

README: Dataset_Creation

Purpose: Derive analytic data files for NHA project analyses

There are 5 steps required for dataset creation:

1. Derive_Survey_and_Claim (Chapters 1-3): extract and process claim and survey data by source and population
2. Within_Impute (Chapter 4): impute for missing values in survey variables by source and population
3. Calibration (Chapters 7-8): calibration of health condition variables by population
4. HMO_Weight_Adjustment (Chapter 5): create weight adjustment for HMO participation, incorporate NHEA adjustment by population
5. Final_Datasets: merge populations and finalize to create final analytic datasets

Within each there may be code for multiple sources and populations where relevant. README files provide further details of order of programs to be run within each data step.