TMX | Toronto Stock Exchange

CFMRC

TSX/CFMRC Database

USER'S GUIDE

2020-05-06
Rev. 2021-05-10

Table of Contents

TSX/CFMRC Database User's Guide

1. Introduction

The TSX/CFMRC Database is a group of research databases compiled and published by the Canadian Financial Markets Research Center (CFMRC) from data provided by The Toronto Stock Exchange (TSX) and other sources. Currently, there are three databases: 1) Summary Information, 2) Toronto Stock Exchange Intra-daily Trade and Quote, and 3) TSX Venture Exchange Intra-daily Trade and Quote.

The Summary Information database contains daily and monthly trading information from the Toronto Stock Exchange about listed securities as well as information on "price adjustments" such as dividends, stock splits, recapitalizations, etc. and ticker changes. The daily and monthly "index" files contain information on daily and monthly index levels as well as selected other financial markets information. Monthly data begins as early as 1949-12 and daily data begins on 1975-01-02.

The two trade and quote (TAQ) databases contain time-stamped and sequencing information on all trades and quotes that occurred on the TSX or TSXV. TAQ data currently begins on 2018-01-02.

Acknowledgments

Data on interest and foreign exchange rates has been made available through the generosity and kind assistance of the Bank of Canada and Statistics Canada. Selected data from Statistics Canada's CANSIM data bank are published with authority of Statistics Canada. CANSIM is Statistics Canada's official Time Series data bank service. For further information about the services contact your nearest Statistics Canada office or call 1-800-263-1136. Where appropriate, the CANSIM data series number is noted.

Disclaimer of Warranties

NEITHER CFMRC NOR ANY THIRD-PARTY SUPPLIERS OF DATA OR SOFTWARE TO CFMRC MAKE ANY WARRANTY, EXPRESS OR IMPLIED, AS TO RESULTS TO BE OBTAINED BY SUBSCRIBER OR OTHERS FROM THE USE OF THE DATABASES OR ANY SOFTWARE PROVIDED HEREUNDER, AND THERE ARE NO EXPRESS OR IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE. The information contained in the Databases shall be obtained from sources considered by CFMRC to be reliable, but the accuracy and completeness thereof are not guaranteed and CFMRC and its third-party suppliers shall have no liability for errors or omissions with respect to the Databases or their delivery, regardless of the cause or source of such error or omission.

Using the CFMRC Databases

The databases are made up of regular text (ASCII) files which may be read with a standard text editor such as Notepad. Some files may be too large[1] to read entirely with some editors – please check your editor's size limits. Although the files can be read with an editor, that is not the common way to use them. Most users will use one of the following methods (or both):

---

[1] The Summary Information database file Data.tse is over 2 gigabytes and is too large for many text editors. Some of the Trade and Quote daily files may be too large, as well.

1) Use the Windows-based TSX/CFMRC Browser to select the data of interest.
2) Write specialized programs to access the databases.

The latter alternative is for users on non-Windows platforms or users who wish to tailor their computations. Most users will probably opt for the ease and simplicity of the browser.

Database and Browser Installation

The TSX/CFMRC Databases are delivered and updated from the cloud and require the TSX/CFMRC Browser v2.3 (the browser), or later, for installation and updates. The browser currently only runs under versions of the Windows operating system (Windows 7, or later, or Server 2008, or later). Windows 10 or Server 2016 (or later) are recommended.

For complete instructions on installing the TSX/CFMRC Browser v2.3 and the CFMRC Databases, please see the document "*TSX_CFMRC Browser v2.3 Installation and User's Guide.pdf*" which is included in the documentation distribution.

Support

If you have any questions, comments, suggestions, or criticisms, please feel free to contact the TSX at:

<div align="center">

TMX Group
300 - 100 Adelaide St. West
Toronto, ON
M5H 1S3

</div>

Or email to: DataSales@tmx.com

Phone: (416) 947-4452

2. Database Directory Structure

When you install the TSX/CFMRC Databases, you will specify a location called the Database Destination. The browser will create a directory structure under this root database directory and install all of the data files and database indices in the appropriate places under this directory. Updates are also downloaded to locations under this directory and, on occasion, backups will also be created there.

The notation [Database Destination]\ refers to the directory named in the Database Destination field on the Settings panel of the browser. Figure 1 shows a typical file and directory listing for a Database Destination.
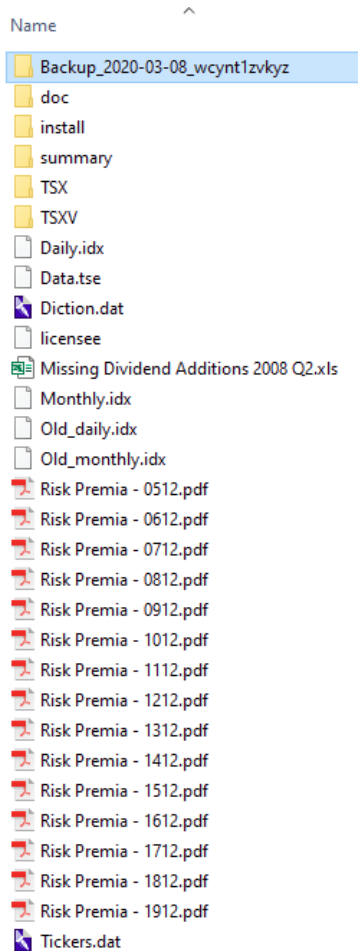

Figure 1. File and Directory Listing for [Database Destination]

There are currently three databases in the TSX/CFMRC Database whose files are stored as follows:

Database	Database Files Location
Summary Information	[Database Destination]\
TSX TAQ	[Database Destination]\TSX\
TSXV TAQ	[Database Destination]\TSXV\

Other directories under [Database Destination]\ have a special purpose and any files in a database location, belong to that database.

Summary Information Database Files

All of the Summary Information Database files for a distribution are found in the [Database Destination]\ directory. The primary summary information files are DAILY.IDX, DATA.TSE, DICTION.DAT, MONTHLY.IDX, and TICKERS.DAT. The other files shown in Figure 1, and the doc directory, contain additional information that is related to the summary information database such as documentation or change information. Periodically, information of interest to CFMRC subscribers is released in an official announcement. These announcements may be found in the announcements directory (rev 2021-05-10).

TSX TAQ Database Files

All of the files associated with the TSX TAQ database are in the [Database Destination]\TSX\ directory. Figure 2 shows a partial listing of the TSX TAQ location.

Name
TSX20200226.zip
TSX20200227.zip
TSX20200228.zip
TSX20200302.zip
TSX20200303.zip
TSX20200304.zip
TSX20200305.zip
TSX20200306.zip
TSX20200309.zip
TSX20200310.zip
TSX20200311.zip
TSX20200312.zip
TSX20200313.zip
TSX20200316.zip
TSX20200317.zip
TSX20200318.zip
TSX20200319.zip
TSX20200320.zip
TSX20200323.zip
TSX20200324.zip
TSX20200325.zip
TSX20200326.zip
TSX20200327.zip
TSX20200330.zip
TSX20200331.zip
TSXBusinessIndex.txt
TSXCUSIPIndex.txt
TSXDateIndex.txt
TSXMasterIndex.txt
TSXMasterIndex.zip
TSXNameIndex.txt
TSXTickerIndex.txt

Figure 2. File Listing for [Database Destination]\TSX

Name
TSXV20200226.zip
TSXV20200227.zip
TSXV20200228.zip
TSXV20200302.zip
TSXV20200303.zip
TSXV20200304.zip
TSXV20200305.zip
TSXV20200306.zip
TSXV20200309.zip
TSXV20200310.zip
TSXV20200311.zip
TSXV20200312.zip
TSXV20200313.zip
TSXV20200316.zip
TSXV20200317.zip
TSXV20200318.zip
TSXV20200319.zip
TSXV20200320.zip
TSXV20200323.zip
TSXV20200324.zip
TSXV20200325.zip
TSXV20200326.zip
TSXV20200327.zip
TSXV20200330.zip
TSXV20200331.zip
TSXVBusinessIndex.txt
TSXVCUSIPIndex.txt
TSXVDateIndex.txt
TSXVMasterIndex.txt
TSXVMasterIndex.zip
TSXVNameIndex.txt
TSXVTickerIndex.txt

Figure 3. File Listing for [Database Destination]\TSXV

Each of the TSXyyyymmdd.zip files, where yyyymmdd represents a date, is a compressed package of TAQ data files for the date specified. The package contains text files with all of the trades and quotes that occurred on the TSX for each security listed on the TSX on that date. A "name index" or mapping of tickers to names and other security information for that date is also included in the package in a text file.

The daily TAQ packages can be read directly with the browser, decompressed into individual files with numerous compression utilities such as WinZip (commercial) or 7-Zip (open source), or even read from your own programs using any .NET language or open source libraries.

You will also find six TSX????Index.txt files, where ???? is one of {Business, CUSIP, Date, Master, Name, Ticker}, in this directory. These files are primarily for the browser to build indices to the TSX TAQ data but may be useful for those accessing the data with their own programs. These files are discussed in detail in the TSX TAQ Data section.

TSXV TAQ Database Files

All of the files associated with the TSXV TAQ database are in the [Database Destination]\TSXV\ directory. Figure 3 shows a partial listing of the TSXV TAQ location.
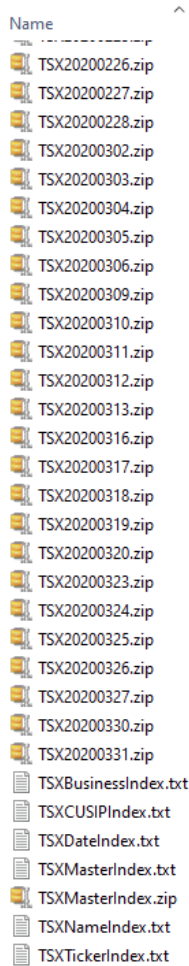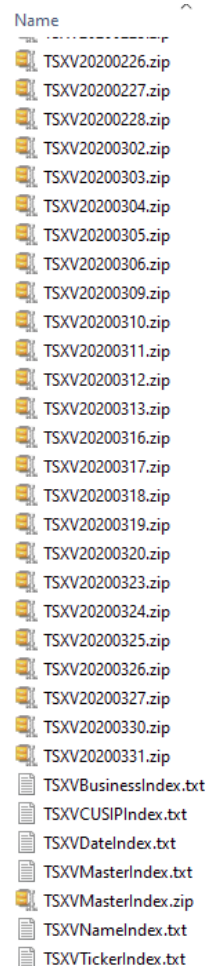
The TSXV TAQ location description is identical to the TSX TAQ location discussion (please see above), except that file names begin with TSXV, not TSX.

Other Special Directories Under [Database Destination]

There are two other directories under [Database Destination]\: install and summary. The install directory contains copies of all the software versions you have downloaded. Each version of the installer will have a name like CFMRC-Setup-23029.exe, where 23029 is the version number. The higher the version number, the more recent the version of the software. If you have multiple locations to install the browser on, you can simply copy the most recent installer file from this directory rather than downloading it again from the web.

Please do not remove files from this directory as network installations may stop updating properly if you do.

The summary directory should only contain one file: Summary_Full.zip. This is the last full update to the Summary Information database that was downloaded. This is already installed on your system by the browser.

The final type of directory you might find under [Database Destination]\ is backups. Whenever part of the Summary Information database is updated, a copy of the existing file is made in a Backup_yyyy-mm-dd_{random letters} directory. This allows the browser to roll back the update if an error occurs or you cancel the update. yyyy-mm-dd is the date of the update. Once you complete an update and are certain there are no issues, you may delete these backups if you need the space (each one is about 2.5 gigabytes).

.BIN Files

Please do not use any files ending in .BIN in any of the database directories. These are for the browser.

WARNING

In general, it is good practice not to edit or change anything in any of the files or directories under [Database Destination]\. Please write your output files to different directories. If you are writing your own programs, it is recommended that you open files in READ ONLY mode. Deleting, renaming or moving a file may only make some data inaccessible but overwriting or changing data in a file is likely to crash the browser (which is required for updates!).

3. Summary Information Data Description

There are seven different types of data in the TSX/CFMRC Summary Information database:

1. Daily index data.
2. Monthly index data.
3. Stock header information.
4. Daily trading data.
5. Monthly trading data.
6. Price adjustment data.
7. Ticker History data.

Each is discussed below.

Changes

This document reflects all changes up to, and including, the **2021 Q1 release** of the Summary Information database. CUSIPs have been added to the first header of all securities in DATA.TSE using previously reserved space in each record.

FOR CUSTOMERS CURRENTLY USING PROGRAMS WRITTEN TO INTERACT WITH PREVIOUS VERSIONS OF THE DATABASE, PLEASE BE AWARE THAT ALL OF THE FILES IN THE SUMMARY INFORMATION DATABASE HAVE CHANGED, BEGINNING WITH THE 2021 Q1 RELEASE. DATES HAVE ALL BEEN EXPANDED TO 8 CHARACTERS (YYYYMMDD) AND SPACE FOR SECURITY SYMBOLS HAS BEEN INCREASED TO 12 CHARACTERS (WHICH INCLUDES TYPE AND CLASS IDENTIFIERS AND THEIR SEPARATORS, '.'). IN ADDITION, THE LAYOUT OF PRICE ADJUSTMENT RECORDS IN DATA.TSE HAS CHANGED. PLEASE SEE CHAPTER 4 FOR COMPLETE DETAILS (REV 2021-05-10).

New data and functionality have been added to the CFMRC Databases however, only the previously unsupported Ticker History, DICTION.DAT, has been added to the Summary Information database. This data and the file format are described below.

Missing Data

Unless otherwise noted, missing data is flagged with -9.0 for real (exceptions are monthly closing prices which use 0.0, and EPS which uses -9999.0) data or -9 for integer data. Data is missing if there is no information available. Thus, if there were no trades in a day for a certain security, the number of trades would be set to 0 and the closing price (for example) would be 0.00 - that is, there is information. Only if there is no trading information at all would trades be set to -9 and closing price to -9.00.

Data Errors

While every effort has been made to eliminate data errors, suspect or erroneous data still creeps into the database. In cases where data is clearly wrong or suspect, data that is only meaningful when positive (shares outstanding or prices, for example) has been flagged by multiplying by -1.0. The user is free to decide on the usefulness of these data and can recover them by multiplying by -1.0 again. Again, monthly closing prices are an exception. See below.

3.1 Daily Index Data (Daily Calendar)

The CFMRC Summary Information database uses relative calendars. That is, each trading day or month is given a sequential day or month number called its "relative calendar" number.

Daily data in the database begins on January 2, 1975 and so 750102 (dates in the database use the YYMMDD format) has a relative daily calendar number of 1. 750103 has a relative daily calendar number of 2 but January 4 & 5, 1975 fell on a Saturday and Sunday so these are non-trading days. Relative day 3 is 750106. To determine the number of trading days between two calendar dates (counting both dates), subtract the relative daily calendar number of the start date from the relative daily calendar number of the end date and add 1. Because of weekends and holidays, there is no simple algorithm for mapping to/from relative calendar number from/to Julian calendar date.

Data available for each day in the relative calendar is:

Date - the Julian calendar date corresponding to this relative daily calendar number.

S&P/TSX Composite Price Index - the closing S&P/TSX Composite Price Index. This index is based solely on security prices. See the *TSX Review* (green pages) for securities included in this index and their weights. Daily data for TSX Indices begins on 19770103 (relative day number 506).

S&P/TSX Composite Total Return Index - the closing S&P/TSX Composite Total Return Index based on prices and distributions (stocks, dividends, etc.). See the *TSX Review* (green pages) for securities included in this index and their weights. Daily data for TSX Indices begins on 19770103 (relative day number 506).

S&P/TSX Sector Indices: Price & Total Return - daily closing values for all ten S&P/TSX Sectors based on closing prices (Price) or including distributions (Total Return). See the *TSX Review* (green pages) for securities included in each sector and their weights. Daily data for S&P/TSX Sectors begins on 19871231. The present sectors are:

> Energy Sector
> Materials Sector
> Industrials Sector
> Consumer Discretionary Sector
> Consumer Staples Sector
> Health Care Sector
> Financials Sector
> Information Technology Sector
> Telecommunications Services Sector
> Utilities Sector

Also included as of 20020603 are the following S&P/TSX Indices (Price and Total Return):

S&P/TSX 60 Index
S&P/TSX Canadian Midcap Index
S&P/TSX Canadian Smallcap Index
S&P/TSX Venture Composite Index

CFMRC Equal Weighted Index - the average daily return for all domestic common equities in the CFMRC database. This is the sum of all defined common equity returns divided by the number of valid equity returns (Canadian based firms only). Returns used in this index are fully adjusted for distributions. See Daily Returns below for the definition of a valid return.

CFMRC Value Weighted Index - the market value weighted average daily return for all domestic common equities in the CFMRC database. This is the sum of all products of defined common equity returns with their market weights. A security's market weight is defined as its market value at the beginning of the current month (shares outstanding times closing price on the last trading day in the previous month) divided by the market value of all securities included in the index (Canadian firms only). Returns used in this index are fully adjusted for distributions. See Daily Returns below for the definition of a valid return.

CFMRC Under $2 Equal Weighted Index - similar to the CFMRC Equal Weighted Index but includes only those securities for which prices fell below $2.00 on either the current or previous trading day.

CFMRC Under $2 Value Weighted Index - similar to the CFMRC Value Weighted Index but includes only those securities for which prices fell below $2.00 on either the current or previous trading day.

CFMRC Over $2 Equal Weighted Index - similar to the CFMRC Equal Weighted Index but includes only those securities for which prices did not fall below $2.00 on either the current or previous trading day.

CFMRC Over $2 Value Weighted Index - similar to the CFMRC Value Weighted Index but includes only those securities for which prices did not fall below $2.00 on either the current or previous trading day.

Call Loan Rate - the Canadian over-night wholesale money market interest rate. Originally, this was CANSIM series B114011. As of July, 2018, this is the Overnight MM Rate from Table 10-10-0139-01.

Foreign Exchange Rate (CA$ per US$) - the daily closing spot exchange rate expressed as Canadian dollars per U.S. dollar. This was CANSIM series B100014. As of July, 2018, it is the FX Rate – Daily from Table 33-10-0036-01.

Historical Daily Index Data

The old Daily Calendar file is included with the database in a flat file named OLD_DAILY.IDX. All daily data fields described above are present in this file except for the S&P/TSX Sector Indices which are replaced with the old S&P/TSX Composite Group Indices, as listed below. This file cannot be viewed with the browser.

S&P/TSX Group Indices: Price & Total Return - daily closing values for all fourteen S&P/TSX Group Indices based on closing prices (Price) or including distributions (Total Return). See the TSX Review (green pages) for securities included in each group and their weights. Daily data for TSX Indices begins on 770103 and ends on 021231. The groups are:

Metals & Minerals
Gold & Silver

Oil & Gas
Paper & Forest Products
Consumer Products
Industrial Products
Real Estate & Construction
Transportation & Environmental
Pipelines
Utilities
Communications & Media
Merchandising
Financial Services
Conglomerates

3.2 Monthly Index Data (Monthly Calendar)

The Summary Information database uses relative calendars. That is, each trading day or month is given a sequential day or month number called its "relative calendar" number.

Monthly data in the database begins with January, 1950 and so 195001 (monthly dates are expressed as YYYYMM) has a relative monthly calendar number of 1. Unlike the daily calendar, relative monthly calendar numbers can be easily computed. If yyyymm (where yyyy is the the year and mm is the month number – for example, 197208 for Aug/1978) is the month in question, then

$$month\,\# = \left(\left(INT\left(\frac{yyyymm}{100}\right) - 1950\right) * 12\right) + \left(yyyymm - \left(INT\left(\frac{yyyymm}{100}\right) * 100\right)\right)$$

The database actually contains price observations (no other data) for many securities for December, 1949. Using the above formula, 194912 has a relative monthly calendar number of 0.

Data available for each month in the relative calendar is:

Month End - the Julian calendar date of the last trading day in the month. For months prior to 197501, the date is given as yyyymm00 (that is, no day).

S&P/TSX Composite Price Index - the closing S&P/TSX Composite Price Index. This index is based solely on security prices. See the TSX Review (green pages) for securities included in this index and their weights. Monthly data for TSX Indices begins on 195601 (relative month number 73).

S&P/TSX Composite Total Return Index - the closing S&P/TSX Composite Total Return Index based on prices and distributions (stocks, dividends, etc.). See the TSX Review (green pages) for securities included in this index and their weights. Monthly data for TSX Indices begins on 195601 (relative month number 73).

S&P/TSX Sector Indices: Price & Total Return - monthly closing values for all ten S&P/TSX Sector based on closing prices (Price) or including distributions (Total Return). See the TSX Review (green pages) for securities included in each sector and their weights. Monthly data for S&P/TSX Sectors begins on 19871231. The present sectors are:

Energy Sector

Materials Sector
Industrials Sector
Consumer Discretionary Sector
Consumer Staples Sector
Health Care Sector
Financials Sector
Information Technology Sector
Telecommunications Services Sector
Utilities Sector

Also included as of 20020603 are the following S&P/TSX Indices (Price and Total Return):

S&P/TSX 60 Index
S&P/TSX Canadian Midcap Index
S&P/TSX Canadian Smallcap Index
S&P/TSX Venture Composite Index

CFMRC Equal Weighted Index - the average monthly return for all domestic common equities in the Summary Information database. This is the sum of all defined common equity returns divided by the number of valid equity returns (Canadian based firms only). Returns used in this index are fully adjusted for distributions. See Monthly Returns below for the definition of a valid return.

CFMRC Value Weighted Index - the market value weighted average monthly return for all domestic common equities in the database. This is the common equity returns weighted with their market weights. A security's market weight is defined as its market value at the beginning of the current month (shares outstanding times closing price on the last trading day in the previous month) divided by the market value of all securities included in the index (Canadian firms only). Returns used in this index are fully adjusted for distributions. See Monthly Returns below for the definition of a valid return.

T-Bill Rate - the average yield on 91-day Government of Canada T-bills based on the weekly auctions. This was CANSIM series B14007. As of July, 2018, this is the 3 mon T-Bill yield from Table 10-10-0122-01.

Long Term Government Bond Rate - the average yield on a portfolio of 10+ year Government of Canada (GOC) Bonds. This was CANSIM series B14013. As of July, 2018, this is the over 10 years GoC from Table 10-10-0122-01.

Corporate Bond Rate - the average yield on a portfolio of high grade, long term corporate bonds. This is CANSIM series B14048. This series was discontinued at the end of June, 2007.

Foreign Exchange Rate (CA$ per US$) - the average noon spot exchange rate expressed as Canadian dollars per U.S. dollar. CANSIM series B3400 prior to July, 2018. Now the FX rate – monthly series from Table 33-10-0163-01.

T-Bill Return - the return on a 91-day T-bill purchased at the end of last month and sold at the end of this month. If r_t is the yield (in percent) at the end of month t, then the price last month, P_{t-1}, of a T-bill with 91 days to maturity is

$$P_{t-1} = 100/(100 + r_{t-1}^{\frac{91}{365}})$$

and the price today of that same bill with today's yield and only 61 days to maturity is

$$P_t = 100/(100 + r_t^{\frac{61}{365}})$$

The return, R_t, is

$$R_t = (P_t - P_{t-1})/P_{t-1}$$

Long Term Government Bond Return - the return on a long term GoC bond with an assumed term to maturity of 17 years purchased at the end of last month (at par, or P_{t-1} = 100) and sold at the end of this month.

If Y_t is the yield at the end of month t, then the price, P_t is simply the present value of the future semi-annual interest payments plus the discounted terminal, or par, value.

$$P_t = \frac{Y_{t-1}/2}{(1 + (\frac{Y_t}{100})/2)^1} + \frac{Y_{t-1}/2}{(1 + (\frac{Y_t}{100})/2)^2} + \cdots + \frac{Y_{t-1}/2}{(1 + (\frac{Y_t}{100})/2)^{2M}} + \frac{100}{(1 + (\frac{Y_t}{100})/2)^{2M}}$$

where M is the number of years to maturity (= 17 in this case). The interest received in month t is $Y_{t-1}/12$ so the total return is

$$R_t = \frac{P_t - 100 + Y_{t-1}/12}{100}$$

Historical Monthly Index Data

The old Monthly Calendar file is included with the database in a flat file named OLD_MONTHLY.IDX. All monthly data fields described above are present in this file except for the S&P/TSX Sector Indices which are replaced with the old S&P/TSX Composite Group Indices, as listed below. This file cannot be viewed with the browser.

TSX Group Indices: Price & Total Return - monthly closing values for all fourteen TSX Group Indices based on closing prices (Price) or including distributions (Total Return). See the TSX Review (green pages) for securities included in each group and their weights. The present groups are:

Metals & Minerals
Gold & Silver
Oil & Gas
Paper & Forest Products
Consumer Products
Industrial Products
Real Estate & Construction
Transportation & Environmental
Pipelines
Utilities

Communications & Media
Merchandising
Financial Services
Conglomerates

3.3 Security Header Information

Security Name - the most recent name of this security while this security was still listed. If a firm changed its name after delisting one of its securities, the change will not be reflected in the database record for that delisted security. Any other securities issued by the firm that were actively listed when the name change occurred will reflect the name change.

Ticker - the most recent ticker symbol for this security.

Usage Number - over time, ticker symbols are reused by the TSX. To uniquely identify securities, the CFMRC has assigned a usage number to each use of a particular ticker which indicates the number of times the ticker has been used before in the CFMRC databases.

For example, between 19450711 and 19730704, Atlantic Sugar Refineries, Inc. used the ticker AC on the TSX. Since this was the first usage of that symbol in the database, it was assigned the usage number 0. So, AC (0) refers to Atlantic Sugar Refineries, Inc. In 1988, the same ticker was issued to Air Canada but since it had been used before, Air Canada was given the usage number 1 so the unique ticker symbol is AC (1).

CUSIP - a nine character (including letters and numbers) code that uniquely identifies a company or issuer and the type of financial instrument. A CUSIP number identifies most financial instruments, including: stocks of all registered U.S. and Canadian companies, commercial paper, and U.S. government and municipal bonds.

Business - a brief description, defined by the TSX, of the type of business this firm is involved in.

Foreign Based - a flag to identify foreign-based versus domestic firms. If this is a domestic firm, this flag is '0' (the character zero). If it is '1', the firm is foreign based and will be excluded from all CFMRC index calculations.

Starting Relative Calendar Day - the relative daily calendar number of the first day of data for this security.

Number of Days - number of daily data records.

Starting Relative Calendar Month - the relative monthly calendar number of the first month of data for this security.

Number of Months - number of monthly data records.

Number of Price Adjustments - number of price adjustment records (dividends, splits, etc.).

3.4 Daily Trading Data

Daily data in the Summary Information database begins on January 2, 1975. However, data for the time between 19750102 and 19770228 inclusive is based on round lot trades only. As of 19770301, data is based on ALL trades (including odd lots). Information reported in the *TSX Review* is based on round lots only.

Opening Price - the price of the first transaction after the market opened. The TSX opened at 10:00 A.M. before 19850930 but has opened at 9:30 A.M. since.

Opening Bid - the last bid price before the market opened. If there were no quotes prior to the market opening, this is the first bid price since the open. The TSX opened at 10:00 A.M. before 19850930 but has opened at 9:30 A.M. since. This data is not available prior to 19770301.

Opening Ask - the last ask price before the market opened. If there were no quotes prior to the market opening, this is the first ask price since the open. The TSX opened at 10:00 A.M. before 19850930 but has opened at 9:30 A.M. since. This data is not available prior to 19770301.

Closing Price - the price of the last transaction of the day.

Closing Bid - the bid price for the last quote of the day before the market closes. This data is not available prior to 19770301.

Closing Ask - the ask price for the last quote of the day before the market closes. This data is not available prior to 19770301.

High - the highest price at which this security traded during the day.

Low - the lowest price at which this security traded during the day.

Return - the fully adjusted daily return calculated as if the security was purchased at the close yesterday and sold at the close today. Let P_t be today's closing price, D_t be the cash or cash equivalent dividend (in $) paid today (that is, today is the ex-dividend date) and S_t be the stock split factor for a stock dividend or split today. The return is

$$R_t = \frac{(P_t + D_t) * S_t - P_{t-1}}{P_{t-1}}$$

If there was no cash dividend, D_t = 0.0 and if there was no stock dividend or split, S_t = 1.0. If a security was reclassified at *t* or *t-1*, the return will be recorded as -9.0 for both. Since open and closing prices are now available in the database, the user can compute quasi-split factors if desired.

Transactions - the number of transactions in this security on day *t*.

Volume - total volume transacted in this stock on day *t*.

Quotes - the number of valid market quotes posted throughout the day for this security. This data is not available prior to 19770301.

Quote Changes - number of times the valid market quoted prices changed throughout the day for this security. Since prices do not always change when a new quote is posted, this number will always be less than or equal to Quotes (above). This data is not available prior to 19770301.

3.5 Monthly Trading Data

Closing Price - the last trade price from the last two days of trading in the given month. If a security did not trade within the last two trading days of the month, -1 times the mean of the last bid and ask quote in the last two days of trading is reported here. If there were no trades or quotes in the last two trading days in the month, 0.0 is reported.

Prices are arbitrarily limited to the last two trading days to ensure that monthly prices are not "stale" and that monthly returns are calculated over a one-month period. This also has the effect of eliminating infrequently traded stocks from the CFMRC indices since they rarely have monthly returns.

Return - the fully adjusted return based on purchasing a share at last month's close and selling it at this month's close. If either month's closing price is less than or equal to 0.0 or an undefined dividend/recapitalization occurred within the month, the return is undefined (-9.0).

EPS - trailing twelve months earnings per share as reported by the TSX. Missing data is flagged with -9999.0 (minus 9999.0).

Beta - the rolling average beta from the CAPM. Given $R_{f,t}$ and $R_{m,t}$, the T-bill return and return on the TSX 300 Total Return Index (see above), and the security return, R_t, beta is calculated as:

$$\log(R_t - R_{f,t}) = a + Beta * \log(R_{m,t} - R_{f,t})$$

When there is no valid return due to missing prices, the mean of the bid and ask is used as a proxy (-1 times the monthly closing price, if valid). A minimum of 24 months of returns over the past 60 months are required before a beta is calculated.

Transactions - total number of transactions in this month.

Total Volume - total volume transacted this month in 100's of shares.

Shares Outstanding - total shares outstanding at month end for this security in 100's of shares.

Dividend Flag - a flag to indicate whether a price adjustment occurred this month. 0 if there were no price adjustments and 1 if there were.

3.6 Price Adjustments Data

Dividend - a decimal amount whose meaning depends on the contents of the Dividend Type flags.

Dividend Type - three flags that indicate the type of dividend and the format of the Dividend field.

The first flag identifies the type of dividend or capital change:

       0 - Cash dividend - Regular
       1 - Cash dividend - Extra
       2 - Stock dividend
       3 - Stock in lieu of cash
       4 - Stock right
       5 - Stock split

6 - Recapitalization
9 - Meaningless data. Record should be ignored.

The second flag identifies the currency in which the dividend was paid:

0 - Canadian dollars
1 - United States dollars
2 - Other foreign currency

The last flag identifies the format of the data in the Dividend field:

0 - Dollar value
1 - Decimal - fraction of a share or share multiple

If the last flag is 1, the type of number in the Dividend field depends on the type of dividend as indicated by flag one. If this is a stock dividend, stock in lieu dividend or stock right, then the number represents a fraction of a share. Otherwise, the number in the Dividend field is a ratio (i.e. 2:1 is 2.0).

Ex-dividend Date - the date upon which the security began trading ex-dividend or ex-split.

Recapitalizations - for recapitalizations and reclassifications, the dividend field is -9.0. Since opening and closing prices are now given in the database, the user can compute quasi-split factors if desired.

3.7 Ticker History

Over time, many firms list new securities, delist expiring securities or change their ticker symbols. The database includes a ticker history with the following information:

Delisting Date - the date the security was delisted in YYMMDD format. If this date ends in 00, either the date preceded the availability of daily data (January 2, 1975) or the ticker is still current and the date (YYYYMM) is the next month that will be added in a future update.

For example, the 2019 Q4 Summary Information database contained many ticker history records with a Delisting Date of 20200100. The last date in that database was 20191231 so these tickers are all current.

Listing Date – the date the ticker was first listed.

4. Summary Information File Formats

All files contain text data and each line is terminated with a linefeed (hex 0A, decimal 10). File formats are given in tabular form detailing the start position, length, field name, and characteristics of each field in a record.

For convenience, Fortran FORMAT statement notation will also be used to summarize the format of each data record in the various data files. The field specifiers are:

Ann   - *nn* spaces of character data

Fx.y   - *x* spaces of floating point data which includes a decimal point ('.') and *y* decimal places.

Inn   - *nn* spaces of integer data

nnX   - *nn* spaces are skipped (but may contain meaningless characters).

Format specifiers are separated by commas and may be grouped with parentheses '( )'. Any specifier or group of specifiers may be repeated *y* times by preceding it with *y*. For example, the format string

$$(A30,2I5,3(1X,F10.4))$$

defines a 30-character long string, 2 five-digit integers and 3 floating point numbers of ten digits each (counting the decimal point and four decimal places) each preceded by one space.

In the data files, numbers are right justified within their fields and strings are left justified. Leading zeros are suppressed.

4.1 Daily Calendar/Index File - **DAILY.IDX** and **OLD_DAILY.IDX[2]**

Each record in the daily index file represents one day. Records are fixed-length and contain 362 data characters and a terminating linefeed for a total of 363 characters per record. Records are organized sequentially starting with relative daily calendar day one. Please see Table 1 for record contents and layout. Note that, presently, each field in the file is separated by one, or more, spaces.

The format is (rev 2021-05-10):

$$(I5,1X,I8,2(1X,F8.2),14(1X,F8.2,1X,F8.2),2(1X,F8.5),3(1X,F9.6,1X,F9.6))$$

4.2 Data File - **DATA.TSE**

Records in the main data file are fixed-length, 112 bytes long, including the trailing linefeed. There is one header record followed by one group of data records for each security in the database.

The file header is simply a 97-character string identifying the version and date of the database. The format is (A97).

Each group of records contains all the data records for a specific security. The order of the data records is:

1. 2 security header records.
2. A variable number (possibly 0) of daily data records.
3. A variable number (possibly 0) of monthly data records.
4. A variable number (possibly 0) of price adjustment records.

---

[2] The length of date fields increased from 6 to 8 characters beginning with the 2021 Q1 release of the database. Thus, records in OLD_DAILY,IDX are two characters shorter than reported here and use the following format:

$$(I5,1X,I6,2(1X,F8.2),14(1X,F8.2,1X,F8.2),2(1X,F8.5),3(1X,F9.6,1X,F9.6))$$

Table 1: Record Layout for DAILY.IDX

Position	Length	Field	Characteristics
1	5	Relative day number	Numeric (integer)
7	8	Date	YYYYMMDD
16	8	S&P/TSX Composite Daily Price Index	DDDDD.dd (numeric, includes '.' and two decimal places)
25	8	S&P/TSX Composite Daily Total Return Index	DDDDD.dd
3	8	Sector 10/Group 1 Daily Price Index	DDDDD.dd
43	8	Sector10/Group 1 Daily Total Return Index	DDDDD.dd
52	8	Sector 15/Group 2 Daily Price Index	DDDDD.dd
61	8	Sector 15/Group 2 Daily Total Return Index	DDDDD.dd
70	8	Sector 20/Group 3 Daily Price Index	DDDDD.dd
79	8	Sector 20/Group 3 Daily Total Return Index	DDDDD.dd
88	8	Sector 25/Group 4 Daily Price Index	DDDDD.dd
97	8	Sector 25/Group 4 Daily Total Return Index	DDDDD.dd
106	8	Sector 30/Group 5 Daily Price Index	DDDDD.dd
115	8	Sector 30/Group 5 Daily Total Return Index	DDDDD.dd
124	8	Sector 35/Group 6 Daily Price Index	DDDDD.dd
133	8	Sector 35/Group 6 Daily Total Return Index	DDDDD.dd
142	8	Sector 40/Group 7 Daily Price Index	DDDDD.dd
151	8	Sector 40/Group 7 Daily Total Return Index	DDDDD.dd
160	8	Sector 45/Group 8 Daily Price Index	DDDDD.dd
169	8	Sector 45/Group 8 Daily Total Return Index	DDDDD.dd
178	8	Sector 50/Group 9 Daily Price Index	DDDDD.dd
187	8	Sector 50/Group 9 Daily Total Return Index	DDDDD.dd
196	8	Sector 55/Group 10 Daily Price Index	DDDDD.dd
205	8	Sector 55/Group 10 Daily Total Return Index	DDDDD.dd
214	8	S&P/TSX 60/Group 11 Daily Price Index	DDDDD.dd
223	8	S&P/TSX 60/Group 11 Daily Total Return Index	DDDDD.dd
232	8	S&P/TSX Mid Cap/Group 12 Daily Price Index	DDDDD.dd
241	8	S&P/TSX Mid Cap/Group 12 Daily Total Return Index	DDDDD.dd
250	8	S&P/TSX Small Cap/Group 13 Daily Price Index	DDDDD.dd
259	8	S&P/TSX Small Cap/Group 13 Daily Total Return Index	DDDDD.dd
268	8	S&P/TSX Venture/Group 14 Daily Price Index	DDDDD.dd
277	8	S&P/TSX Venture/Group 14 Daily Total Return Index	DDDDD.dd
286	8	Call Loan Interest Rate	DD.ddddd (numeric, includes '.' and 5 decimal places)
295	8	Daily Foreign Exchange Rate (Cdn$ per US$)	DD.ddddd
304	9	CFMRC Daily Equal Weighted Index	DD.dddddd (numeric, includes '.' And 6 decimal places)
314	9	CFMRC Daily Value Weighted Index	DD.dddddd
324	9	CFMRC Equal Weighted Under $2 Index	DD.dddddd
334	9	CFMRC Value Weighted Under $2 Index	DD.dddddd
344	9	CFMRC Equal Weighted Over $2 Index	DD.dddddd
354	9	CFMRC Value Weighted Over $2 Index	DD.dddddd

4.2.1 Security Header Records

The first header record contains identifying information about the security in the group and the second describes the records in the group.

Table 2: Record Layout for DATA.TSE Group Header Records

Position	Length	Field	Characteristics
Group Header 1			
1	12	Ticker	Alphanumeric
14	1	Usage Number	Integer
16	9	CUSIP	Alphanumeric
26	45	Name	Alphanumeric
72	24	Business	Alphanumeric
97	1	Foreign Flag	1 if Foreign, 0 otherwise
Group Header 2			
1	5	Daily data start date relative calendar number	Integer
7	5	Number of daily data records	Integer
13	5	Monthly data start date relative calendar number	Integer
19	5	Number of monthly data records	Integer
25	5	Number of price adjustments	Integer

The format of the first group header is (rev 2021-05-10):

$$(A12,1X,I1,1X,A9,1X,A45,1X,A24,1X,I1)$$

The second header record has the format:

$$(5(I5,1X))$$

4.2.2 Daily Data Records

The first record immediately following the second security header line is the first daily data record if the number of daily data records in the second header line is greater than 0. The date for this record can be found by looking up the Julian calendar date for the daily data start date relative calendar number in the relative calendar/index file (DAILY.IDX). Subsequent daily data records have sequential daily relative calendar numbers.

Please see Table 3 for the content and layout of each daily record. Daily data records have the following format:

$$(8F9.3,I8,I9,2I6,F10.6)$$

Table 3: Record Layout for DATA.TSE Daily Data Records

Position	Length	Field	Characteristics
1	9	Opening price	Numeric, DDDDD.ddd
10	9	Opening bid	Numeric, DDDDD.ddd
19	9	Opening ask	Numeric, DDDDD.ddd
28	9	Closing price	Numeric, DDDDD.ddd
37	9	Closing bid	Numeric, DDDDD.ddd
46	9	Closing ask	Numeric, DDDDD.ddd
55	9	High	Numeric, DDDDD.ddd
64	9	Low	Numeric, DDDDD.ddd
73	8	Transactions	Integer
81	9	Volume	Integer
90	6	Quotes	Integer
96	6	Quote changes	Integer
102	10	Return	Numeric, DDD.dddddd

4.2.3 Monthly Data Records

The first record immediately following the last daily data record is the first monthly data record if the number of monthly data records in the second header line is greater than 0. As with the daily data records, the dates associated with monthly records can be found by looking them up in the monthly calendar/index file (Monthly.idx).

See Table 4 for the content and layout of each monthly record. The format of monthly data records is:

$$(F9.3, I9, I10, I11, F13.6, F7.3, F11.6, 1X, I1)$$

Table 4: Record Layout for DATA.TSE Monthly Data Records

Position	Length	Field	Characteristics
1	9	Closing price	Numeric, DDDDD.ddd
10	9	Total transactions	Integer
19	10	Total volume	Integer, in 100s of shares
29	11	Shares outstanding	Integer, in 100s of shares
40	13	Earnings per share	Numeric, DDDDDD.dddddd
53	7	Beta	Numeric, DDD.ddd
60	11	Monthly return	Numeric, DDDD.dddddd
72	1	Dividend flag	1 if there are price adjustments this month, 0 otherwise

4.2.4 Price Adjustment Data Records

Each price adjustment record contains from one to five data clusters and each cluster contains data on one price adjustment. Cluster contents are:

Table 5: Record Layout for DATA.TSE Price Adjustment Clusters
(up to 5 per record)

Position	Length	Field	Characteristics
1	11	Dividend	Numeric, DDDD.dddddd
12	1	Dividend flag 1	Integer, single digit
13	1	Dividend flag 2	Integer, single digit
14	1	Dividend flag 3	Integer, single digit
15	8	Ex-dividend date	YYYYMMDD

Each cluster is formatted as (rev 2021-05-10)

$$(F11.6,3I1,I8)$$

So, the format for a complete record (five clusters) is

$$(5(F11.6,3I1,I8))$$

Note that there is one cluster for every price adjustment indicated in the second stock header record. If this number, say *N*, is greater than 0, there will be

$$INT\left(\frac{N-1}{5}\right) + 1$$

price adjustment records for this security.

4.3 Monthly Calendar/Index File - **MONTHLY.IDX** and **OLD_MONTHLY.IDX**[3]

Each record in the monthly index file represents one month. Records are fixed-length and contain 359 data characters and a terminating linefeed for a total of 360 characters per record. Records are organized sequentially starting with relative monthly calendar month number one. Please see Table 6 for record contents and layout. Note that, presently, each field in the file is separated by one, or more, spaces.

The format is (rev 2021-05-10):

$$(I4,1X,I8,1X,F8.5,1X,F9.6,3(1X,F8.5),15(1X,F8.2,1X,F8.2),3(1X,F9.6))$$

---

[3] The length of date fields increased from 6 to 8 characters beginning with the 2021 Q1 release of the database. Thus, records in OLD_MONTHLY.IDX are two characters shorter than reported here and use the following format:

$$(I4,1X,I6,1X,F8.5,1X,F9.6,3(1X,F8.5),15(1X,F8.2,1X,F8.2),3(1X,F9.6))$$

Table 6: Record Layout for MONTHLY.IDX

Position	Length	Field	Characteristics
1	4	Relative month number	Numeric (integer)
6	8	End of Month Date	YYYYMMDD
15	8	91 Day T-Bill Rate	DD.ddddd (numeric, includes '.' and five decimal places)
24	9	30-day Return on T-Bills	DDDDD.ddd
34	8	Long Term Government Bond Rate	DDDD.ddd
43	8	Corporate Bond Rate	DDDD.ddd
52	8	Foreign Exchange Rate (Cdn$ per US$)	DDDD.ddd
61	8	S&P/TSX Composite Monthly Price Index	DDDDD.dd
70	8	S&P/TSX Composite Monthly Total Return Index	DDDDD.dd
79	8	Sector 10/Group 1 Monthly Price Index	DDDDD.dd
88	8	Sector 10/Group 1 Monthly Total Return Index	DDDDD.dd
97	8	Sector 15/Group 2 Monthly Price Index	DDDDD.dd
106	8	Sector 15/Group 2 Monthly Total Return Index	DDDDD.dd
115	8	Sector 20/Group 3 Monthly Price Index	DDDDD.dd
124	8	Sector 20/Group 3 Monthly Total Return Index	DDDDD.dd
133	8	Sector 25/Group 4 Monthly Price Index	DDDDD.dd
142	8	Sector 25/Group 4 Monthly Total Return Index	DDDDD.dd
151	8	Sector 30/Group 5 Monthly Price Index	DDDDD.dd
160	8	Sector 30/Group 5 Monthly Total Return Index	DDDDD.dd
169	8	Sector 35/Group 6 Monthly Price Index	DDDDD.dd
178	8	Sector 35/Group 6 Monthly Total Return Index	DDDDD.dd
187	8	Sector 40/Group 7 Monthly Price Index	DDDDD.dd
196	8	Sector 40/Group 7 Monthly Total Return Index	DDDDD.dd
205	8	Sector 45/Group 8 Monthly Price Index	DDDDD.dd
214	8	Sector 45/Group 8 Monthly Total Return Index	DDDDD.dd
223	8	Sector 50/Group 9 Monthly Price Index	DDDDD.dd
232	8	Sector 50/Group 9 Monthly Total Return Index	DDDDD.dd
241	8	Sector 55/Group 10 Monthly Price Index	DDDDD.dd
250	8	Sector 55/Group 10 Monthly Total Return Index	DDDDD.dd
259	8	S&P/TSX 60/Group 11 Monthly Price Index	DDDDD.dd
268	8	S&P/TSX 60/Group 11 Monthly Total Return Index	DDDDD.dd
277	8	S&P/TSX Mid Cap/Group 12 Monthly Price Index	DDDDD.dd
286	8	S&P/TSX Mid Cap/Group 12 Monthly Total Return Index	DDDDD.dd
295	8	S&P/TSX Small Cap/Group 13 Monthly Price Index	DDDDD.dd
304	8	S&P/TSX Small Cap/Group 13 Monthly Total Return Index	DDDDD.dd
313	8	S&P/TSX Venture/Group 14 Monthly Price Index	DDDDD.dd
322	8	S&P/TSX Venture/Group 14 Monthly Total Return Index	DDDDD.dd
331	9	CFMRC Monthly Equal Weighted Index	DD.dddddd (numeric, includes '.' and six decimal places)
341	9	CFMRC Monthly Value Weighted Index	DD.dddddd
351	9	30 Day Return on Long Term GoC Bonds	DD.dddddd

4.4 Tickers & Names File - **TICKERS.DAT**

The ticker file contains a list of all securities included in the current version of the database. Only securities listed in this file are in DATA.TSE and vice versa. There is one 104-character record for each security which, as with other files, is terminated with a linefeed for a 105-character record. See Table 7 for record contents.

Table 7: Record Layout for TICKERS.DAT

Position	Length	Field	Characteristics
1	12	Ticker	Alphanumeric
14	1	Usage Number	Integer, single digit
16	9	Header Record Number in **Data.tse**	Long Integer
26	45	Name	Alphanumeric
72	31	Business	Alphanumeric
104	1	Common Flag	"C" – common, "N" - otherwise

The format is:

$$(A8,1X,I1,1X,I9,1X,A45,1X,A31,1X,A1)$$

4.5 Ticker Dictionary File - **DICTION.DAT**

Every security listed on the TSX has a record in the dictionary file describing ticker changes over time. Each record contains a Delisting Date and one, or more, groups of {Ticker, Usage Number, and Listing Date}.

The format is (rev 2021-05-10):

$$(I8, n\{1X,A12,1X,I1,1X,I8\})$$

where n is the number of ticker groups.

Table 8: Record Layout for DICTION.DAT

Position	Length	Field	Characteristics
1	8	Delisting Date	Integer
10	12	Ticker	Alphanumeric
23	1	Usage Number	Integer, single digit
25	8	Listing Date	Integer

When a record has more than one group, the ticker has changed since initial listing. The left-most group describes the current or last used ticker. Each group to the right then describes the immediately prior ticker and so on, until you reach the last group in the record (right-most) which is the initial ticker.

The Delisting Date is the last date that any ticker was used for this security on the Exchange. There may be trades and/or quotes for this security on the Delisting Date. The Listing Date, of which there may be several in a ticker history record, is the first date on which a particular ticker was used for this security on the TSX.[4]
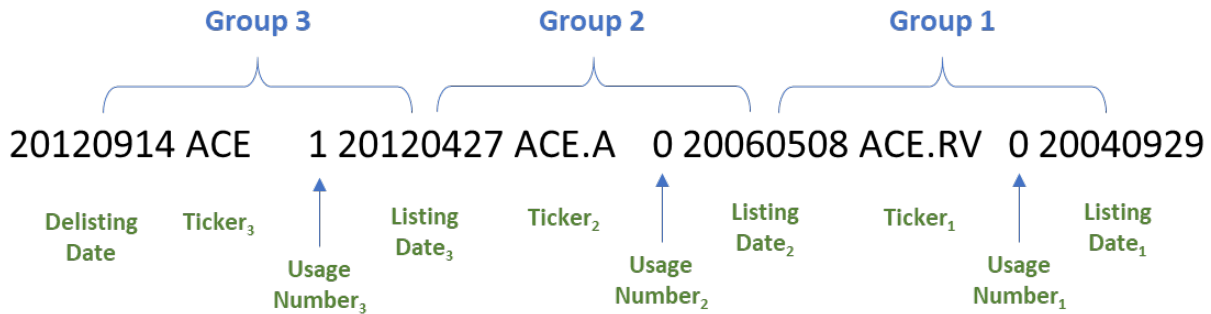

Figure 4. Example of a Ticker History Record

Figure 4 shows a ticker history record for ACE(1) (the most recent ticker in the chain). Group 1 shows us this security began with a ticker of ACE.RV(0) which was listed on 2004-09-29. Per the information in Group 2, this security began trading on 2006-05-08 under the ticker ACE.A(0) consequently, it stopped trading under the ticker ACE.RV(0) after the *previous trading day* (2006-05-05, see DAILY.IDX). The last ticker in the database for this security is ACE(1) which it changed to on 2012-04-27 (see Group 3). ACE.A(0) stopped being used after the previous trading day, in this case 2012-04-26.

The delisting date in the record is 2012-09-14 so there is no information on this security in the database after this date.

All data in the Summary Information database is reported under the most recent ticker for a security. Thus, you will find trade and quote information for the security in Figure 4 under ACE(1) in the database files.[5]

5. Intra-daily Trade and Quote (TAQ) Data Description

The TSX/CFMRC TAQ Databases include the TSX TAQ Database and the TSXV TAQ Database which contain intra-daily data for all Toronto Stock Exchange listed securities and all TSX Venture Exchange listed securities, respectively. With the exception that file and directory names are slightly different depending on the database, data files, definitions, and field specifications are identical for both databases. All TSX TAQ database file names begin with TSX… and are in the [Database Destination]\TSX\

---

[4] There are cases where some new issues begin trading on the TSX on a "when issued" basis. In the past, it wasn't always possible to tell when this happened. Securities are added to the Summary Information database when they officially become listed. Currently, most "when issued" securities are traded with a .W ticker extension (e.g. AEC.W) and the ticker changes when the security becomes officially listed. In these cases, the "when issued" ticker is added to the database and a ticker change is noted on the official listing date.

[5] This **DOES NOT** apply to the Trade-and-Quote databases. Tickers in the TAQ databases are the tickers used by securities on the date the data was observed on the exchange.

directory. All TSXV TAQ database file names begin with TSXV… and are in the [Database Destination]\TSXV\ directory. For brevity, examples will only use the TSX TAQ database.

There are three different types of data in the TSX/CFMRC TAQ databases:

1. Intra-daily trade and quote data.
2. Daily name index data.
3. Master TAQ index.

As mentioned above, TAQ data is organized in daily compressed packages for each exchange. Each package contains one TAQ text file for each security that was listed on the relevant exchange on the package date. In addition, the daily packages contain an index file that lists all active securities on the relevant exchange on the relevant date.

These data are explained below.

5.1 Intra-daily Trade and Quote Data

All records in the daily TAQ files begin with a record type. The remaining information depends on the record type.

Record Type – a single character that indicates the type of data contained in the record. "D" is a date record, "T" is a trade record, and "Q" is a quote record.

5.1.1 Date Records

There is only one Date record in each daily TAQ file. The Date record identifies the characteristics common to all records in the particular daily file.

Date – the date in YYYYMMDD format.

Ticker – the security ticker in use on the date of the TAQ file. Unlike the Summary Information database, ticker changes are not tracked in the TAQ databases. Ticker changes will be reflected in daily TAQ files on, and after, the date a firm changes its ticker.

Name – the name of the security on the date of the TAQ file. Name changes will be reflected in daily TAQ files on, and after, the date a firm changes its name.

CUSIP – a nine character (including letters and numbers) code that uniquely identify a company or issuer and the type of financial instrument. A CUSIP number identifies most financial instruments, including: stocks of all registered U.S. and Canadian companies, commercial paper, and U.S. government and municipal bonds.

Business - a brief description, defined by the TSX, of the type of business this firm is involved in.

5.1.2 Trade and Quote Records

Time – the time of the transaction in nanoseconds (HHMMSSnnnnnnnnn). There is no decimal point but the last nine digits are nanoseconds.

Sequence Number – a unique number that sequences trade and quote records across all securities. When two transactions have the same Time, the Sequence Number may be used to identify which was executed first.

5.1.3 Trade Records

Price – the price of the trade per share in tenths of a cent ($$$$CCC). There is no decimal point but the last three digits are thousandths of a dollar.

Shares – number of shares traded.

Buyer – the Trading Number of the Participating Organization on the buyer side of the trade. See the *TSX eReview* for the date of interest or go to the TSX.com POs page (currently: https://www.tsx.com/trading/accessing-our-markets/member-firm-directory?l=A) for firm names and their associated number. A Buyer number of 1 means the buyer is anonymous (Rev 20-06-30).

Seller - the Trading Number of the Participating Organization on the selling side of the trade. See Buyer for information on how to look these up. A Seller number of 1 means the seller is anonymous (Rev 20-06-30).

Odd Lot – a marker indicating whether a trade was for an "odd lot" – that is, not a round lot. "1" indicates an odd lot trade, otherwise blank (" ").

Round lot size for a security depends on the closing price of the security *on the previous trading day*:

Previous Day Close (per unit)	Board Lot Size
less than $0.10	1,000 units
$0.10 to less than $1.00	500 units
$1.00 or more	100 units

Trade Session – a single character that indicates the type of trading session (please see https://www.tsx.com/trading/calendars-and-trading-hours/trading-hours for more information):

"A" - *Continuous Market* - 9:30 AM to 4:00 PM, during this time, all regular order types are accepted.

"O" - *Market on Open* (MOO) or *Open* – 9:30 AM, all matching orders are executed at a single opening trade price with any remaining orders carrying through to the continuous limit order book.

"M" - *Market on Close* (MOC) – 3:40 PM to 4:00 PM, cut-off for MOC orders is at 3:40. MOC imbalance is published at 3:40 after which Limit on Close (LOC) orders opposite to the imbalance side are accepted (subject to a price collar) into the non-displayed MOC facility. Trades publish at 4:00 unless a Price Movement Extension is required, in which case trades publish at 4:10.

"C" - *Extended Trading Session* or *Crossing Session* -4:10 PM – 5:00 PM, orders at the last sale price are accepted, but trades may only occur at the last sale price except for regulatory approval of a specialty price cross. Day orders participate in this session. MBF session for option expiry takes place during Extended Trading once per month, the evening before an option expiry day.

Cancellation – a marker on a previously entered trade that indicates the trade is cancelled. The time on this record reflects the time the trade was cancelled. "1" indicates a Cancellation, otherwise blank (" ").

Cancelled – a marker on the original trade to indicate the trade has been cancelled (there will also be a subsequent Cancellation record). "1" indicates trade was Cancelled, otherwise blank (" ").

Correction – a marker indicating a trade correction. "1" indicates a Correction, otherwise blank (" ").

Delayed Delivery – a marker indicating the trade is a special terms trade in which there is a clear understanding between the buying and selling parties that the delivery of the securities will be delayed beyond the usual three-day settlement period. "1" indicates trade is Delayed Delivery, otherwise blank.

Cash – a marker indicating special terms attached to an equity trade that requires the trade to be settled either the same day or the following business day for cash. "1" indicates trade is Cash, otherwise blank.

Non Net – a marker indicating there is a clear understanding between the buying and selling parties that they will settle the trade directly with each other. "1" indicates trade is Non Net, otherwise blank.

Special Terms – a marker indicating the trade has special conditions. For example, a cash trade will be settled sooner than the usual three-day settlement period.

Specialty Cross – a single character that indicates whether a cross is in one of the four "Specialty Price Cross" categories. Generally, a cross is a trade that occurs when two accounts within the same Participating Organization/Member wish to buy and sell the same security at an agreed price and volume. Please see https://www.tsx.com/trading/tsx-alpha-exchange/order-types-and-features/cross-facilities?lang=en for more details on TSX Crossing Facilities.

> "B" – *Basis Cross* - A trade whereby a basket of securities or an index participation unit is transacted at prices achieved through the execution of related exchange-traded derivative instruments which may include index futures, index options and index participation units in an amount that will correspond to an equivalent market exposure.

> "V" – *VWAP Cross* - A transaction for the purpose of executing a trade at a volume-weighted average price of a security traded for a continuous period on or during a trading day.

> "C" – *Contingent Cross* - A trade resulting from a paired order placed by a Participant on behalf of a client to execute an order on a security that is contingent on the execution of a second order placed by the same client for an offsetting volume of a related security.

> "I" – *Internal Cross* - An intentional cross at or between the best bid/ask and between two client accounts of a Participating Organization which are managed by the Participating Organization as a portfolio manager with discretionary authority and is not subject to interference.

> "S" – *Special Trading Session Cross* - A closing price cross resulting from an order placed by a Participant on behalf of a client for execution in the Special Trading Session at the last sale price.

> Blank – not a Specialty Cross

5.1.4 Quote Records

Bid Price – the price bid per share, in tenths of a cent ($$$$CCC).

Ask Price – the price asked per share, in tenths of a cent ($$$$CCC).

Bid Size – how many shares the bidder is willing to buy at the bid price. In Boardlots, not shares.

Ask Size – how many shares the seller is willing to sell at the ask price. In Boardlots, not shares.

Halted – a marker that indicates whether the market for this security was Halted at the time of the quote. "1" indicates market was Halted, otherwise blank.

5.2 Daily Name Index Data

Ticker – the security ticker in use on the date of the TAQ file.

Transactions – the number of trade and quote records present in the daily TAQ file for this ticker on this date. A value of 0 means the security was listed on this date but had neither quotes nor trades. In this case, there will not be a daily TAQ ticker file for this security in the compressed package.

Name – the name of the security on the date of the TAQ file.

CUSIP – the CUSIP of the security on the date of the TAQ file.

Business – a brief description, defined by the TSX, of the type of business this firm is involved in.

5.3 Master TAQ Index

Each TAQ database also has a relational master index composed of a Master table and five index tables.

Record Numbers in the Master TAQ Index

Both the Master table and the index tables use record numbers to refer to records in the master TAQ index tables. These *record numbers are 0-based references* (that is, the first record is record 0, not record 1) and, in cases where files may include comments or headers, only refer to actual data records (comments and headers are *not included* in the record count).

5.3.1 Index Tables

The index tables all follow the same format. Each index record contains a key (a possible value for the index variable) and a list of references to records in the Master table that contain this key. Key values are unique within an index table and index tables only include keys used in the Master table. The list of references is of variable length, depending on how often the key is used in the Master table. The five index key types are: *Ticker*, *Name*, *CUSIP*, *Business*, and *Date*.

Key – a value for the index variable of the index table. Keys are unique within a table and are present in one, or more, records in the Master table.

List of References – a variable length list of integer references to records in the Master table that contain this key value. References are 0-based record numbers and do not count comments or headers in table files.

5.3.2 Master Table

The Master table contains information on each daily TAQ ticker file in the database and makes it faster to select securities that meet various criteria. References in the Master table are 0-based record numbers and refer to records in the corresponding Index table. Each record in the Master table refers to one daily TAQ ticker file and all daily TAQ ticker files are included in the table once. Master table records include the following information about the data in each daily TAQ file:

Ticker Reference – the 0-based record number of the Ticker in the *Ticker Index* table.

Name Reference – the 0-based record number of the Name in the *Name Index* table.

CUSIP Reference – the 0-based record number of the CUSIP in the *CUSIP Index* table.

Business Reference – the 0-based record number of the Business in the *Business Index* table.

Date Reference – the 0-based record number of the Date in the *Date Index* table.

TAQRecords – the number of trade and quote records in the file.

6. TAQ Database File Formats

Because there are so many different possible file names in the TAQ databases, the following notation is used for brevity. Please replace the terms in [] with one of the following:

> [Exchange] – one of: TSX or TSXV.
> [YYYYMMDD] – a date including the 4-digit year and 2-digit month and day.
> [Ticker] – any possible ticker, including extensions separated by '.' (e.g. ABK.PR.C).
> [Index] – one of: Ticker, Name, CUSIP, Business or Date.

6.1 Daily TAQ Files – [Exchange]_[YYYYMMDD]_[Ticker].TXT

These files are found in the daily compressed packages found in the [Database Destination]\[Exchange]\ directory. Daily compressed packages have file names that follow the [Exchange][YYYYMMDD].ZIP format.

The daily TAQ ticker files contain one record per line. There are three different record types (Date, Trade, and Quote) and each type has a different record length so records usually need to be read sequentially. The type of each record is determined by the first character of the record and record data begins with the second character.

6.1.1 Date Records

Date records use a tab-delimited format: fields (aside from the record type) are terminated with a tab character (<Tab>, decimal 9). Fields have no white space (additional spaces) before or after their data. The last field in the record is terminated with a linefeed (<LF>).

Table 9: Record Layout of Date Records in [Exchange]_[YYYYMMDD]_[Ticker].TXT

Position[6]	Length	Field	Characteristics
1	1	Record Type	"D" – Date record, single character
2	<Tab>	Date	YYYYMMDD
(3)	<Tab>	Ticker	Alphanumeric, 12 characters, max.
(4)	<Tab>	Name	Alphanumeric, 45 characters, max.
(5)	<Tab>	CUSIP	Alphanumeric, 9 characters
(6)	<LF>	Business	Alphanumeric, 31 characters, max.

6.1.2 Trade Records

Table 10: Record Layout of Trade Records in [Exchange]_[YYYYMMDD]_[Ticker].TXT

Position	Length	Field	Characteristics
1	1	Record Type	"T" – Trade record, single character
2	15	Time	HHMMSSnnnnnnnnn (nanoseconds)
17	9	Sequence Number	Numeric
26	7	Price	Numeric, $$$$CCC
33	9	Shares	Integer
42	3	Buyer	Integer
45	3	Seller	Integer
48	1	Odd Lot	Marker: On = "1", Off = blank
49	1	Trade Session	"A" - Continuous Market; "O" - Open Trade; "M" - Market On Close Trade; "C" - Crossing Session, single character
50	1	Cancellation	Marker: On = "1", Off = blank Cancellation record appearing at the time the trade is cancelled
51	1	Cancelled	Marker: On = "1", Off = blank Marker on original trade to indicate the trade has been cancelled
52	1	Correction	Marker: On = "1", Off = blank
53	1	Delayed Delivery	Marker: On = "1", Off = blank
54	1	Cash	Marker: On = "1", Off = blank
55	1	Non Net	Marker: On = "1", Off = blank
56	1	Special Terms	Marker: On = "1", Off = blank
57	1	Specialty Cross	"B" - Basis Cross; "V" - VWAP Cross; "C" - Contingent Cross; "I" - Internal Cross; "S" - Special Trading Session Cross, single character

---

[6] A position enclosed in () means this is the field number in records with delimited fields rather than a character or byte location within the record.

6.1.3 Quote Records

Table 11: Record Layout of Quote Records in [Exchange]_[YYYYMMDD]_[Ticker].TXT

Position	Length	Field	Characteristics
1	1	Record Type	"Q" – Quote record, single character
2	15	Time	HHMMSSnnnnnnnnn (nanoseconds)
17	9	Sequence Number	Numeric
26	7	Bid Price	Numeric, $$$$CCC
33	7	Ask Price	Numeric, $$$$CCC
40	3	Bid Size	Integer, in Boardlots
43	3	Ask Size	Integer, in Boardlots
46	1	Halted	Marker: On = "1", Off = blank

6.2 Daily Name Index Files – **[YYYYMMDD][Exchange].INDEX**

Each daily compressed package contains a Daily Name Index File. Please look for these files in the corresponding [Exchange][YYYYMMDD].ZIP file.

Daily Name Index file records use a tab-delimited format: fields are terminated with a tab character (<Tab>, decimal 9). Fields have no white space (additional spaces) before or after their data. The last field in the record is terminated with a linefeed (<LF>).

Table 12: Record Layout for [YYYYMMDD][Exchange].INDEX

Position	Delimiter	Field	Characteristics
(1)	<Tab>	Ticker	Alphanumeric, 12 characters, max.
(2)	<Tab>	Transactions	Integer
(3)	<Tab>	Name	Alphanumeric, 45 characters, max.
(4)	<Tab>	CUSIP	Alphanumeric, 9 characters
(5)	<LF>	Business	Alphanumeric, 31 characters, max.

6.3 Index Table Files – [Exchange][Index]INDEX.TXT

There are five Index Table files for each TAQ database, one for each of the following indices: *Ticker*, *Name*, *CUSIP*, *Business*, and *Date*. These files all follow the same formatting rules.

Each record has a *Key*, and a *List of References*. Because the length of the List of References is not bounded, records in these files often won't fit on a single line. Therefore, records in the Index Table files are delimited with the "{" special character or an end-of-file (<EOF>). The "{" character actually begins a new record and explicitly terminates the previous record.

Records begin with a unique, variable length key which is terminated with a "}<LF>" (that is, the key is the only thing on the first line of a record). Lines following the key in record contain 0-based references to records in the Master table that contain this key. There is only one reference per line but there may be many lines and thus, references. The keys in a particular index table include all values in the TAQ files for

that variable and only values that are in the TAQ files. The keys in the TSXTickerIndex.txt file are all tickers, and include all tickers in the TSX TAQ files and only those tickers. Likewise, the keys in the Date index table are dates, etc.
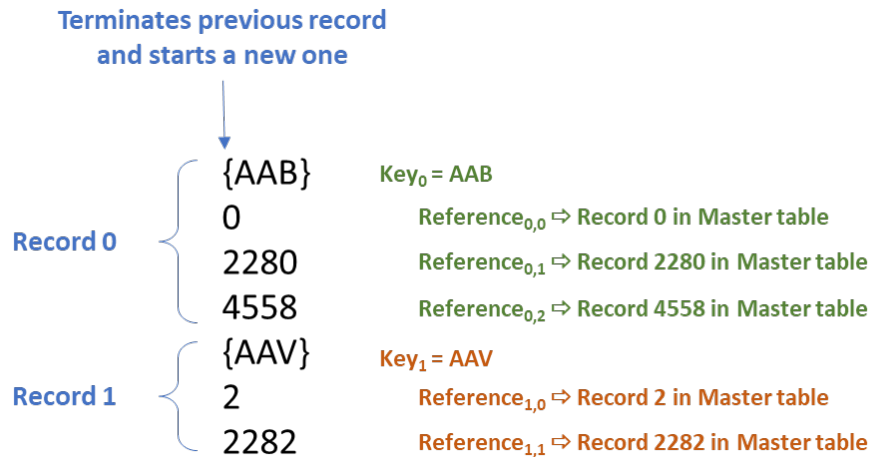
Figure 5 shows a simple Ticker Index table.

Figure 5. Example of a Ticker Index Table.

{1ST ASSET GLOBAL FINANCIAL SECTOR ETF UN}
802
3082
{GREAT-WEST LIFECO INC. 3.65% 1ST PR SERIES N}
906

Figure 6. Example of a Name Index Table.

Figure 6 shows a simple Name Index table. Notice that the keys here are security names. The references in the index records are still to Master table records but they are now Master table records that use these security names.

6.4 Master Table File – **[Exchange]MASTERINDEX.TXT**

The Master table file (there is one for each TAQ database) begins with a header record. The header record begins with the "#" character and continues with a comma-delimited list of variable names or headers that reflect the data in each of the corresponding fields in each Master record. This record is strictly for convenience – it isn't used anywhere in the databases. The current header is:

#TickerIndex,NameIndex,CUSIPIndex,BusinessIndex,DateIndex,TAQRecords

Each line following the header line is a Master table record and contains six comma-delimited integer fields. Master table records are 0-based referenced so the first record after the header is Master record 0. The first five fields in the Master table records reference records in each of the Index tables: *Ticker*, *Name*, *CUSIP*, *Business*, and *Date* for fields one through five, respectively. Field six is the number of TAQ records in the corresponding daily TAQ ticker file.

By using the references in the Master table record, you can find the key values for each of the variables (ticker, name, CUSIP, business, and date) associated with the record and quickly look up related Master table records.

Table 13: Record Layout for [Exchange]MASTERINDEX.TXT

Position	Delimiter	Field	Characteristics
(1)	","	Ticker Reference	Integer
(2)	","	Name Reference	Integer
(3)	","	CUSIP Reference	Integer
(4)	","	Business Reference	Integer
(5)	","	Date Reference	Integer
(6)	<LF>	Transactions	Integer