

Updates on Editing and Imputation at the National Agricultural Statistics Service

Darcy Miller and Linda J. Young, PhD

September 2015



“ . . . providing timely, accurate, and useful statistics in service to U.S. agriculture.”



National Agricultural Statistics Service (NASS)

- “The National Agricultural Statistics Service provides timely, accurate, and useful statistics in service to U.S. Agriculture.”
- Over 400 Survey Reports Annually
- Census of Agriculture every 5 years

Updating Editing and Imputation Programs at NASS

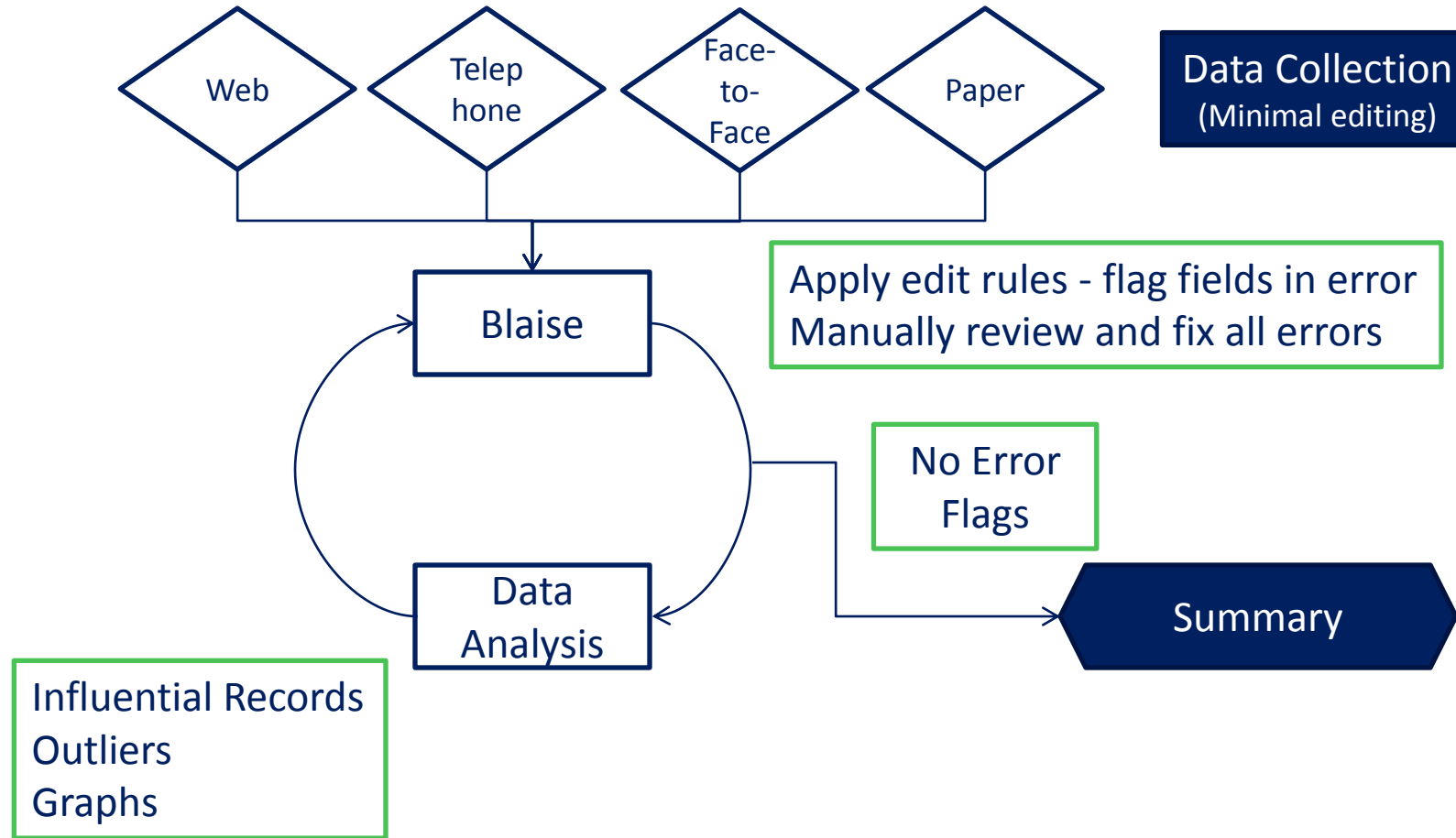
- Significance Editing (SignEdit)
- Iterative Sequential Regression (ISR)
- IVEware

SIGNEDIT (BANFF + SELECTIVE EDIT)

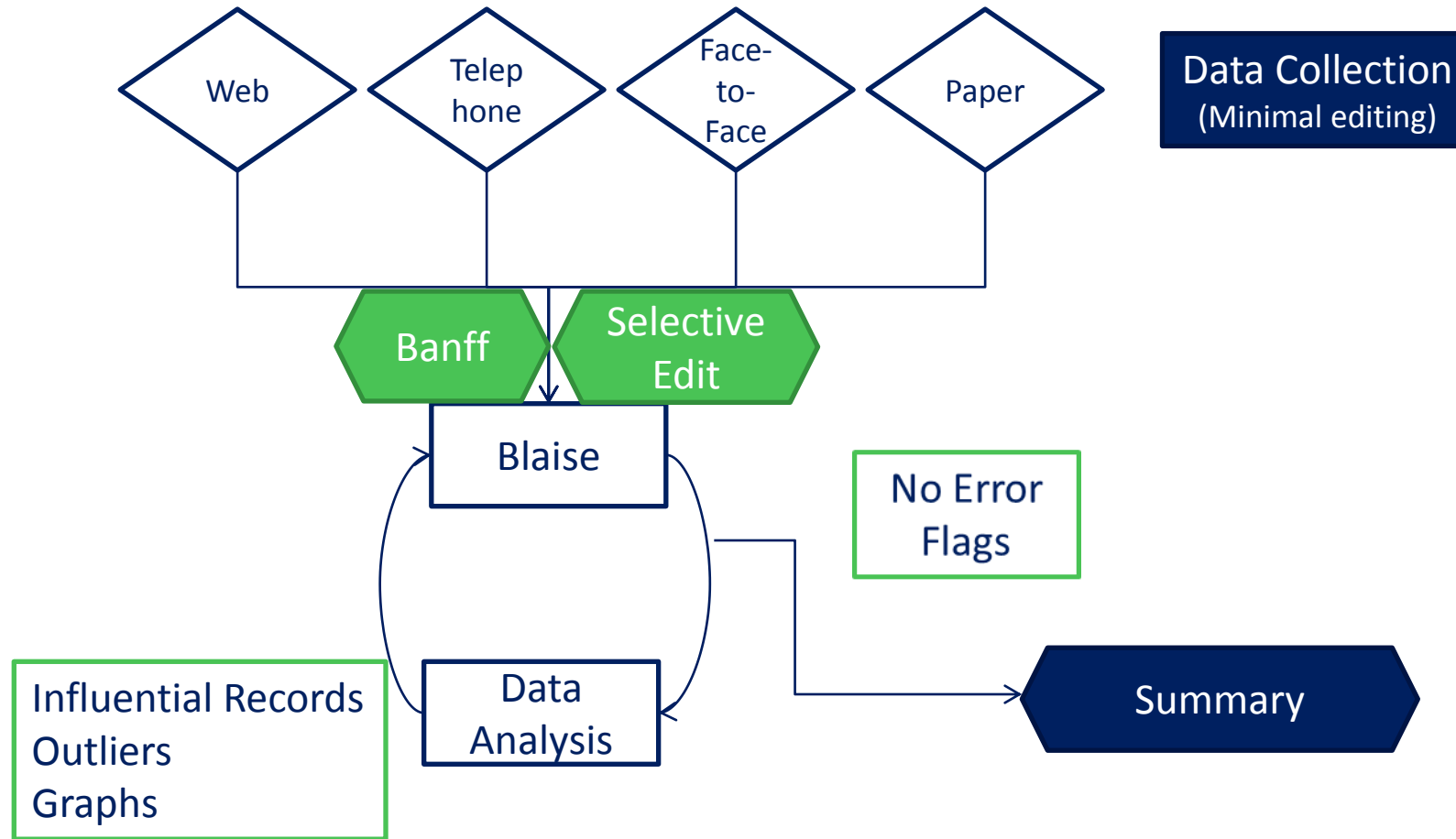


Hog Survey Processing

Current



Hog Survey Processing SignEdit (December 2015)



Selective Edit

- A way to view automated changes after Banff
- An item-level score is the percentage of impact from Banff change to final total of each commodity
- The unit score of a record is the maximum of all item scores
- Analysts use a custom viewer to review top “most impactful” changes based on unit score and make changes if necessary

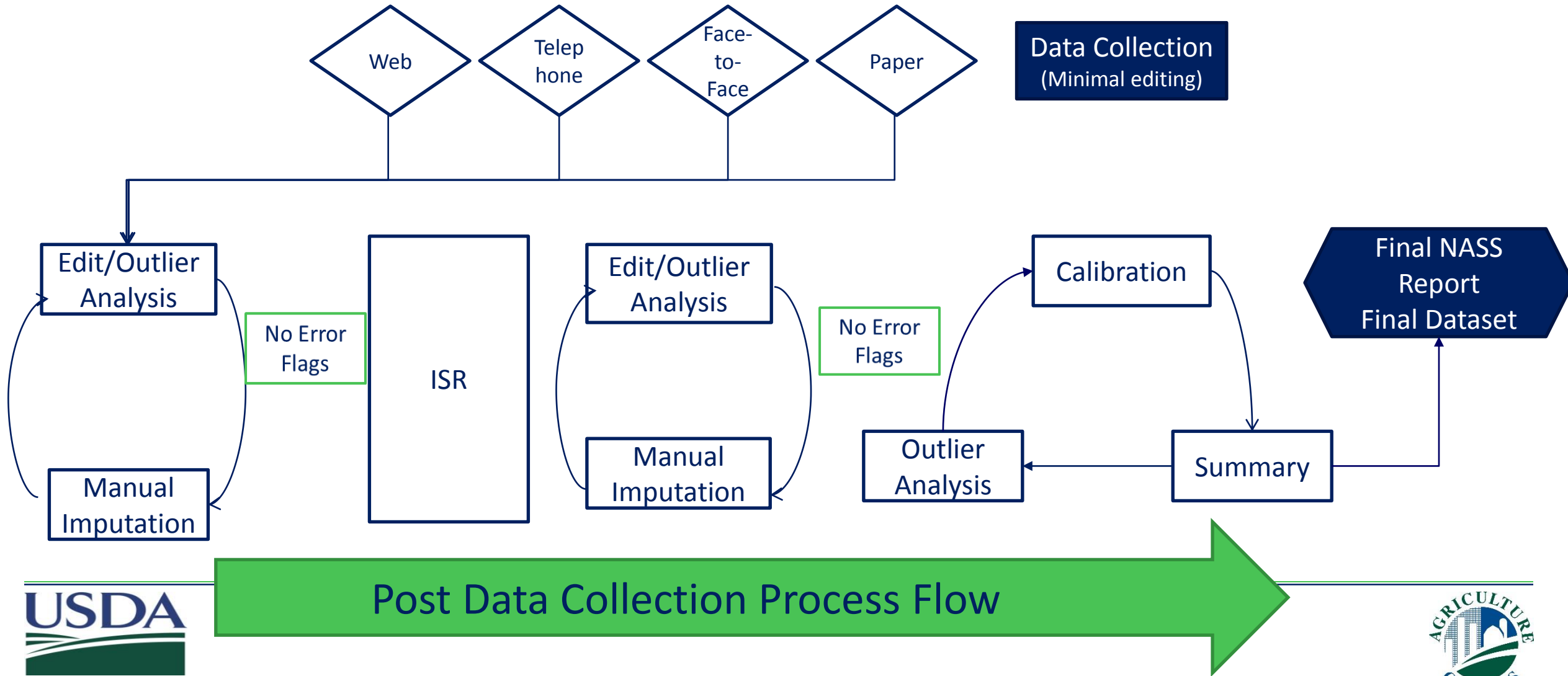
ITERATIVE SEQUENTIAL REGRESSION (ISR)



ISR

- Replaced conditional mean methodology to better maintain relationships for items of interest
- Implemented for 2014 survey year
- SAS Interface with in-house program written in R
- Blend of data augmentation (DA) and fully conditionally specified (FCS) models, having the covariate choice flexibility of the FCS methods but the theoretical background of the DA methods (Robbins, M. et al, JASA, 2013)

ARMS III Processing



IVEWARE



IVEware

- University of Michigan program implementing FCS methodology, Sequential Regression Multiple Imputation (SRMI), described by Ragunathan (2001)
- Similar to ISR (ARMS III)
- SAS Implementation
 - Free
 - Easy implementation
 - Options for a variety of data types, step-wise regression covariate selection, bounded imputations, restricted imputation, and more

Conclusion

- NASS is implementing statistical editing and imputation methods such as ISR, IVEware, and SignEdit
- Cultural shift from manual handling every record to modelling distributions
- Extend evaluation and updates to other survey programs