Log of Suspected Issues under Investigation in the GSS Cumulative File, Panel Datafiles, and Data Explorer

Curation of GSS survey data is an ongoing activity, and we rely on the user community to help us report possible issues with the data. Users are encouraged to report potential data issues to gss@norc.org. Making the data available to the community in a timely manner is a foundational principle for our work. For example, the GSS team holds no data embargos.

The GSS team corrects suspected errors as soon as possible but we do not have the resources to send out a notification each time one is confirmed and corrected. Therefore, we plan to release updated data about twice a year (Spring and Summer). However, if you have a question about a specific question and you believe you may have discovered an error or want to check in on one you notified us about, please get in touch with us directly so we can discuss it.

Current known errors that are scheduled to be fixed in the cumulative datafile, panel datafiles, and single-year datafiles are listed below, along with notes about anticipated corrections for the data.

This list will be constantly updated. Once the issues have been addressed and updated data has been released, the item will be removed from this list, and corrections will be included in release notes. Details on the actual updates will be provided in the GSS data documentation.

**CONINC, CONRINC, REALINC, REALRINC**

**Years affected:** Multiple years

**Datasets affected:** All data files


The values assigned to the top-coded categories may not be consistent for certain of these variables for certain years. This affects 1994, 2010, 2018, 2021, and 2022.

The GSS team is actively working to correct these issues.

**Expected release:** Spring 2024

**Exports using the GSSDE-generated .dat file (I)**

**Years affected:** All years

**Datasets affected:** GSS Data Explorer

GSS Data Explorer utilizes negative values to denote missing values (e.g., don't know, inapplicable), but many statistical software packages (e.g., SAS, Stata, SPSS) do not recognize these values as missing. Users may need to recode these negative values to missing, according to their chosen package's specifications, to exclude these cases from analysis.

In addition, the associated syntax files generated for Stata try to assign decimal values a label, even when the value is just a number. Stata does not allow decimal values to be assigned labels and the associated syntax will produce an error.

The GSS team is actively working to find a fix for these issues.

**Expected release:** Winter 2024