
Assessment of Manual vs Automated Survey Editing and Imputation

Sean Rhodes, Megan Lipke, Jennifer Maiwurm, Darcy Miller, Denise A. Abreu, Linda J. Young (U.S. Department of Agriculture National Agricultural Statistics Service, United States of America)

sean.rhodes@usda.gov

Abstract

The United States (U.S.) Department of Agriculture's National Agricultural Statistics Service (NASS) conducts hundreds of surveys each year to estimate nearly every facet of agriculture in the U.S. One of these surveys is the Crops Acreage, Production, and Stocks (APS) Survey. Crops APS is conducted quarterly in the months of June, September, December, and March. The September Crops APS collects data on acres planted, acres harvested, yield or production for small grains, and grain stored. Once data have been collected, responses are then reviewed and edited by NASS's regional field offices for outliers, edit failures, and item imputation. The current review process and item-level imputations are done through a manual, interactive process, which can be time consuming and costly for the Agency. NASS has taken steps to develop a generalized system called IDEAL (Imputation, Deterministic Edits, Automation and Logic) to automate the imputation and error correction occurring after data collection but before the outlier review process. This automated imputation and error correction was applied to historical data to evaluate the imputation methodology, efficiency, and data quality. This paper presents the results from testing on historical data which compares the data quality and efficiency of the manual process to the data quality and efficiency of the automated process using the historical data. These testing results are critical to implementing the new automatic editing system in the production environment for September Crops APS 2024.