conditions need to be met for inclusion:

Metric	Threshold	Condition

30-day average price difference (absolute)	10%	Less
30-day average daily volume	5 million USD	Greater
Liquidity factor	50%	Greater

6.3.3 **Illiquid pairs**

Any pair that is not in the list of liquid pairs are reviewed as 'illiquid pairs'. The main motivation for this review is to provide the best price discovery possible. These pairs are less liquid, therefore data sources are scarce. As of writing this report there are around 14 thousand other pairs from 30 thousand exchange pairs. This means in average an illiquid pair has 2 exchanges.

6.3.3.1 Dry Pairs (Less than 4 exchanges)

Pairs with less than 4 exchanges are called 'dry pairs,' and will allow all price feeds to be included (up to 3 exchanges), unless price difference is an order of magnitude high (10x). The safety net for including all exchanges up to 3 exchanges in the CCCAGG outlier methodology which automatically excludes a 3x difference. A minimum of 3 exchanges is required for the outlier methodology to work.

6.3.3.2 Top-Tier Exchanges To Include

All top-tier exchanges are reviewed for illiquid pairs that are not dry pairs. Although top-tier exchanges are trusted for their data quality, they may have illiquid markets too as they launch new products. The following metrics are used for the review: 30-day average price difference and liquidity ratio.

Metric	Threshold	Condition
30-day average price difference (absolute)	5%	Less
Liquidity factor	50%	Greater

7. Ongoing Maintenance

7.1 Methodology Review and any Changes to Methodology

The methodology is reviewed at least every quarter by the Technical Committee to ensure that it remains representative of the relevant market or economic reality that it is intended to measure. If the Technical Committee requires any material changes to the methodology, any change must be signed off by the Oversight Function before entering a period of public consultation of no less than thirty (30) days. The Oversight Function shall undertake an internal review of the methodology at least annually.

In accordance with Article 13(1)(c) of BMR, the consultation exercise provides notification to users, at a minimum, of the key elements of the methodology that would be affected by the proposed material change. CCCAGG consumers will be notified of the methodology consultation and proposed changes via the API newsletter and other direct client communication channels, and the proposed methodology amendments will be available for review at data.cryptocompare.com/research, where users will be able to submit feedback on any proposed change.

7.1.1 Backtesting and Benchmarking

In order to maintain confidence that the CCCAGG index is representative and replicable, each quarter the following tests are conducted:
- Compare daily CCCAGG values for the last 3 months with the median price of the constituent exchanges. It is expected that CCCAGG follows the market median closely
- Compare daily CCCAGG volatility to exchange volatility in the last 3 months. It is expected that CCCAGG is less volatile than each individual exchanges
- Recalculate daily CCCAGG values using raw trade data for the last 3 months. This ensures that the index is replicable and transparent.

7.2 Constituent Exchange Review

Constituent exchange selection is reviewed by the Technical Committee at least once every calendar month or on an ad-hoc basis, when market or technical events require. Market or technical events may include:

- Suspended Trading
- False Data Provision
- Service Outage

The Exchange Benchmark, which informs the qualitative aspects of the Constituent Exchange Selection and Review occurs on a semi-annual basis and is reviewed and updated no less than twice annually.

The constituent exchange selection and review methodology, outlined in section 6 of this document, describes the process and criteria by which constituent exchanges are selected and considered toward the calculation of CCCAGG.

7.3 Discretion Regarding the Use of Input Data

Pursuant to Art. 12 No.1. (b) of the BMR, CryptoCompare has established the following rules identifying how and when discretion may be exercised in the administration of an index.

In cases where input data is or appears to be qualitatively inferior or different sources provide different data, or a situation is not covered by the index rules, CryptoCompare may use or change the data at its own discretion according to the following discretion policy after a plausibility check. This may include

- Liquidity and size data,
- Event information,
- Classifications and other secondary data.

Any changes to input data that CryptoCompare intends to apply because of missing data, different data from different sources, or other information concluding the inappropriateness or

incorrectness of data must be subject to reasonable discretion. The decision on any change must be required, appropriate, commensurable, and in line with the respective index scope and objective and must reasonably consider in a balance weight the interest of users, investors in related products and the integrity of the market.

The Technical Committee ensures consistency in the use of discretion in its judgement and decision. Employees involved in the Technical Committee must have shown the respective experience and skills. Significant decisions are subject to sign-off by a supervisor. In case of material changes to data the relevant situation will be analysed in detail, described and presented to the Oversight Function and discussed and reviewed with the Oversight Function.

The broad range of possible data quality problems does not allow to define specific steps for each possible instance. CryptoCompare will always weigh the different interests of the index users, the integrity of the market and other involved parties and determine the least disadvantageous measure that equally considers the relevant interests best.

In order to avoid individual decisions on the use of data in similar cases for the future, an update of the index rules can be taken into consideration if applicable. Other possible mitigation measures may include the change of input data sources or providers and/or own data research where possible and reasonable.

Records are kept about material judgement or discretion and will include the reasoning for said judgement or discretion.

7.4 Potential Limitations

CCCAGG methodology works best for liquid markets where multiple exchanges provide data sources for the index. If the number of exchanges is low or each individual exchange is illiquid, the volume weighted average price will give an indicative value that might not sufficiently reflect the market.

The CCCAGG index also relies on frequent trade updates. If no trading occurs on a market for more than 24 hours, the index will become stale.

8. Dissemination

CCCAGG indices are disseminated via REST API and Websocket API. Read the API documentation at https://min-api.cryptocompare.com/documentation.

9. Disclaimer

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher, except in the case of brief quotations embodied in critical reviews and certain other non-commercial uses permitted by copyright law.

CC Data Limited is a registered company in England with company number 10966788 and registered address of 162 Main Road, Danbury, Chelmsford, England CM3 4DT

10. Contact

CryptoCompare

13 Charles II Street

London, SW1Y 4QU

United Kingdom

data.cryptocompare.com

research@cryptocompare.com