FRED-SD
Data Appendix V1.0

Data Appendix

This version of the appendix summarizes the availability of data and general cleaning procedures used to assemble the FRED-SD database. Sections proceed as follows:

I.   Availability of Raw Data
II.   Cleaning Raw Data
III. Assigning Vintage Dates
IV.   Missing Values
V.   Accessing Data

The FRED-SD database includes 28 variables for the 50 U.S. states and Washington D.C. Tables 1 and 2 include names and further description of these variables. All data is sourced from ALFRED (Archival FRED). Throughout this appendix, note that the date format mm-YYYY (ex. Aug-2019) refers to observation dates and YYYY-MM (ex. 2019-08) refers to vintage dates.

I.   Availability of Raw Data

a.   Substituted Series
The following series are substituted in the FRED-SD database because the desired series does not exist (in ALFRED).

i.   **MFGHRS: Manufacturing Hours** is substituted with Goods Producing Hours (which aggregates Manufacturing and Mining Hours) in the following series:
1.   DCMFGHRS: Manufacturing Hours for Washington D.C.
2.   MTMFGHRS: Manufacturing Hours for Montana
3.   NVMFGHRS: Manufacturing Hours for Nevada
4.   NMMFGHRS: Manufacturing Hours for New Mexico
5.   WYMFGHRS: Manufacturing Hours for Wyoming

ii. **CONS: Construction Employment** is substituted with Mining, Logging and Construction Employment in the following series:
1.   DECONS: Construction Employment for Delaware
2.   DCCONS: Construction Employment for Washington D.C.
3.   HICONS: Construction Employment for Hawaii

b.   Unavailable Series
The following series are omitted from the FRED-SD database because they do not exist (in ALFRED) and lack adequate substitutes.

i.   **RENTS: Real Estate, Rental and Leasing Employment**
1.   NMRENTS: Real Estate, Rental and Leasing Employment for New Mexico
2.   RIRENTS: Real Estate, Rental and Leasing Employment for Rhode Island
3.   SDRENTS: Real Estate, Rental and Leasing Employment for South Dakota

ii. **MINNG: Mining and Logging Employment**
1.   DEMINNG: Mining and Logging Employment for Delaware
2.   DCMINNG: Mining and Logging Employment for Washington D.C.

3. HIMINNG: Mining and Logging Employment for Hawaii

c. **Omitted Series**

The following series are omitted from the FRED-SD database because they are currently discontinued and contain only a small number of observations and vintages:

i. DCCONSTNQGSP: Construction GSP for District of Columbia
ii. DCMANNQGSP: Manufacturing GSP for District of Columbia
iii. DENATURNQGSP: Agriculture and Mining GSP for Delaware
iv. RINATURNQGSP: Agriculture and Mining GSP for Rhode Island
v. RICONSTNQGSP: Construction GSP for Rhode Island
vi. WYAGRNQGSP: Agriculture, Forestry, Fishing and Hunting GSP for Wyoming [i]
vii. WYMANNQGSP: Manufacturing GSP for Wyoming

II. **Cleaning Raw Data**

a. **Adjusting Units**
   i. **OTOT: Personal Income**
   Vintages from 2013-09 to 2018-09 are published in thousands of dollars for all fifty states and DC. These vintages are adjusted reflect to millions of dollars for comparison to previous and future vintages in ALFRED.
   ii. **LF: Labor Force**
   Vintages from 2007-06 to 2016-01 are published in thousands of persons for all fifty states and DC. These vintages are adjusted to reflect number of persons for comparison to current and future vintages in ALFRED.**Aggregating Series**

The following series aggregate related GSP categories from ALFRED and create new GSP categories to reduce the number of variables in the FRED-SD database.

   i. **NATURNQGSP: Agriculture and Mining GSP**
      1. AGRNQGSP: Agriculture, Forestry, Fishing, and Hunting GSP
      2. MINNQGSP: Mining GSP
   ii. **FIRENQGSP: Finance and Real Estate GSP**
      1. FININSNQGSP: Finance and Insurance GSP
      2. RERENTLEANQGSP: Real Estate, Rental and Leasing GSP
   iii. **PSERVNQGSP: Other Private Industries GSP**
      1. WHOLENQGSP: Wholesale Trade GSP
      2. RETAILNQGSP: Retail Trade GSP
      3. TRANSWARENQGSP: Transportation and Warehousing GSP
      4. PROBUSNQGSP: Professional, Scientific, and Technical Services GSP
      5. MNGCOENTPRNQGSP: Management of Companies and Enterprises GSP
      6. ADMINWASTNQGSP: Administrative and Waste Management Services GSP
      7. EDCATNQGSP: Educational Services GSP
      8. HLTHSOCASSNQGSP: Health Care and Social Assistance GSP
      9. ARTENTRECNQGSP: Arts, Entertainment, and Recreation GSP
      10. ACCOMDNQGSP: Accommodation and Food Services GSP
      11. OTHSERVENQGSP: Other Services, Except Government GSP

For each aggregated series, all the subcomponent series were available for the same observation period with the same number of vintages, with the exception of one aggregated series:

    i. **DCNATURNQGSP: Agriculture and Mining GSP for District of Columbia**
    Mining and Agriculture data had slightly different observation and vintage dates. The combined series here represents only common observations. This resulted in dropping 4 observations from DCAGRNQGSP (Jan-2005 to Oct-2005) and one vintage from DCMINNQGSP (2016-08). [ii]

c. **Aggregating Frequency**
    i. **ICLAIMS**: **Initial Claims**
    Vintages are aggregated from weekly to monthly frequency. The monthly vintage reflects the average value of weeks observed in a month using the last weekly vintage in a given month.

d. **Seasonal Adjustments**
The following series were not available with seasonal adjustment (in ALFRED) but have been manually adjusted in the FRED-SD database:

    i. Applied to all 50 states and DC:
        1. MFGHRS: Manufacturing Hours
        2. STHPI: Home Price Index
        3. ICLAIMS: Initial Claims
    ii. Applied to select states:
        1. SDCONS: Construction Employment for South Dakota
        2. TNCONS: Construction Employment for Tennessee
        3. WVINFO: Information Employment for West Virginia
        4. SDMINNG: Mining and Logging Employment for South Dakota
        5. TNMINNG: Mining and Logging Employment for Tennessee
        6. DCBPPRIVSA: New Private Housing Units for District of Columbia

III. **Assigning Vintage Dates**

a. **Monthly Data**
Generally, a new observation is added each month for a series, and the vintage date is forced to the last date of the respective month.

    i. **Delays in Data due to Release Schedule**
    The State Employment and Unemployment news release presents data from the Local Area Unemployment Statistics and State and Metro Area Employment, Hours, & Earnings programs. The *State Employment and Unemployment Report* typically publish data on a one month lag; however, data for January is not released in February. As a result, January and February observations are both released in March.
    The FRED-SD database preserves the real-time availability of these employment series: no new observation in February and two new observations in March. In some cases, the second data release was not captured in ALFRED, so a pseudo vintage was created to reflect the real-time availability of these series.
    Link explaining releases: https://www.bls.gov/bls/news-release/laus.htm#2014

    ii. **Delays in Data due to Government Shutdown**
The partial government shutdown from December 22, 2018 to January 25, 2019 delayed the release of some data series. FRED-SD reflects the real-time availability of these series. As a result, some vintages will not add a new observation or will increase by more than one observation. For example:

       1. BPPRIVSA: New Private Housing Units
          a. 2018-12: zero new observations
          b. 2019-02: zero new observations
          c. 2019-03: three new observations

Census Release Schedule can be found here:
https://www.census.gov/construction/bps/schedule.html

b. **Quarterly Data**
FRED-SD assigns quarterly series vintage dates to the middle month of the quarter, with the exception of Personal Income series which is assigned to the last month of the quarter. By using this method, quarterly data in FRED-SD is consistent timing between observations but may reflect data released *later* than the real-time availability in some instances.

    iii. For instance, most quarterly series in FRED-SD are released at the beginning of the second month in the quarter. If the 1st falls on a weekend, the data is sometimes released at end of the first month. However, FRED-SD would not recognize the observation until the end of the second month in the quarter.

IV. **Missing Values**
The following series have at least one missing observation in FRED/ALFRED

a. **Filling in Missing Values**
    i. **Using Data from Another Vintage**
       1. When data is missing, the first procedure is to fill in the observation using the previous available vintage.
       2. When a previous vintage is not available (i.e. a missing observation occurs in the first vintage), the second procedure is to backfill the observation from the next available vintage.

    ii. **Using Data from Original Source**
       1. In some rare cases, an observation was not recorded in FRED for any vintage, in which case, the third procedure is to fill in the observation using the most recent value from the original data source. For example:
          a. DCBPPRIVSA: New Private Housing Units for District of Columbia
Observation for Sep-2006 was not available via FRED. Most recent Census data used to fill in missing value across all available vintages.
Data sourced from: https://www.census.gov/construction/bps/txt/tb2u200609.txt.

      b. **Missing Values Due to Series Discontinuation**

      The following series are intentionally left as missing values from FRED-SD because the series was discontinued at some point in time:

            i. MENATURNQGSP: Agriculture and Mining GSP for Maine

            This series was temporarily discontinued in 2008, 2009, and 2012.

      c. **Omitting Values Due to Inconsistent Report**

      In some cases, either vintages or early observations of a series were omitted from the FRED-SD database because of inconsistent reporting between ALRED vintages or the original source. For example:

            i. IMPTOT: Import of Goods

                1. Although the first vintage in FRED is 2019-03, vintages begin at 2019-04 in FRED-SD.

            iv. CTICLAIMS: Initial Claims for Connecticut

                1. Although the first observation in FRED is 10/1/1985, the series begins in 2/1/1986 in FRED-SD due to missing weekly observations.

V. Accessing Data

      a. Downloading FRED-SD Database

            i. Data can be downloaded from the following link: _____ .

      b. How to Read the FRED-SD Database

      The database is structured after FRED-MD, where

            i. Each file name corresponds to the vintage date.

            ii. Each tab in the file corresponds to a variable.

            iii. Each column corresponds to a U.S. state or Washington D.C.

            iv. Each row corresponds to an observation date

Endnotes

---

[i] The AGRNQGSP series does not appear directly in the FRED-SD database but is aggregated with MINNQGSP to create the NATURNQGSP series. In this case, WYNATURNQGSP represents only WYMINNQGSP.

[ii] The AGRNQGSP series does not appear directly in the FRED-SD database but is aggregated with MINNQGSP to create the NATURNQGSP series. In this case, only the DCAGRNQGSP series was adjusted and later merged with DCMINNQGSP.