

CRSP/COMPUSTAT RELEASE NOTES

MARCH 2020 MONTHLY 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 MARCH 2020 CRSP/COMPUSTAT MERGED DATABASE

The cut date for these release notes is 20200314.

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	19251231-20200313	1-24843
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	19260102-20200313	1-4915

DATABASE CHARACTERISTICS

The following table summarizes general characteristics of the CCM database:

ATTRIBUTE	CCM FORMAT DATABASE
SOURCE	COMPUSTAT XPRESSFEED
VOLUME LABELS	CMZ20200313 - CRSPACCESS CFZ20200313 - SAS/ASCII/R CLZ20200313 (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	50,268
TOTAL GVKEY COMPANIES	48,070
TOTAL GVKEYX INDEXES	2,198
TOTAL GVKEY/IID SECURITIES	62,105
TOTAL CRSP SECURITIES LINKED	30,068
TOTAL GVKEYS IN LINK	48,020
TOTAL LINK RECORDS	105,815
TOTAL LINK RECORDS WITH A LINK TO CRSP	34,890

DATABASE SIZES

CMZ	11.67 GB
CFZ - ASCII	10.58 GB
CFZ - SAS	29.09 GB
CFZ - R	1.41 GB
CLZ - ASCII	1.09 GB
CLZ - SAS	1.76 GB
CLZ - R	44.6 MB

MARCH 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
DERAC	DERIVATIVE ASSETS - CURRENT	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
DERALT	DERIVATIVE ASSETS LONG-TERM	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
DERLC	DERIVATIVE LIABILITIES- CURRENT	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
DERLLT	DERIVATIVE LIABILITIES LONG-TERM	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
GLASA_DC	GAIN/LOSS AFTER-TAX DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLA_DC	GAIN/LOSS AFTER-TAX DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLDQ_DC	GAIN/LOSS DILUTED EPS EFFECT DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLDSA_DC	GAIN/LOSS DILUTED EPS EFFECT DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLDY_DC	GAIN/LOSS DILUTED EPS EFFECT DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLD_DC	GAIN/LOSS DILUTED EPS EFFECT DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLPSA_DC	GAIN/LOSS PRETAX DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
GLP_DC	GAIN/LOSS PRETAX DC	7	SFAS	INDUSTRIAL FORMAT, PRE-FASB REPORTING, STANDARDIZED PRESENTATION
HBINVNO	INVENTORIES NOT OWNED/DEPOSITS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
OGI	IMPAIRMENT OF OIL & GAS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
OGIQ	IMPAIRMENT OF OIL & GAS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
OGIQ_DC	IMPAIRMENT OF OIL AND GAS DATA CODE	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
OGIY	IMPAIRMENT OF OIL & GAS	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION
OGIY_DC	IMPAIRMENT OF OIL AND GAS DATA CODE	5	PFO	INDUSTRIAL FORMAT, PRO FORMA REPORTING, STANDARDIZED PRESENTATION

MARCH 2020 LINK EDITS

GVKEY 9835 PIONEER ENERGY SERVICES CORP

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19810101	19810203	0	0	00X	NR	C
19810204	99999999	63335	4202	01	LC	C

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
19810101	19810203	0	0	00X	NR	C
19810204	20190813	63335	4202	1	LC	C
20190814	99999999	0	0	1	NR	C

GVKEY 18844 SHARING ECONOMY INTL INC - OLD

OLD:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20060101	20071129	0	0	00X	NR	C
20071130	20091228	0	0	01	NR	P
20091229	99999999	93191	53312	01	LC	P

NEW:

LINKDT	LINKENDDT	LPERMNO	LPERMCO	LIID	LINKTYPE	LINKPRIM
20060101	20071129	0	0	00X	NR	C
20071130	20091228	0	0	01	NR	P
20091229	20181204	93191	53312	01	LC	P
20181205	20191231	0	0	01	NR	P

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.](#)