GENERAL DATABASE INFORMATION

- SAS data sets are provided, which can be used with Excel, R, and SAS respectively with no conversion needed.
- The CRSP US Treasury Databases have been tested on Windows 10 and Linux.

For complete details and instructions, please refer to the CRSP Treasuries Guides on our website at www.crsp.org/documentation.

RECOMMENDED USE

Our recommendation is to use the format of the CRSP US Treasury Databases that requires the least amount of data conversion. Therefore if you plan to use use Excel, 'R', or SAS for analysis using the corresponding data set will be the most straight forward.

DELIVERY METHOD

Sift Professional and Enterprise are available to all CRSP subscribers for downloading on our MOVEit site, https://crsp.moveitcloud.com in the Utility_Downloads folder.

Sift Professional is supported on Windows 10. Sift Enterprise supports Servers 2008 R2 and 2012 R2. See details in the Sift Release Notes on our website at http://www.crsp.org/products/software-access-tools/crspsift

ZIP INSTALLS

CRSP uses a zip process for data installation. ZIP utilities that are standard on all platforms should work to unzip the databases. On Windows computers, CRSP recommends using 7-ZIP, which can be downloaded from this site: http://www.7-zip.org/

ZIP FILE STRUCTURE AND PROCESS

There are four separate zip files; each will have to be extracted separately. At the time the files are unzipped, make sure to set the directory path to the intended home directory (for example, in Windows C:\ , in Red Hat /home). The unzip process will extract to that directory under the subdirectory \crspdata.

Note that the destination directories have also changed as part of this process.

ZIP FILE ON DISK	DESCRIPTION	INSTALL PATH	FILE SIZE (MB)	
			ZIPPED	UNZIPPED
TRZYYYYMM_ASCII	CRSP ASCII FLAT FILES	\CRSPDATA\ TRZYYYYMM_ASCII	180.49	813.87
TRZYYYYMM_CADB	CRSP ACCESS FILES	\CRSPDATA\ TRZYYYYMM_CADB	135.46	305.26
TRZYYYYMM_EXCEL	CRSP EXCEL FILES	\CRSPDATA\ TRZYYYYMM_EXCEL	110.82	117.67
TRZYYYYMM_R	CRSP R FILES	\CRSPDATA\TRZYYYYMM_R	133.95	136.67
TRZYYYYMM_SAS	CRSP SAS FILES	\CRSPDATA\ TRZYYYYMM_SAS	163.55	497.87

The tables below list version specific information for the Daily and Monthly CRSP US Treasury Databases.

DECEMBER 2019 FILE VERSION SPECIFICS

	MONTHLY BOND	DAILY BOND
DATA RANGE	19251231-20191231	19610614-20191231
TOTAL ISSUES	7197	5349
MAXIMUM ACTIVE ISSUES	400	400
ACTIVE ISSUES THIS UPDATE	358	355

JUNE 2019 EDITS

TREASNO	CRSPID	DESC
205091	19470915.201250	CHANGED TOTAL DEBT FROM 687 TO 1687 FOR 19441230-19470331

SUPPORT

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

ANNOUNCING 'R' FORMAT DATASETS – AVAILABLE JANUARY 31, 2020

Starting with the December 2019 data release, the CRSP US Treasury Database will also be available in 'R' format datasets. The 'R' datasets will be in the zip file trz201912_r.zip in the same folder on MOVEit Cloud as the other formats of the data. The first release of these files will be January 31, 2020. In future months, the 'R' files will be released at the same time as the other Treasury formats.

MINOR CHANGES FOR INCREASED CONSISTENCY ACROSS FORMATS

The CRSPAccess/Sift formats remain unchanged.

Starting with the December 2019 data, released in January 2020, changes were made to improve the consistency across the ASCII, Excel, 'R', and SAS formats. The changes fall into three groups: missing values, numeric precision/display, and miscellaneous other changes.

- **Missing Values Changes** - 29 missing values conventions changes – See Appendix I in this document for details.
- **Numeric Precision/Display**
  - Increased Precision - 96 fields among 14 files have been increased from 13 significant digits scientific ("E") notation in ASCII to 14 significant digits to better match the available precision in the 8-byte (64-bit) floating point numeric fields used by SAS, Excel, and R - See Appendix II.A for details.

    Example:
    | 1.234567890123E-01 | to | 1.2345678901234E-01 |
- **Change to Fixed Format** – Three fields, one in each of three files have been changed from 13 significant digits scientific ("E") notation in an ASCII fixed notation to remove unnecessary trailing zeros. – See Appendix II.B for details.

  Example:
  | 4.750000000000E+00 | to | 4.750 |
- **Change T-Bill-only Duration Fields to Integer Format** – Fourteen fields among six files have been changed from 13 significant digits scientific ("E") notation in an ASCII integer format to remove unnecessary trailing zeros for duration fields that contain only T-Bills – See Appendix II.C for details.

  Example:
  | 4.5000000000E+01 | to | 45 |
- **Miscellaneous Other Changes**
  - **Remove Paid Interest Rows before 12/31/1925** –The minimum value of TPQDATE is 12/31/1925.

    The TFZ_PAY (paid interest file) previously had records of interest payments for a handful of issues with payment dates (TPQDATE) prior to 12/31/1925. These rows were not used for any calculations and have been deleted from the files.
  - **Minor Display Changes to Excel Files** – The mechanism used to create the Excel file was changed, and there are two changes to default display format.
    - The Header Row appearance has changed
    - The default display format for many numeric variables has changed (e.g. values previously displayed as 9.9663E-05 and 99.959167 are now displayed as 0.0001 and 99.9592, respectively), but the underlying values used for any Excel calculation remains unchanged.
  - **Adding a Header Row to ASCII Files** – The ASCII files now contain a header row with the column names.
  - **Sort order changes to TREASNO from CRSPID** - four files reflect this new sort order:
    - TFZ_DLY – Daily Time Series
    - TFZ_ISS – Issue Descriptions
    - TFZ_MTH – Monthly Time Series
    - TFZ_PAY – Paid Interest Event Series

APPENDIX I - MISSING VALUE DETAIL

- Six SAS and Excel code fields changed from SAS and Excel from missing to zero, where the zero is a documented code value:
  - TFZ_ISS file – IWHY, TNIPPY, IFCPDTF, TNOTICE, IUNIQ
  - TFZ_MTH_TS – RMTREASNO
- Seven numeric fields that had missing values in SAS and Excel, but were represented by zeros in ASCII have been changed to an empty string in ASCII in order to be consistent with SAS and Excel:

- ○ TFZ_ISS file – IYMCN
- ○ TFZ_DLY file – TDPUBOUT and TDTOTOUT
- ○ TFZ_MTH file – TMPUBOUT and TMTOTOUT
- ○ TFZ_IDX file - TTERMMIN and TTERMMAX
- Eight numeric fields in SAS, Excel, and ASCII changed to missing values (empty string) where they previously were coded as -99:
  - ○ TFZ_DLY file – TDRETNUA, TDYLD, TDDURATN
  - ○ TFZ_MTH file – TMRETNUA, TMYIELD, TMDURATN, TMPCYLD, TMRETNX
- Six numeric fields in SAS, Excel, and ASCII changed to missing values (empty string) where they previously were 0:
  - ○ TFZ_DLY file – TDBID, TDASK, TDNOMPRC
  - ○ TFZ_MTH file – TMBID, TMASK, TMNOMPRC

APPENDIX II.A – INCREASED PRECISION – ALL FORMATS

The following 96 fields in 14 files have been increased from 13 significant digits scientific ("E") notation in ASCII to 14 significant digits (e.g. +1.234567890123E+12 to +1.2345678901234E+12) to better match the available precision in the 8-byte (64-bit) floating point numeric fields used by SAS, Excel, and R.

1. TFZ_ISS –TVALFC
2. TFZ_DLY –TDBID, TDASK, TDNOMPRC, TDACCINT, TDRETNUA, TDYLD, TDDURATN, TDPDINT
3. TFZ_MTH –TMBID, TMASK, TDNOMPRC, TMACCINT, TMRETNUA, TMYLD, TMDURATN, TMPCYLD, TMRETNXS, TMPDINT
4. TFZ_PAY –PDINT
5. TFZ_MTH_BP –TMEWRETD
6. TFZ_MTH_FB –TMNOMPRC, TMYTM
7. TFZ_DLY_FT –TDYEARSTM, TDDURATN, TDRETADJ, TDYTM, TDBID, TDASK, TDNOMPRC, TDACCINT
8. TFZ_MTH_FT –TMYEARSTM, TMURATN, TMRETADJ, TMYTM, TMBID, TMASK, TMNOMPRC, TMACCINT
9. TFZ_MTH_RF –TMBIDYTM, TMASKYTM, TMYTM
10. TFZ_DLY_RF2 –TDBIDYLD, TDASKYLD, TDYLD
11. TFZ_MTH_RF2 –TMBIDYLD, TMASKYLD, TMYLD
12. TFZ_MTH_TS –TMBID, TMBIDRET, TMBIDYLD, TMBIDFWD, TMASK, TMASKRET, TMASKYLD, TMASKFWD, TMNOMPRC, TMAVERET, TMAVEYLD, TMAVEFWD
13. TFZ_DLY_TS2 –TDBID, TDASK, TDNOMPRC, TDBIDYLD, TDASKYLD, TDYLD, TDBIDFWD1, TDASKFWD1, TDAVEFWD1, TDBIDFWD4, TDASKFWD4, TDAVEFWD4, TDBIDHLD1, TDASKHLD1, TDAVEHLD1, TDBIDHLD4, TDASKHLD4, TDAVEHLD4
14. TFZ_MTH_TS2 –TMBID, TMASK, TMNOMPRC, TMBIDYLD, TMASKYLD, TMYLD, TMBIDFWD1, TMASKFWD1, TMAVEFWD1, TMBIDFWD4, TMASKFWD4, TMAVEFWD4, TMBIDHLD1, TMASKHLD1, TMAVEHLD1, TMBIDHLD4, TMASKHLD4, TMAVEHLD4

APPENDIX II.B – CHANGED TO FIXED NOTATION - ASCII

The following three fields, one in each of three files, have been changed from 13 significant digits scientific ("E") notation in an ASCII fixed notation to remove unnecessary trailing zeros:

Example:
| 4.750000000000E+00 | to | 4.750 |

- File TFZ_ISS –TCOUPRT – 7.3 format (e.g | 12.123 |)
- File TFZ_DLY_CD –TDRATE – 6.2 format (e.g | 12.12 |)
- File TFZ_MTH_CD –TMRATE – 6.2 format (e.g | 12.12 |)

APPENDIX II.C – CHANGED T-BILL-ONLY DURATION TO INTEGER FORMAT - ASCII

Fourteen fields among six files have been changed from 13 significant digits scientific ("E") notation in an ASCII integer format to remove unnecessary trailing zeros for duration fields that contain only T-Bills:

Example:
| 4.500000000000E+01 | to | 45 |

- File TFZ_MTH_RF – 1 Field – TMDURATN
- File TFZ_DLY_RF2 – 1 Field – TDDURATN
- File TFZ_MTH_RF2 – 1 Field – TMDURATN
- File TFZ_MTH_TS – 1 Field – TMDURATN
- File TFZ_DLY_TS2 – 5 Fields – TDDURATN, TDDURFWD1, TDDURFWD2, TDDURHDL1, TDDURHDL4
- File TFZ_MTH_TS2 – 5 Fields – TMDURATN, TMDURFWD1, TMDURFWD2, TMDURHDL1, TMDURHDL4