MORNINGSTAR

CRSP US Stock Databases Calculations & Index Methodologies Guide

July 2026

Table of Contents

MORNINGSTAR

Chapter 1: Introduction

Center for Research in Security Prices (CRSP) was established in 1960 as a research center at the University of Chicago to support Professor Lawrence Fisher and James Lorie's seminal research paper, "Rates of Return on Investment in Common Stocks." Since that time, CRSP has been curating data for scholarly researchers and investment practitioners. From time-to-time CRSP makes refinements to its databases when new incremental data becomes available. To incorporate new data and enhance the user experience, CRSP developed and in 2022 released a new US Stock & Indexes Database Flat File Format 2.0 (CIZ).

These changes were motivated by CRSP's commitment to provide subscribers with the most complete and accurate data available by reflecting additional data from new sources. This also allowed CRSP to reconsider how best to present monthly data derived from daily data, which is now complete back to 1925. The new data allows for more accurate and precise monthly stock and index return calculations which are provided in CIZ. These return calculation changes were announced in 2022 when CIZ was introduced to subscribers. CIZ is available in SAS, ASCII and R. The December 2024 data which was released in January 2025, was the last CRSP US Stock & Indexes release produced in the Legacy (FIZ) and Flat File Format 1.0 (SIZ) formats.

This guide contains formulas and methodologies used to derive CRSP variables in the stock and index files and generated by the CRSP data utilities.

Chapter 2: Index Calculations

Adjusted Data

Price, dividend, shares, and volume data are historically adjusted for split events to make data directly comparable at different times during the history of a security. CRSP provides unadjusted price, dividend, shares, and volume data. Cumulative adjustment factors can be used to adjust data at different times to the same basis.

An adjustment base date is chosen as the anchor date. All data on this date are unadjusted, and other data is converted based on the split events between the base date and the time of that data. The adjustment base date is usually chosen to be the last available day of trading.

Split events always include stock splits, stock dividends, and other distributions with price factors such as spin-offs, stock distributions, and rights. Shares and volumes are only adjusted using stock splits and stock dividends. Split events are applied on the Ex-Distribution Date.

Price and dividend data are adjusted with the calculation:

$A(t) = \mathrm{P(t)} \,/\, \mathrm{C(t)},$

where A(t) is the adjusted value at time t, P(t) is the raw value at time t, and C(t) is the cumulative adjustment factor at time t.

Share and volume data are adjusted with the calculation:

$A(t) = \mathrm{P(t)} * \mathrm{C(t)},$

where A(t) is the adjusted value at time t, P(t) is the raw value at time t, and C(t) is the cumulative adjustment factor at time t.

In both cases, where C0 is the adjustment base date, the cumulative adjustment factor is:

if $t = C_0$,

$$C(t) = 1.0$$

if $t > C_0$ and no split events since t-1,

$$C(t) = C(t-1)$$

if $t > C_0$ and a split event with factor f since t-1,

$$C(t) = C(t-1) * f$$

if $t > C_0$ and split event change

$$C(t-1)/f$$

if $t < C_0$ and a split event change

$$C(t+1) * f$$

where factor is typically the Factor to Adjust Price variable + 1

If there is a gap in trading where possible split events are not known, all adjusted values are set to missing when the gap is between the observation and the adjustment base date.

Monthly: If monthly summary data (id or Low Price, Ask or High Price, and Volume Traded) are adjusted, the adjustment factor cannot take into account adjustments that take place in the middle of the month. Therefore, the result assumes all

adjustment events occur on the last trading day of the month. A more accurate monthly adjusted value can be derived by adjusting and re-summarizing the underlying daily data.

Annualized Return

Annualized Return is the constant annual return applied to each period in arrays that would result in the actual compounded return over that range. An Annualized Return is a special case of a Geometric Average Return where the time periods are expressed in terms of years.

Associated Portfolio Return

Associated Portfolio Returns are a composite of a group of portfolio index series based on a time-dependent portfolio assignment for a security. They are built for each security based on assignments within the specified portfolio type. The associated portfolio return at any time is the return of the portfolio to which the security belongs at that time. If the security is not assigned to a portfolio of that type at the time, the associated portfolio return is set to a missing value.

Cumulative Return

A Cumulative Return is a compounded return from a fixed starting point. Each period in a time series of Cumulative Returns contains the compounded return from the first period in the time series to the end of that period.

Delisting Return

Delisting Return is the return of security after it is delisted. It is calculated by comparing a value after delisting against the price on the security's last trading date. The value after delisting can include a price on another exchange or the total value of distributions to shareholders. If there is no opportunity to trade a stock after delisting before it is declared worthless, the value after delisting is zero. Delisting Returns are calculated similarly to total returns except that the value after delisting is used as the current price.

Valid delisting payment information is either a valid price with at least a bid and ask quote within ten trading periods, or a complete set of payments received for the shares. If information after delisting is insufficient to generate a return, a missing value is reported.

Monthly: The monthly Delisting Return is calculated from the last month ending price to the last daily trading price if no other delisting information is available. In this case the delisting payment date is the same as the delisting date. If the return is calculated from a daily price, it is a partial-month return. The partial-month returns are not truly Delisting Returns since they do not represent values after delisting, but allow the researcher to make a more accurate estimate of the Delisting Returns. The change to use daily compounded returns rather than monthly also applies to delisting returns (dlret).

When valuing a portfolio, the Delisting Return or other representation can be used to assign a value to the delisted security. The researcher must decide whether to assign alternate estimated values based on the Delisting Code when delisting payment information is unavailable. If using monthly data and an alternate estimate for Delisting Return is used, partial month returns should also be adjusted by this factor.

Daily Ordinary Dividend Amount (DlyOrdDivAmt) & Daily Non-Ordinary Dividend Amount (DlyNonOrdDivAmt)

The dividend amounts are the cash amounts used in a holding period return time period to calculate returns. They are an adjusted summation of all distribution cash amounts available in the distribution history with Ex-distribution dates after the previous period and up to and including the current period, adjusted to the basis at the end of the previous period.

The dividend amounts include non-ordinary and ordinary types. Non-ordinary dividends include return of capital distributions. Ordinary dividends are excluded from capital appreciation returns calculations.

**TS_Print item names are Odivamt and TDivamt*

Excess Return

An Excess Return is defined as the return in excess of a comparable benchmark. The benchmark can be a single associated index series or a composite of a group of portfolio index series based on security and time-dependent portfolio assignments.

If an Excess Return is based on a single index series, the Excess Return for a period is

$$E(t) = R(t) - I(t),$$

where

$E(t)$ is the Excess Return at time t

$R(t)$ is the security return at time t

$I(t)$ is the index return at time t.

If the security return R(t) is based on a previous price t' that is not the previous time period, I(t) is the compounded index return from t' + 1 to t.

If an Excess Return is based on associated portfolios, the Excess Return for a period is

$$E(t) = R(t) - I(p(t), t)$$

where

$E(t)$ is the Excess Return at time t

$R(t)$ is the security return at time t

$p(t)$ is the portfolio assignment of the security at time t

$I(p(t), t)$ is the return of that portfolio at time t.

If the security return R(t) is based on a previous price t' that is not the previous time period, I(p(t),t) is the compounded return of the security's portfolio return from t' + 1 to t. If the security is not assigned a portfolio assignment of the given type at time t, E(t) is set to a missing value.

When cumulating Excess Return, the security returns and the index returns are cumulated separately before subtracting the difference.

Daily Factor To Adjust Prices In Period (DlyFacPr)

Factor to Adjust Prices in Period is the amount the current price is multiplied by in returns calculations so that current and previous prices are on the same split-adjusted basis. Factor to Adjust Prices in Period is derived from the Factor to Adjust Price field of distributions with Ex-Distribution Dates after the previous period and up to and including the current period. In simple stock splits, Factor to Adjust Prices in Period is distribution Factor to Adjust Price plus one.

**TS_Print item name is Facprc*

Geometric Average Return

A Geometric Average Return is the constant return applied to each period in a range that would result in the compounded return over that range.

The Geometric Average Return is calculated using the formula below:

$$g_n = (1 + r_c)1/\mathrm{n} - 1$$

where

g_n = the Geometric Average Return applicable on each subset period n

r_c = the cumulative return over the entire period

n = the number of equal subset periods to average the return

Income Return

Income Return is the return on the ordinary dividends paid to shareholders of a security. It is the ratio of the amount of ordinary dividends since the end of the previous period up to and including the end of the period of interest to the price at the end of the previous period. It is similar to a dividend yield.

Income Return is calculated by CRSP as the difference of the Total Return and Capital Appreciation Return, as follows.

$$DlyIncRet_t = DlyTotRet_t - aret_t$$

where

$DlyIncRet_t$ is the income return for time t

$DlyTotRet_t$ is the total return for time t,

$DlyPrcRet_t$ is the capital appreciation return for time t

Index Count

Index Count is the number of securities in the portfolio during a period (period type can be daily or monthly). Rules are based on the specific index or portfolio methodology. See Total and Used Counts for more details.

Index Level

Index Level is the value of an investment relative to its value at one fixed point in time. Index Levels allow convenient comparison of the relative performance of the different portfolios or asset classes. Differences arise when indexes are based on different underlying databases such as daily and monthly CRSP stock products.

The initial date and value are set arbitrarily, but must be consistent if comparing multiple indexes. The Index Level for any series at any time before the initial point, indicates the value invested at that time that will result in the initial value at the initial point. The Index Level of a series is set to missing prior to the period before its first available return.

Let:

I_t = Index Level for any series at time t

R_t = return for the period t-1 to t

F = First Return. The time of the first non-missing return of the series

D = Initial Date. An arbitrary date where the level is set to the initial value

L = Initial Level. An arbitrary value the level is set to on the initialization date

then

if t = D, then $I_t = L$

if t > D, then $I_t = I_{t-1} * (1 + R_t)$

if t < D, then $I_t = \frac{I_{t\text{-}1}}{1 + R_{t+1}}$

if t-1 < F then I_t is set to missing– Note: Missing values are file format specific.

Defined CRSP indexes use the following initial dates and levels:

CRSP Stock File Indexes	
Initial Level	100.00
Initial Date	December 29, 1972
CRSP Cap-Based Portfolios	
Initial Level	1.00
Initial Date	December 31, 1925
CRSP US Government Treasury and Inflation Indexes	
Initial Level	100.00
Initial Date	December 29, 1972

Publicly available indexes such as for the S&P 500 Composite and NASDAQ Composite have initial values set by their creators and differ from the CRSP initializations.

Index Return

An Index Return is the change in value of a portfolio over some holding period. The return on an index (R_t) is calculated as the weighted average of the returns for the individual securities in the index:

$$R_t = \frac{\sum(w_{i,t} * r_{i,t})}{\sum(w_{i,t})}$$

where:

R_t is the index return

$w_{i,t}$ is the weight of security i at time t

$r_{i,t}$ is the return of security i at time t

In a value-weighted index, the weight ($w_{i,t}$) assigned is its total market value; see Index Weight below. In an equally-weighted index, the weight is equal and by convention $w_{i,t}$ is set to one for every stock. Such an index would consist of n stocks, with the same dollar amount invested in each stock.

The security returns can be total returns or capital appreciation (returns without dividends). This determines whether the index is a total return index or a capital appreciation index.

In an index where the individual components are not known, but an index level is available from an external source, such as the Standard & Poor's 500 Composite Index, the return R_t is calculated as follows:

$$R_t = \frac{I_t}{I_{t-1}} - 1$$

where

R_t is the index return for time t

I_t is the index level at time t

I_{t-1} is the index level at end of the previous period (time t-1)

Index Weight

The weight of an index for a time period is the total market value of the securities in the index at the end of the previous trading period.

$$V_t = \sum(w_{i,t}) = \sum(v_{i,t})$$

where:

$$v_{i,t} = p_{i,t-1} * s_{i,t-1}$$

in which:

$v_{i,t}$ is value of security i at time t

$p_{i,t-1}$ is the price of security i at the end of the previous trading period (time t-1).

$s_{i,t-1}$ is the number of shares outstanding of security i at the end of the previous trading period (time t-1).

Market Capitalization

Market Capitalization (in 1000s) is a measurement of the size of a security defined as the price multiplied by the number of shares outstanding. CRSP uses the closing price or the absolute value of the bid/ask average from the Price or Bid/Ask Average variable and the applicable shares observation from the Shares Outstanding Observation Array for each calendar period to calculate Market Capitalization.

Rebasing Index Levels

It is possible to rebase an index to make index levels of two index level series comparable. To rebase an index, choose a new initial date and value, find the current index level on the new initial date, and multiply the levels on all dates by the new initial value divided by the old initial date index level:

$$N_t = I_t * \frac{L}{I_D}$$

where:

I_t = Original Index Level for the series at time t

N_t = New Index Level for the series at time t

D = New Initial Date

I_D = Original Index Level for the series on the new initial date

L = New Initial Level

Return

A Return is the change in the total value of an investment in a security over some period of time per dollar of initial investment. Total Return is the Holding Period Total Return for a sale of a security on the given day, taking into account and reinvesting all distributions to shareholders. It is based on a purchase on the most recent time previous to this day when the security had a valid price. Usually, this time is the previous calendar period, but may be up to ten calendar periods prior to the calculation.

Returns are calculated as follows:

$$r(t) = \frac{p(t)f(t) + d(t)}{p(t')} - 1$$

For time t (a holding period), let

t' = time of last available price < t

r(t) = return on purchase at t', sale at t

p(t) = last sale price or closing bid/ask average at time t

d(t) = dividend amount for t

f(t) = factor to adjust price in period t

p(t') = last sale price or closing bid/ask average at time of last available price < t

t' is usually one period before t, but t' can be up to ten periods before t if there are no valid prices in the interval. If there is a trading gap with unknown status between t and t', the previous price is considered invalid.

In daily databases, dividends are reinvested in the security on the Ex-Distribution Date. In monthly databases, the returns are holding period returns from month-end to month-end, not compounded daily returns, and dividends are reinvested in the security at month-end.

The Factor to Adjust Prices in Period is derived from the distribution history Factor to Adjust Price using all distributions with Ex-Distribution dates after the previous period and up to the end of the current period. The dividend amount is derived from the distribution history Dividend Cash Amount and Factor to Adjust Price in the same range. For example, if a 2-for-1 split is the only distribution event in the time range, Factor to Adjust Price is 1.0, Factor to Adjust Prices in Period is 2.0, and Dividend Cash Amount is 0.0. If a one dollar dividend is the only distribution event in the time range, both Dividend Cash Amount and dividend amount are 1.0.

A series of special return codes specify the reason a return is missing:

Code	Description
NA	Not Applicable
MP	Missing Price
NT	Not Tracked
GP	Gap Between Prices Too Large
NS	New Security
MV	MIssing Corporate Action Value
DM	Delisting Price/Amount Missing
DG	Delisting Price is more than 10 PERIODs from the...
RA	Return after Not Tracked PERIOD
DP	Delisting Pending

New calculation summary flags allow efficient filtering and grouping of data. DlyDistRetFlg (Daily Distribution Return Flag) provides information about the summary distribution values so that returns with specific categories of distributions (e.g., No distribution, ordinary cash only, stock split only, etc.) can be easily identified for further analysis. DlyRetDurFlg (Daily Return Duration Flag) provides information about the number of periods and days between the prices used to calculate the daily return, whether there were any missing prices, and if there were no missing prices, whether they were adjacent days or intervening weekends or holidays.

Scholes-Williams Beta

Beta is a statistical measurement of the relationship between two time series, and has been used to compare security data with benchmark data to measure risk in financial data analysis. CRSP provides annual betas computed using the methods developed by Scholes and Williams (Myron Scholes and Joseph Williams, "Estimating Betas from Nonsynchronous Data," Journal of Financial Economics, vol 5, 1977, 309-327).

$$\beta_i = \frac{\sum(lr_{i,t} * M3_t) - \left(\frac{1}{n_i}\right) * (\sum lr_{i,t}) * (\sum M3_t)}{\sum(lM_t * M3_t) - \left(\frac{1}{n_i}\right) * (\sum lM_t) * (\sum M3_t)}$$

where:

β_i is the Beta for security i for the year being calculated

$r_{i,t}$ is the return of security i at day t

$lr_{i,t}$ = ln(1+$r_{i,t}$) is the natural log of the return of security i at time t+1 or the continuously compounded return

M_t is the value-weighted market return at time t

lM_t = ln(1+M_t) is the natural log of the value-weighted market return at time t+1 or the continuously compounded return

$M3_t = lM_{t-1} + lM_t + lM_{t+1}$ is the three-day moving window of the above market return

n_i is the number of non-missing returns for security i during the year

where the summations are over t and include all days on which security i traded, beginning with the first trading day of the year and ending with the last trading day of the year. There are two index families based on Scholes-Williams Beta calculations: NYSE/NYSE American and NASDAQ-only.

In the NYSE/NYSE American family, only trading prices are considered in the beta calculation, and a security must have traded half the days in a year to be given a non-missing beta for that year. The index used in the calculation is the total returns on the Trade-only NYSE/NYSE American Value-Weighted Market Index.

Betas for the NASDAQ family do not use the standard Scholes-Williams trade-only data restriction, since most NASDAQ securities were not required to report transactions until 1992. Removing bid/ask averages would restrict NASDAQ data to only NASDAQ National Market securities after 1982. NASDAQ returns based on bid/ask averages have different characteristics from trade-based returns, and betas are provided for comparison. NASDAQ betas are based on the total returns on the NASDAQ Value-Weighted Market Index.

Standard Deviation

Standard Deviation is a statistical measurement of the volatility of a series. CRSP provides annual standard deviations of daily returns using the following calculations:

$$\sigma_i = \sqrt{\frac{\sum(r_{i,t}{}^2) - \left(\frac{1}{n_i}\right) * \left(\sum r_{i,t}\right)^2}{n_i - 1}}$$

where:

σ_i is the standard deviation for security i for the year being calculated

$r_{i,t}$ is the return of security i at time t

n_i is the number of non-missing returns for security i during the year

where the summations are over t and include all days on which security i had a non-missing return, beginning with the first trading day of the year and ending with the last trading day of the year. A security must have valid returns for eighty percent of the trading days in a year to have a Standard Deviation calculated. There are two families of indexes provided by CRSP with annual standard deviations as the statistic, the NYSE/NYSE American Standard Deviation Portfolios and the NASDAQ Standard Deviation Portfolios.

Total Counts (DlyTotCnt/MthTotCnt) & Used Counts (DlyUsdCnt/MthUsdCnt)

Total Counts and Used Counts are provided for all indexes and portfolios. The following table identifies differences.

Total Count	Used Count
Current Day closing price required for inclusion	Previous day & current day closing prices required for inclusion
On same date the Total Count will always be greater than or equal TO the Used Count. The difference will be the number of securities with missing prices on the previous day (usually adds)	The Total Count on Day t will be greater than or equal TO the Used Count on Day t+1. The difference will be the number of securities with missing prices on t+1 (usually the drops)
Total Count will fluctuate throughout the year.	Used Count will fluctuate throughout the year.

Total Value (DlyTotVal/MthTotVal) & Used Value (DlyUsdVal/MthUsdVal)

Total Value and Used Value are provided for all CRSP stock indexes. The following table identifies differences.

Total Value	Used Value
Current Day market value of eligible securities - price and shares for the current day are required for inclusion	For value-weighted indexes, this is the Index weight - market value of eligible securities with - price for the current day and price and shares for the previous day are required for inclusion
On same date the Total Value will always be greater than or equal to the Used Value.	

Trade-Only Data

CRSP provides Price or Bid/Ask Average as the standard daily price field, and derives returns from this field. Bid/ask averages are marked as negative numbers by convention. A trade-only price is derived from Price or Bid/Ask Average by setting all bid/ask average prices to missing. Trade-only returns are calculated using trade-only prices. A trade-only index is calculated using trade-only prices and returns.

Unadjusted Data

Unadjusted Data is price, dividend, shares, and volume data reported in the amounts reported at the time of the observations. All CRSP data are provided unadjusted. However, the distribution history can be used to generate Adjusted Data from the raw data.

Weighted Return

Weighted Return is the relative weight of a security within a portfolio or index multiplied by its return. In a value-weighted portfolio, Weighted Return is the capitalization at the end of the previous period multiplied by the return for the period.

Chapter 3: Index Methodologies

Stock File Indexes

The CRSP Stock File Indexes are a set of Market Indexes and Fractile Portfolio Indexes provided daily, monthly, quarterly, and annually for five market groups of securities. The market groups of securities for which indexes are calculated are the individual NYSE, NYSE American, NASDAQ and Arca markets, the NYSE/NYSE American, NYSE/NYSE American/NASDAQ, and the NYSE/NYSE American/NASDAQ/Arca market combinations. Published S&P 500 and NASDAQ Composite Index Data are also included.

The ranges for individual exchange data are listed below. The series containing combinations of exchanges begin at the earliest point that data for any of the exchanges is available.

- The New York Stock Exchange (NYSE) all series begins December 31, 1925
- NYSE American all series begins July 2, 1962
- The NASDAQ Stock Market (NASDAQ) all series begins December 14, 1972
- The Arca Exchange (Arca) all series begins March 8, 2006

Note: Quarterly and annual index returns are not available for the series including Arca.

Daily and monthly index returns are calculated based on daily and monthly security holding period returns, respectively. Quarterly and annual frequency index returns are calculated by compounding monthly index returns.

CRSP Market Indexes

An Equal-Weighted Index and a Value-Weighted Index are calculated for each market group. Each index contains index returns with and without dividends, index weights and counts.

The Equal-Weighted Index is an Equal-Weighted Portfolio built each calendar period from all issues listed on the selected exchanges with valid prices on the current and previous periods.

The Value-Weighted Index is a Value-Weighted Portfolio built each calendar period using all issues listed on the selected exchanges with available shares outstanding and valid prices in the current and previous periods, excluding American Depositary Receipts. Issues are weighted by their Market Capitalization at the end of the previous period.

An additional daily trade-only value-weighted index is available for NYSE/NYSE American. This index uses the same methodology as the NYSE/NYSE American Value-Weighted Market Index, but only includes non-ADR securities with trades on current and previous trading days.

Index Levels of CRSP Market Indexes are set to 100.0 on December 29, 1972.

The NYSE/NYSE American/NASDAQ/Arca Market Indexes are available in Daily and Monthly Stock Files. Other exchange combinations are available in the CRSP US Index Database and Security Portfolio Assignment Module.

Published S&P 500 and NASDAQ Composite Index Data

The S&P 500 Composite Index is a value-weighted index created by Standard & Poor's. Since March 1957, the index contains 500 securities. Prior to that time the index contained 90 securities. These have been combined into a single time series. S&P Composite levels are collected from public sources such as the Dow Jones New Service, the Wall Street Journal and the Standard & Poor's Statistical Service.

The NASDAQ Composite Index is a value-weighted index created by the NASDAQ Stock Market.

Published S&P 500 and NASDAQ Composite Index Data are provided with the daily and monthly CRSPAccess Stock Files. Index levels and returns exclude dividends. As a result, the Return with Dividends variable, returns a missing value for both indexes. Total returns and membership data for the S&P 500 are available to subscribers of the CRSPAccess Index Files.

CRSP Stock File Capitalization Fractile Indexes

CRSP Stock File Capitalization Fractile Indexes are calculated for each of the Stock File Indexes market groups. All securities excluding American Depositary Receipts on a given exchange or combination of exchanges are ranked according to capitalization and then divided into ten equal parts each rebalancing period.

The portfolios are rebalanced each year, using the security market capitalization at the end of the previous year to rank the securities. If a security starts trading in the middle of a year, its first capitalization of the year is used in the ranking. The largest securities are placed in portfolio 10 and the smallest in portfolio 1. A security not assigned to a portfolio is not used in the index and has its Portfolio Assignment set to 0.

Value-Weighted Index Returns including all dividends are calculated on each of the ten portfolios. Index levels are calculated based on an initial value of 100.0 on December 29, 1972.

Each set of decile indexes represents one Index Group of index results and one Portfolio Type of portfolio assignments and statistics. Ten Index Series are created for each Portfolio Type.

Monthly Security Cap INDNOs and daily Security Cap INDNOs are distinguished using different index families, with INDFAMs 11004xx representing the monthly version. Please refer to the CRSP Index Series table in the US Stock and Index Guide for details.

CRSP Stock File Risk-Based Fractile Indexes

CRSP Stock File Risk-Based Fractile Indexes are created for the daily NYSE/NYSE American and NASDAQ market combinations for two risk-based criteria. In these Market Segment Indexes, portfolios are created by ranking securities according to a measurement of the risk of their returns. One ranking uses beta values computed using the methods developed by Scholes and Williams (Myron Scholes and Joseph Williams, “Estimating Betas from Nonsynchronous Data”, Journal of Financial Economics, vol 5, 1977, 309-327). The other ranking uses the annual standard deviation of the daily returns for its ranking.

The methodologies used to calculate these statistics are described in the CRSP Calculations section under Scholes- Williams Beta and Standard Deviation.

CRSP Stock File Risk-Based Fractile Indexes are rebalanced each year by ranking the statistics at the end of the previous year. If there is no data for the previous year for an issue but a valid statistic can be calculated for the current year, that statistic is used in the rankings.

CRSP Beta Fractiles are ranked with Portfolio 1 containing the securities with the largest positive betas and 10 containing securities with the smallest and most negative.

CRSP Standard Deviation Fractiles are ranked with Portfolio 1 containing the securities containing the largest standard deviations and portfolio 10 containing securities with the lowest.Once securities are assigned to portfolios, an equal-weighted total return index is calculated for each portfolio each calendar period. Trade-only security total returns are used for the NYSE/NYSE American Beta Portfolios only. Index levels are calculated based on an initial value of 100.0 on December 29, 1972.

Each set of fractile indexes represents one Index Group of index results and one Portfolio Type of portfolio assignments and statistics. Ten Index Series are created for each Portfolio Type.

CRSP Cap-Based Portfolios

CRSP Cap-Based Portfolio Index data are a monthly series based on portfolios that are rebalanced quarterly. The methodology used to calculate the series differs from the CRSP Stock File Capitalization Fractile Indexes.

The universe includes all common stocks listed on the NYSE, NYSE American, and NASDAQ National Market excluding Unit Investment Trusts, Closed-End Funds, REITs, Americus Trusts, foreign stocks and American Depositary Receipts. Eligible companies with primary listings on the NYSE are ranked into equally populated fractiles. The largest capitalizations in each fractile serve as the breakpoints that are applied to various exchange groupings of the universe.

Fractile results are created for three exchange groups:

- NYSE only
- NYSE and NYSE American. NYSE American data are added beginning July 1962
- NYSE, NYSE American and the NASDAQ National Market. The NASDAQ National Market data are added beginning April 1982

Index family 1180xxx (INDNOs 1080xxx) is based on NYSE only universe and is used to calculate and create breakpoints for all issuer cap-based indexes. The NYSE Breakpoint Issuer Statistics series contains the issuer cap statistics values used to determine breakpoints for Issuer Cap indexes. These statistics are used solely for breakpoint determination and are not used for issuer cap portfolio assignment. Both the rebalancing and membership data of this index family are included.

Individual Fractile portfolios are created for each exchange group, the largest being in fractile 1 and the smallest in fractile 10. In addition to each fractile portfolio, returns are calculated for the following: CRSP 1-2, CRSP 3-5, CRSP 6-8, CRSP 9-10, CRSP 6-10 and CRSP 1-10.

Companies becoming eligible or ineligible during a quarter are handled with the following rules:

- Previous period market capitalizations are used for assigning fractiles and weights.
- Securities added during a quarter are assigned to appropriate portfolios when two consecutive month-end prices are available.
- When a security's last price is a month-end price, its month's return is included in the portfolios' quarterly return.
- When the month-end price is missing, a replacement month-end value is derived from the delisting return including merger terms, regional exchanges, etc. If the derived replacement month-end price is not available, the last available daily price is used.
- If an issue becomes ineligible for an index in the middle of a quarter but is still active, such as after an exchange change or because the issue is leaving the NASDAQ National Market, the issue is considered held until the end of the month and then dropped.
- Index Total Returns, Index Capital Appreciation, and Index Income Returns are calculated from a value-weighted portfolio of securities in the portfolio each period. Index Levels are calculated for each of these returns series based on an investment of one dollar on December 25, 1925.

- Only monthly indexes and portfolio assignments are calculated for the Cap-Based Portfolios. Each of the three sets of Cap-Based Indexes represents one Index Group of index results and one Portfolio Type of portfolio assignments and statistics. Seventeen Index Series, one for each fractile and each composite, are created for each Portfolio Type.

CRSP Indexes for the S&P 500 Universe

CRSP Indexes for the S&P 500® Universe are standard CRSP Market Indexes derived from CRSP Stock Files but include only issues from the CRSP stock data that are in the S&P 500® universe.

The CRSP Indexes for the S&P 500® series contain value- and equal-weighted returns with and without dividends for a market of stocks in the S&P 500® universe. Daily and monthly data beginning December 25, 1925 are provided. The published S&P 500® index and returns are also included for comparison. For a security to be included in the CRSP indexes for the S&P 500 Universe, it must have a price at the end of the current period, a price at the end of the previous period, and it must be a member of the S&P 500 Universe at the end of the current period. See CRSP Market Indexes for the variables calculated and the methodology used.

Prior to March, 1957, the index contains 90 issues. CRSP does not have data for two securities between 1925 and 1931 as follows.

Company Name	Start Date	End Date
INT'L MERCANTILE MARINE PFD	December 31, 1925	July 22, 1929
STANDARD POWER & LIGHT "B"	February 6, 1930	November 16, 1931

Due to differences in handling mergers, reorganizations, and other major corporate actions, CRSP data and the S&P 500® universe do not always have a one-to-one mapping. In some cases this results in a short period where CRSP is missing prices or has multiple prices per company listed by S&P.

The Count of Securities Used is not always 500 (90 prior to March 1957) due to missing prices. Known reasons for missing prices are when-issued trading, halts, and suspensions.

CRSP Portfolios for the S&P 500 Universe

The CRSP Portfolios for the S&P 500 Universe include an alternate value- and equal-weighted version of the CRSP indexes for the S&P 500 Universe. The methodology differences are:

- Issues are selected based on membership in the S&P 500 at the end of the previous period instead of the end of the current period.
- Delisting returns are used to evaluate the value of securities that delist before the end of a period they were selected.

CRSP Treasury and Inflation Indexes

The CRSP US Treasury and Inflation Series (CTI) Files are provided on a monthly frequency. The series contains returns adapted from the CRSP US Treasury Fixed Term Index Series, the CRSP Risk Free Rates File, and the US Government Consumer Price Index. These derived files offer 10 groups of indexes: 30 year, 20 year, 10 year, 7 year, 5 year, 2 year, 1 year, 90 day, and 30 day target maturity indexes, as well as the Consumer Price Index.

For fixed-term series with maturities of one year or greater, a representative Treasury bond or note for each series is selected. Available issues are filtered on the basis of their characteristics. Each month, the most recent non-callable, non-flower, and

fully taxable issue closest to the target maturity is selected. If none are found, a second pass allows flower bonds. Note that all these series begin in 1941 or 1942 due to the lack of suitable issues in the early history.

For thirty and ninety day risk-free series, a representative Treasury bill for each series is selected. Each month the issue maturing closest to the target duration, as measured from the end of the previous month, is selected. Bills must have at least thirty days to their maturity date to be selected for the thirty day series. However, for ninety day series, bills with less than ninety days to maturity may be selected. Due to the lack of data, the selection process in periods prior to 1942 is somewhat subjective and the maturities of the selected issues may deviate several days from the thirty and ninety day targets. Where bills were not available, certificates or notes may have been used. Exclusions may include:

- suspicious quotes
- issues that did not mature on their next coupon payment data
- bid quotations that implied negative yields

Each monthly return is calculated as price change plus interest, divided by last month's price. The returns and corresponding index values are set to -99 for months in which a return cannot be calculated, i.e. if the price is missing for either this month or last month, or if no valid issue was available.

The issue chosen for the 30, 20, 10, 7, 5, 2, and 1 year Fixed Term Index series for a given date was selected based on its length to maturity as of the date. The returns contained in these series are calculated under the assumption that the relevant issue is bought one month prior to the quote date and sold on the date.

The issue chosen for the 90 and 30 day Treasury Bill series on a given date was selected based on its length to maturity as of the month immediately prior to the date. The 90 and 30 day series returns were calculated on the basis of buying the relevant issue one month prior to the date and selling it on the date. For example, a 90 day bill return is calculated between a date approximately 90 days prior to the bill's maturity, and the date which is a month after this date. Likewise, a 30 day bill return is calculated between a date approximately 30 days prior to the bill's maturity, and the date which is a date one month later. In cases where the date chronologically approached or exceeded the maturity date, thereby making a final price unavailable, the return was calculated based on a final price of $100.

The associated index levels of the CRSP US Treasury and Inflation Series all have been initialized so that December 29, 1972 (19721229) equals 100. This facilitates comparison between the CTI Indexes and Stock File Indexes.

CRSP Select File Specifications

Long Term Bond Selection

1. Select the 20-year bond that is the closest to having a term of at least 19.5 years to maturity at the beginning of the year. If more than one exists, choose the bond with the most current dated date (i.e. most recently issued).
2. If a 20-year bond does not meet the above criteria, choose the 25-year bond with at least 19.7 years to maturity at the beginning of the year. If more than one exists, choose the bond closest to 20 years to maturity.
3. If a 25-year bond does not meet the above criteria, choose the 30-year bond with at least 19.7 years to maturity at the beginning of the year. If more than one exists, choose the bond closest to 20 years to maturity on the quote date.

The bond chosen under any of the categories above cannot be dated any later than December 1st of the previous year for which the bond is being considered for inclusion in the index (i.e. dated date + one month < = quote date). Before 1942, only partially tax-exempt bonds (itax=2) are chosen because of the limited number of fully taxable bond issues. After 1942, only fully taxable issues are chosen (itax=1).

The bond is held for one full year in the index. Bonds chosen for this index are either non-callable or callable Treasury bonds with a type of Bond or Callable Bond. A 20-year bond can be selected from a universe of bonds that were issued as having a term to maturity of 7305-7693 days, a 25-year bond from an issue of 8766-9892 days, and a 30- year bond from an issue of 10955-11288 days.

Intermediate Term Bond Selection

1. Select the most currently issued 5 year bond with at least 5 years to maturity at the beginning of the calendar year.
2. If a 5-year bond does not meet the above criteria select the next shortest maturity that is closest to 5 years to maturity on the quote date. For example, if a 7-year bond exists, choose the 7-year closest to 5 years to maturity. If a 7-year bond does not exist move up to the next highest maturity and so forth.

For the period 1934-1942, always choose a non-flower bond (iflwr of 1) and preference is given to a bond that is partially tax exempt (itax of 2). If a partially tax-exempt bond does not meet the above criteria, choose a wholly tax exempt bond (itax of 3). After 1942 only fully taxable non-flower bonds are chosen. Callable and non-callable U.S. Treasury bonds and notes are considered for index inclusion. The issues are chosen from a universe of bonds issued with a term to maturity between 1000 to 7000 days to maturity.

Short Term Bond Selection

Choose the Treasury Bill closest to 90 days to maturity on the quote date. A bill can be within 4 days of target maturity, i.e. 90 days plus or minus 4 days. If a bill is not available, use a certificate or a note.

About CRSP Research Data Products

Center for Research in Security Prices (CRSP), originally established at the University of Chicago in 1960, is widely recognized as a leading provider of research quality historical market and returns data. Built on rigorous academic standards, its research data products are trusted by academic, commercial, and government institutions worldwide that rely on accurate, transparent data for meticulous financial analysis, economic research, and policy development where precision and historical continuity are essential.

In February 2026, Morningstar completed the acquisition of CRSP, integrating CRSP's research data products—renowned for their academic rigor, historical depth, and accuracy—into Morningstar's global data and research platform. This combination enhances Morningstar's equity research capabilities while continuing CRSP's legacy of providing high quality data to support institutional research, benchmarking, and investment decision making.

indexes.morningstar.com/research-data-products

Contact Us

rdp@morningstar.com

MORNINGSTAR