

CRSP/COMPUSTAT RELEASE NOTES

JULY 2020 QUARTERLY UPDATE

NEW! CRSP/COMPUSTAT SAS, ASCII, AND R FILES

CRSP is providing new access to our subscribers for the CRSP\Compustat Merged Database (CCM) in SAS, ASCII, and R file formats. These files are intended for use with the CRSP SAS and ASCII stock files and Compustat data files that contain GVKEYs and IIDs.

While a full subscription is required for either, for convenience and efficiency for some users, two packages are available:

- CFZ – Complete set of all files
- CLZ – Link package - a subset of the full CRSP/Compustat SAS, ASCII and R data files. The set consists only of files used for linking the CRSP and Compustat data. These files are intended for use with the CRSP SAS, ASCII and R stock files and Compustat data files that contain GVKEYs and IIDs.

These files are covered in existing CCM license agreements. To request them for download through CRSP's MOVEit site, please email support@crsp.org or call 312-263-6400, Option 2.

ABOUT THE JULY 2020 CRSP/COMPUSTAT MERGED DATABASE

The cut date for these release notes is 20200711.

CALENDARS

The CCM database utilizes the CRSP stock calendars to align properly with CRSP data. Data in the CRSP daily Stock (and Indexes) file use daily and annual calendars. Data in the monthly files use monthly, quarterly and annual calendars.

CALENDAR	DATE RANGE	DATE INDEX RANGE
DAILY TRADING CALENDAR	19251231-20200710	1-24925
MONTHLY TRADING CALENDAR	19251231-20231231	1-1189
ANNUAL TRADING CALENDAR	19251231-20201231	1-97
QUARTERLY TRADING CALENDAR	19251231-20201231	1-385
SEMI-ANNUAL TRADING CALENDAR	19251231-20201231	1-193
WEEKLY TRADING CALENDAR	19260102-20200710	1-4932

DATABASE CHARACTERISTICS

The following table summarizes general characteristics of the CCM database:

ATTRIBUTE	CCM FORMAT DATABASE
SOURCE	COMPUSTAT XPRESSFEED
VOLUME LABELS	CMZ20200710 - CRSPACCESS CFZ20200710 - SAS/ASCII/R CLZ20200710 (LINK ONLY) - SAS/ASCII/R
DELIVERY	MOVEIT CLOUD
CRSPACCESS VERSION	3.22 AND HIGHER – TS_PRINT CCM_PRINT
CRSPSIFT VERSION (WINDOWS ONLY)	YES – ALL CURRENTLY SUPPORTED VERSIONS
SASEXCCM ENGINE	VERSION 9.4
C	YES – CA3.22 AND HIGHER
NUMBER OF DATA ITEMS	OVER 4,000 QUARTERLY & ANNUAL DATA ITEMS

DATABASE PARAMETERS

TOTAL COMPUSTAT RECORDS	50,576
TOTAL GVKEY COMPANIES	48,378
TOTAL GVKEYX INDEXES	2,198
TOTAL GVKEY/IID SECURITIES	62,588
TOTAL CRSP SECURITIES LINKED	30,231
TOTAL GVKEYS IN LINK	48,328
TOTAL LINK RECORDS	106,420
TOTAL LINK RECORDS WITH A LINK TO CRSP	35,076

DATABASE SIZES

CMZ	11.79 GB
CFZ - ASCII	10.68 GB
CFZ - SAS	29.36 GB
CFZ - R	1.43 GB
CLZ - ASCII	1.10 GB
CLZ - SAS	1.79 GB
CLZ - R	45.09 MB

JULY 2020 LINK EDITS

GVKEY 111629 CREXENDO INC

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19970701	19980629	0	0	00X	NR	C
19980630	19991117	0	0	01	NR	P
19991118	99999999	87404	17399	01	LU	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19970701	19980629	0	0	00X	NR	C
19980630	19991117	0	0	01	NR	P
19991118	20150112	87404	17399	01	LC	P
20150113	99999999	0	0	01	NR	P

JUNE 2020 NEW DATA ITEMS

Listed in the tables below are new data items and new keysets for existing data items that have been added this month.

ITEM MNEMONIC	ITEM NAME	KEYSET	KEYSET_TAG	DESCRIPTION
CEIEXBILL_FN	COST & EARNINGS IN EXCESS OF BILLINGS FN	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
CEIEXBILL	COST & EARNINGS IN EXCESS OF BILLINGS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION

JUNE 2020 LINK EDITS

GVKEY 61140 WINDTREE THERAPEUTICS INC

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19940101	19950807	0	0	00X	NR	C
19950808	20170504	82156	13915	01	LC	P
20170505	99999999	0	0	01	NR	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19940101	19950807	0	0	00X	NR	C
19950808	20170504	82156	13915	01	LC	P
20170505	20200519	0	0	01	NR	P
20200520	99999999	82156	13915	01	LC	P

GVKEY 64229 DELMAR BANCORP

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19970331	99999999	0	0	01	NU	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19970331	20200505	0	0	01	NR	P
20200506	99999999	19396	57086	01	LC	P

GVKEY 140023 DELCATH SYSTEMS INC

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19990101	20011021	0	0	00X	NR	C
20011022	20170914	89169	40199	01	LC	P
20170915	99999999	0	0	01	NR	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19990101	20011021	0	0	00X	NR	C
20011022	20170914	89169	40199	01	LC	P
20170915	20200430	0	0	01	NR	P
20200501	99999999	89169	40199	01	LC	P

GVKEY 34417 BOXWOOD MERGER CORP

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20181130	20181231	0	0	01	NR	P
20190101	20190106	0	0	01	NR	C
20190107	20200228	0	0	01	NR	N
20190107	99999999	18357	56586	02	LC	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20181130	20181231	0	0	01	NR	P
20190101	20190106	0	0	01	NR	C
20190107	20200228	0	0	01	NR	N
20190107	20200217	18357	56586	02	LC	P
20200228	20200228	0	0	03	NR	J

GVKEY 20810 INVENTERGY GLOBAL INC

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20130101	20140608	0	0	00X	NR	C
20140609	20151207	87604	17932	01	LC	P
20151208	99999999	87604	17932	01	LC	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20130101	20140608	0	0	00X	NR	C
20140609	20170606	87604	17932	01	LC	P
20170607	99999999	0	0	01	NR	P

GVKEY 108645 ULTRA PETROLEUM CORP

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20010118	20160502	88882	41178	01	LC	P
20160503	20170331	0	0	01	NR	P
20170401	20170412	0	0	01	NR	C

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20170413	99999999	0	0	01	NR	N
19900731	20010117	0	0	01C	NU	C
20010118	20040331	88882	41178	01C	LX	N
20170413	99999999	16682	41178	02	LC	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20010118	20160502	88882	41178	01	LC	P
20160503	20170331	0	0	01	NR	P
20170401	20170412	0	0	01	NR	C
20170413	99999999	0	0	01	NR	N
19900731	20010117	0	0	01C	NU	C
20010118	20040331	88882	41178	01C	LX	N
20170413	20190807	16682	41178	2	LC	P
20190808	99999999	0	0	2	NR	P

MAY 2020 DATA ITEMS

Listed in the tables below are new data items and new keysets for existing data items that have been added this month.

ITEM MNEMONIC	ITEM NAME	KEYSET	KEYSET_TAG	DESCRIPTION
HFO_FN	OTHER % OF REVENUE FOOTNOTE	1	STD	INDUSTRIAL FORMAT, CONSOLIDATED INFORMATION, STANDARDIZED PRESENTATION
HFMCD_FN	MEDICAID % OF REVENUE FOOTNOTE	1	STD	INDUSTRIAL FORMAT, CONSOLIDATED INFORMATION, STANDARDIZED PRESENTATION
HFMGD_FN	MANAGED CARE % OF REVENUE FOOTNOTE	1	STD	INDUSTRIAL FORMAT, CONSOLIDATED INFORMATION, STANDARDIZED PRESENTATION
BILLEXCEQ	BILLINGS IN EXCESS OF COST & EARNINGS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
CEIEXBILLQ_FN	COST & EARNINGS IN EXCESS OF BILLINGS FN	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
CEIEXBILLQ	COST & EARNINGS IN EXCESS OF BILLINGS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION

MAY 2020 LINK EDITS

GVKEY 183612 ZOOM TECHNOLOGIES INC

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20091010	99999999	77011	10759	01	LC	C

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20090101	20091009	0	0	00X	NR	C
20091010	20140829	77011	10759	01	LC	C
20140830	99999999	0	0	01	NR	C

CHANGES

For subscribers of the existing CCM database, changes to organization of the data has been kept to a minimum. Reviewing the data at the level of detail required to create these files benefits all subscribers through data corrections and cleanup.

CRSP AND COMPUSTAT LINKING

Files include PERMNO and PERMCO fields already associated with GVKEYs and IIDs. Linking information is available for details, but the work is done!

METADATA FILES

Metadata files are included with the SAS, ASCII and R formats:

- D_ItemInfo
- D_FileInfo
- D_ItemRelationship
- D_ItemDensity
- D_KeysetInfo
- D_CommonItems
- D_FiscalPeriod

DATA ITEM NAMES & DESCRIPTIONS

Item Names have been cleaned. Names exceeding field width have been shortened using conventions. Names are no longer truncated. Typographical errors have been corrected. CRSP has made every effort to make item names match those that are provided by Compustat.

MNEMONICS

CRSP database structure requires that no itemids or mnemonics are duplicated. In cases where there are duplicates, CRSP has modified the mnemonic used to identify the item.

- idx_ Prefixes - all quarterly and annual index items mnemonics have been prepended with "idx_"
- sec_ Prefixes – Three items in the SecurityTimeSeriesData table have had prefixes added in order to retain uniqueness:
 - sec_curcdq – ISO Currency Code
 - sec_nisa – Net Income (Loss)
 - sec_nisa_dc – Net Income (Loss) Data Code

TABLE ORGANIZATION

- Annual Balance Sheet data is split into two files:
 - Bal Ann Indl (Industrial)
 - Bal Ann Fin (Financial Services)
- Income Statement Annual is split into two files:
 - IS Ann Indl (Industrial)
 - IS Ann Fin (Financial Services)
- New groups
 - Annual Data – group Index Annual is new. All items are prefixed with IDX_*
 - Quarterly Data – group Index Quarterly is new. All items are prefixed with IDX*_

REFERENCE FILES

Most Compustat reference codes have been consolidated into three files, and eight remained as stand-alone files.

- ReferenceCodes - Reference Codes
 - Accounting Standard
 - Acquisition Method
 - Bal Sheet Presentation
 - Company Status
 - Comparability Status
 - Country
 - Data Code
 - Data Format
 - Dividend Tax Marker
 - Exchange Rate Type
 - Footnote
 - IC Auditor Opinion
 - Income Statement Model
 - Industry Format
 - Industry Presentation
 - Issue Status Alert
 - Issue Type
 - Level of Consolidation
 - MIC Reference
 - Note Type
 - Officer Certification
 - Officer Title
 - Oil & Gas Method

- Reason for Deletion
- Sec Fin Ann Footnote
- State/Province
- Status Alert
- Stock Ownership
- ReferenceNumbers – Reference Numbers
 - Auditor
 - Auditor Opinion
 - Cash Flow Format
 - Inventory Valuation
 - Price Status
 - S&P Economic Sector
 - S&P Industry Sector
 - Source Document
 - Stock Exchange
 - Update
- ReferenceNumberswithStatus – Reference Numbers with Status
 - Fortune Industry
 - NAICS
 - S&P Industry Index
 - S&P Major Index
 - SIC
- ReferenceConstituents - Constituents
- ReferenceExchangeTier - Exchange Tier
- ReferenceFootnote - Footnote
- ReferenceGICS - GICS
- ReferenceIndex - Index
- ReferenceMajorIndexes - Major Index
- ReferenceMarketHoliday - Market Holidays
- ReferenceNoteSubtype - Note Subtype

FISCAL PERIOD DATE CONVENTIONS

Format of date items used to access fiscal data has changed. YYYY.Q, where Q identifies quarter 1-4, is now presented as and integer YYYYQ. This change allows for this to be a more efficient relational database key for joining tables.

STEPS TO INSTALL:

- From the directory containing your downloaded zip file, right click (in Windows) or run File Roller (in Red Hat)
- Set the extract destination to the current higher level directory where data is normally stored and unzip will add the rest of the path (default is C:\ in Windows or /home in Red Hat)
- For example, in Windows, extracting to C:\ will install the data under C:\crspdata

Contact CRSP Client Services for assistance at: support@crsp.org or call 312-263-6400, Option 2.

SETTING ENVIRONMENT VARIABLES

The zip process will not automatically set environment variables for use with command line tools. The following steps will do so. NOTE: CRSPSift does not use and therefore does not require environment variables.

LINUX

NOTE: The new zip installs are now saving to directories with lower case letters – for example, ~/crspdata/diz201308

Check the current state of the environment variables by entering this command: `env | grep crsp`

- If you have admin privileges on your machine, open the .profile file, and make updates to the following lines to update the paths for the new data installed:
 - `CRSP_CCM=<full path name>/<compustat directory>` (example: `CRSP_CCM=~/.crspdata/cmz201309`)
- If you are using the bash shell (or any Bourne-style shell), enter the same commands above and export each one. This will make changes for your current session only and does not save to the .profile file.
- A shell script is available to make changes that can be reused for each update done. Note that this will only make changes to your current session and does not save to the .profile file.
 - Navigate to where the CUPL tools are loaded and then enter: `cd accbin`

- Start the script by entering: `./crsp_setup.sh`
- The script will prompt for the following:
 - File name (*will be appended with .kshrc*)
 - Top level directory (this should be where CUPL tools are located (example: `~/crsp`)
 - Scratch directory for error messages – name can be anything
 - Enter path for daily stock directory
(*example: `~/crspdata/diz201308`*)
 - Enter path for monthly stock directory
(*example: `~/crspdata/miz201308`*)
 - Enter path for index directory
(*example: `~/crspdata/ifz201308`*)
 - Enter path for CCM directory – enter to leave blank
(*example: `CRSP_CCM=~/crspdata/cmz201309`*)
 - Enter path for CST directory – enter to leave blank
 - Execute script by running: `./<file name created above>.kshrc` (*note there is a space between the two dots*)

WINDOWS:

The CRSP Client Environment utility provides users a way to check the location of the software and to set environment variables. The utility can be started by locating it under installed programs CRSPAccess/CRSP Client Environment, or by navigating to the /accbin folder where the CUPL tools are installed and running: `Client_Environment.exe` . After making any changes, be sure to log off and back into Windows for the settings to be saved.

[Click here to view a copy of the CRSP/Compustat Merged Database Guide.](#)