



## Data Revisions with FRED®

[Diego Mendez-Carbajo, PhD](#), Federal Reserve Bank of St. Louis

### GLOSSARY

**Methodology:** Any formal procedure used to gather and measure information to produce data.

**Tradeoff:** An exchange or a compromise between two different outcomes or features.

**Vintage:** A version of a data series available at a particular moment in time.

### Compelling Question

How do revisions and updates improve data?

### Description

FRED® provides access to current data from more than 100 sources. Some of those sources revise and update their data, and ALFRED® stores all previous versions. This article describes to new data users why data are revised and updated and can serve as a reference to advanced data users.

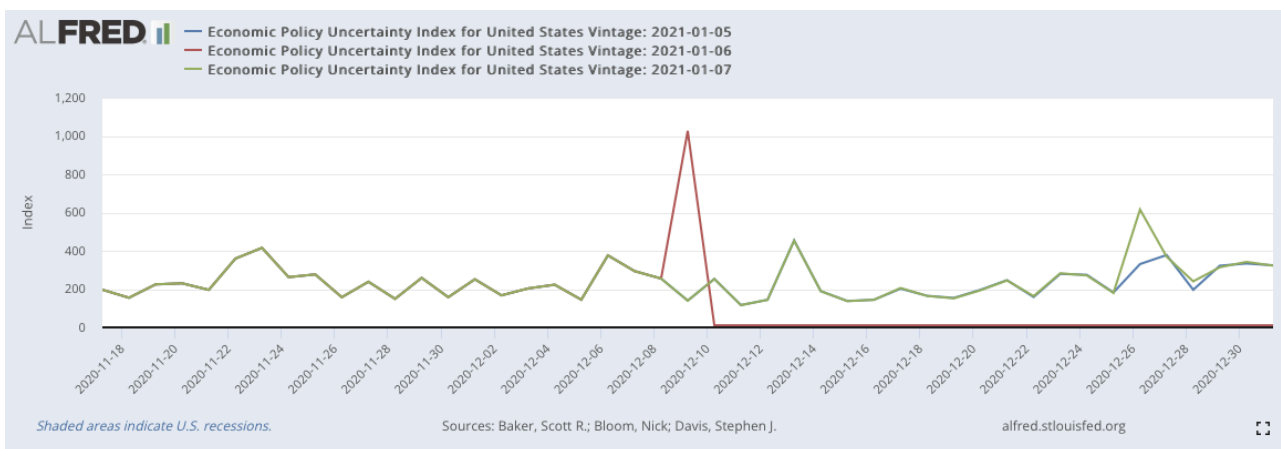
### Introduction

Sources revise data for several different reasons and at multiple moments in time. Occasionally, some data contain errors. Regularly, some data are revised to provide a more complete picture of current or very recent economic conditions. Periodically, some data are updated to produce a more accurate account of long-term trends and patterns.

Every time a new version, or **vintage**, of a data series is released, FRED displays the latest version, and the replaced version is archived in ALFRED. Economic researchers and consumers of data must reference the data vintage they use by citing the date they access the data.<sup>1</sup> Otherwise, discrepancies between data vintages can undermine the credibility of their work.

This article describes how headline economic data are revised and updated to provide a more complete and accurate description of economic conditions. High-quality data facilitate good decisionmaking, so the time and energy the data providers spend to improve their data benefits everybody who uses those data.

Figure 1  
Economic Policy Uncertainty Index for the United States



SOURCE: Baker, Scott R., Bloom, Nick and Davis, Stephen J. via FRED®, Federal Reserve Bank of St. Louis; <https://alfred.stlouisfed.org/graph/?g=O10h>, accessed April 7, 2022.

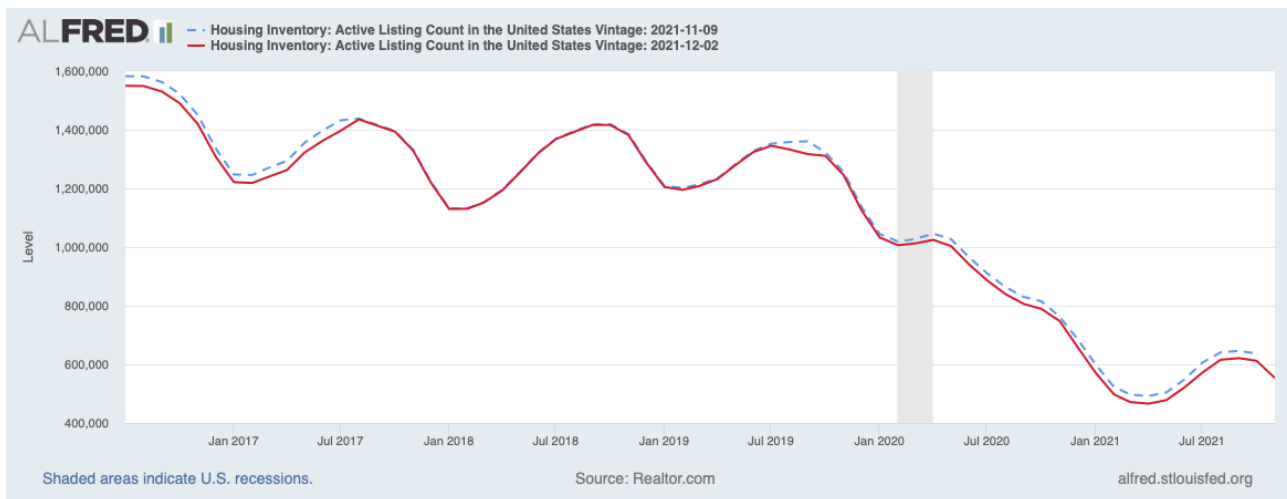
### Revising Data to Correct Errors

The process of collecting information to produce economic statistics is complex, and, despite quality-assurance checks along the way, sometimes erroneous data are released to the public. Not unlike proofreading a text to identify typos, plotting the data in a graph can help identify some of those errors.

Figure 1 shows three consecutive vintages of the same daily data series: The blue line shows the data released on January 5, 2021; the red line shows the data released the following day; and the green line shows the data released the day after that. Every day of the week a new data point is added, so each new line is longer than the previous one.<sup>2</sup> Moreover, with the addition of each new observation, the data provider revised previously released data after new information came to light.

For much of the graph, all three lines match up, indicating there were no discrepancies between the vintages. There are some small differences in the data values for the last week of December 2020 between the January 5 (blue) and January 7 (green) data releases. However, the big red spike in the graph reflects the largest discrepancy in data values: The data for December 9, 2020, that were released on January 6, 2021, indicated a 300% increase in value over the vintage released on January 5, 2021. And all the data points in the January 6 release after December 9 had a constant value (almost undistinguishable from zero in the graph), which is also a stark discrepancy from the values in the releases both the day before and day after. Whether this was the result of a coding error or some other reason, the next day's data vintage, on January 7, corrected these erroneously reported values.

Figure 2  
**Housing Inventory: Active Listing Count in the United States**



SOURCE: Realtor.com via FRED®, Federal Reserve Bank of St. Louis; <https://alfred.stlouisfed.org/graph/?g=O119>, accessed April 7, 2022.

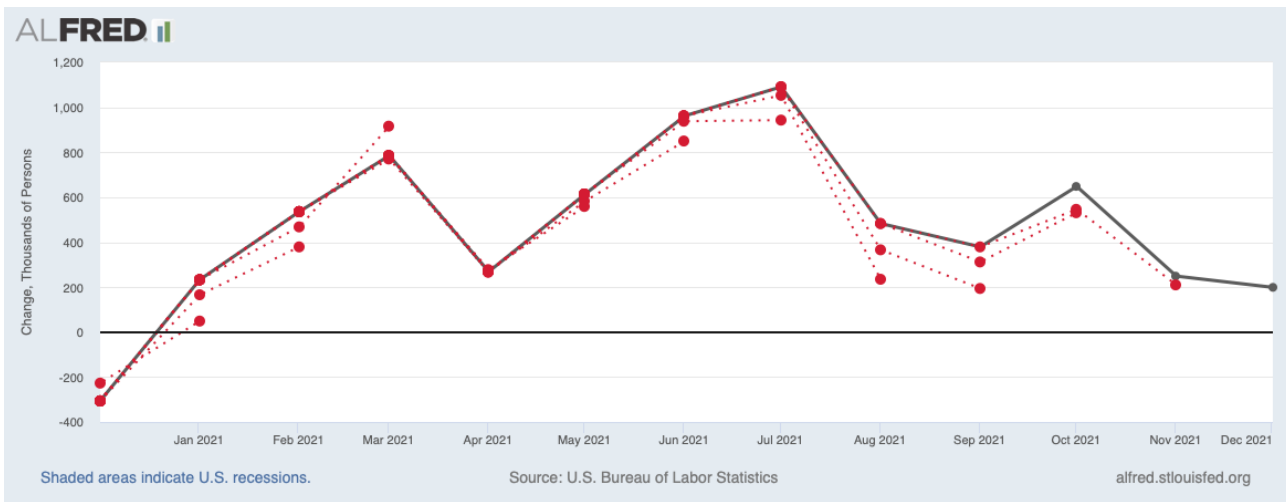
### Updating the Data Collection Methodology

The process of collecting information to produce economic statistics evolves and improves over time. Changes in **methodology** frequently result in revised data.<sup>3</sup> Not unlike comparing the text between two editions of a book, plotting the data produced using two different methodologies can help identify the scope and impact of the changes involved.<sup>4</sup>

Figure 2 shows two different vintages of the same monthly data series: The dashed blue line shows the data released on November 9, 2021, and the solid red line shows the data released on December 2, 2021—after a revision in the methodology. The gaps between the blue and red lines in the graph represent the differences in the reported statistic resulting from the revised methodology. You can observe relatively small differences in data values between vintages during some years and almost none during others. Overall, the impact of the change in methodology is most noticeable for the data reported since 2020.

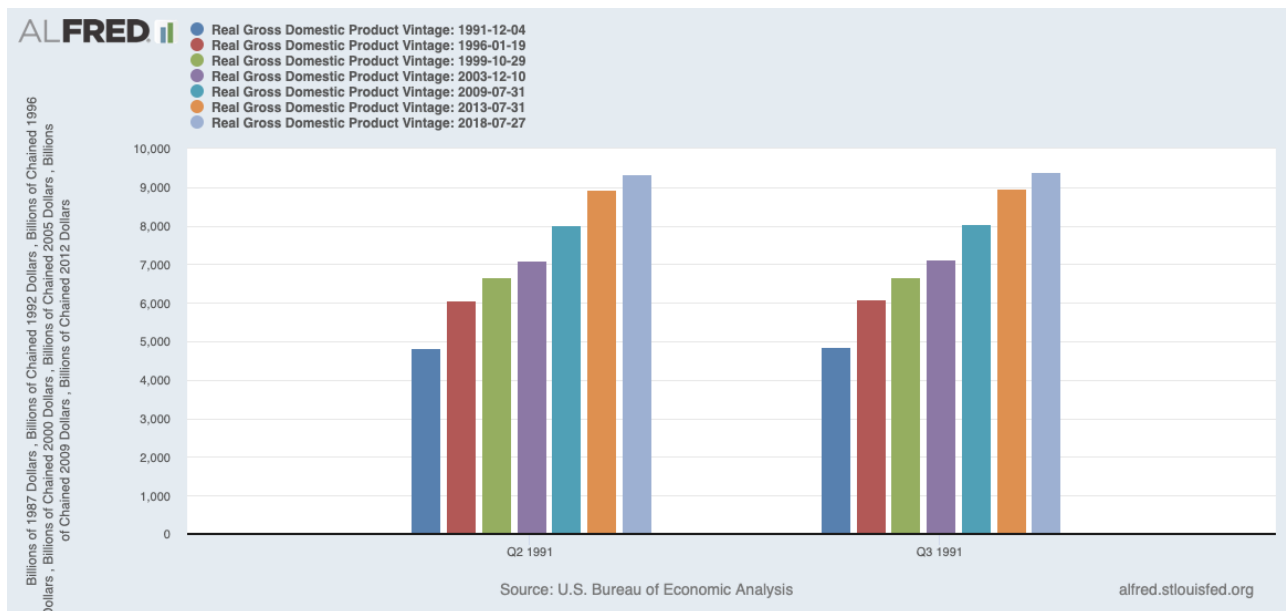
Besides comparing data vintages, you can sometimes identify changes in data collection methodologies through a one-time large variation in the value of the data.<sup>5</sup> In those cases, the historical record of the data themselves isn't revised, so the analysis of data values before and after the change in methodology must account for it.

Figure 3  
Change in All Employees: Total Nonfarm Payrolls



SOURCE: U.S. Bureau of Labor Statistics via FRED®, Federal Reserve Bank of St. Louis; <https://alfred.stlouisfed.org/graph/?g=KCs9>, accessed April 7, 2022.

Figure 4  
Real Gross Domestic Product



SOURCE: U.S. Bureau of Economic Analysis via FRED®, Federal Reserve Bank of St. Louis; <https://alfred.stlouisfed.org/graph/?g=PKSP>, accessed April 18, 2022.

### Improving the Accuracy and Completeness of the Data

The process of collecting information to produce economic statistics is time consuming. Because of that, there is frequently a **tradeoff** between the timeliness and the accuracy of information provided by the data. Not unlike a copyedited and revised draft of a manuscript, the final version of data provides more complete and precise information about current or very recent economic conditions.

Figure 3 shows 12 different vintages of the same monthly data series: The solid black line shows, as of January 7, 2022, the monthly changes in total nonfarm employment for each month in 2021. For each month in the graph, the three red dots show the initially announced value for the monthly change in total nonfarm employment, the first revision, and the second and final revision. The distance between the red dots and the black line shows the magnitude of the data revisions. In some cases the employment figures were revised up, and in other cases they were revised down.

Monthly revisions to employment data are a matter of common practice.<sup>6</sup> So are quarterly revisions to gross domestic product (GDP) and annual revisions to the consumer price index (CPI). These revisions reflect additional information collected from consumers, businesses, and government agencies since the initial (or preliminary) data were released.

Finally, headline economic indicators are periodically updated to reflect current economic practices and patterns. For example, the relative weights of goods and services purchased across the different categories of spending included in the CPI are updated every two years. At those times, the Bureau of Labor Statistics recalculates the CPI over the previous five years.<sup>7</sup>

Figure 4 shows seven different vintages of real GDP for the second and third quarter of 1991. The vintage dates included in the series names are the dates when the Bureau of Economic Analysis released a comprehensive update to the standards and methods it uses to measure economic activity. On those dates, the source also updated its GDP statistics by moving forward the reference year for its inflation adjustment and price measures.

The legend on the left axis of the graph lists the reference years for the seven comprehensive updates to the real GDP series available in FRED: 1987, 1992, 1996, 2000, 2005, 2009, and 2012. As the general price level rises over time, each successive bar is taller than the previous one. However, absent substantial changes to the types of economic activity being measured, the compounded annual rates of growth of real GDP are very similar across vintages.<sup>8</sup>

### Summary

More accurate data facilitate better decisionmaking. When data sources employ time and resources to improve their data, everybody who uses those data benefits. Because there can be significant differences in data values across series vintages, it is important to provide a complete citation, including the date you accessed the series, when using data. ■

## Additional Resources

Bureau of Economic Analysis. "Revising Economic Indicators: Here's Why the Numbers Can Change." *BEA Wire* (blog), 2018; <https://www.bea.gov/news/blog/2013-07-08/revising-economic-indicators-heres-why-numbers-can-change>.

Croushore, Dean and Stark, Tom. "Does Data Vintage Matter for Forecasting?" Federal Reserve Bank of Philadelphia Working Paper No. 99-15, October 1999; <https://www.philadelphiafed.org/-/media/frbp/assets/working-papers/1999/wp99-15.pdf>.

Jordà, Òscar; Kouckekinia, Noah; Merrill, Colton and Sekhposyan, Tatevik. "The Fog of Numbers." Federal Reserve Bank of San Francisco *Economic Letter*, July 2020; <https://www.frbsf.org/economic-research/publications/economic-letter/2020/july/fog-of-numbers-gdp-revisions/>.

Nardone, Thomas; Robertson, Kenneth and Maxfield, Julie H. "Why Are There Revisions to the Jobs Numbers?" U.S. Bureau of Labor Statistics *Beyond the Numbers: Employment & Unemployment*, 2013, 2(17); <https://www.bls.gov/opub/btn/volume-2/revisions-to-jobs-numbers.htm>.

## Notes

<sup>1</sup> For more on how to cite data in FRED, see the following article: Mendez-Carbajo, Diego. "Data Citations with FRED." Federal Reserve Bank of St. Louis *Page One Economics Data Primer*, October 2020; <https://research.stlouisfed.org/publications/page1-econ/2020/10/21/data-citations-with-fred>.

<sup>2</sup> For a discussion of an error in the Economic Policy Uncertainty Index for the United States data, see <https://fredblog.stlouisfed.org/2021/01/a-friendly-warning-data-arent-perfect/>.

<sup>3</sup> For a discussion of a change in the methodology used to calculate the number of active real estate listings, see <https://fredblog.stlouisfed.org/2021/12/a-change-in-measuring-active-real-estate-listings/>.

<sup>4</sup> For a discussion of a change in the methodology used to calculate the St. Louis Fed's Financial Stress Index, see <https://fredblog.stlouisfed.org/2022/01/the-st-louis-feds-financial-stress-index-version-3-0/>.

<sup>5</sup> For a discussion of a change in the methodology used to calculate the monetary aggregate M1, see <https://fredblog.stlouisfed.org/2021/05/savings-are-now-more-liquid-and-part-of-m1-money/>.

<sup>6</sup> For a discussion of revisions to employment data during 2020 and 2021, see <https://fredblog.stlouisfed.org/2022/01/revisions-to-employment-data-during-2020-and-2021/>.

<sup>7</sup> For a discussion of revisions and updates to the consumer price index, see <https://fredblog.stlouisfed.org/2022/04/revisions-and-updates-to-cpi-data/>.

<sup>8</sup> For a discussion of comprehensive updates to gross domestic product, see <https://fredblog.stlouisfed.org/2022/04/comprehensive-updates-to-real-gdp/>.

Please visit our website and archives at <http://research.stlouisfed.org/publications/page1-econ/> for more information and resources.

© 2022, Federal Reserve Bank of St. Louis. The views expressed are those of the author(s) and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis or the Federal Reserve System.