Steps in data cleaning and analysis

SAS program:

"T:\USER\Zeping\Nonresponse in ATUS\NonresponseATUS\_update2018.sas"

1. Input
   1. Check duplicates
2. Linking datasets
   1. Link CPS
      1. Check duplicates
      2. Preprocess linking variables
      3. Merging CPS with ATUS
   2. Checking missing values
3. Preprocessing
   1. Making format labels
   2. Reconstruct interview outcome variables
   3. Make two variables about presence of children
   4. Make spouse work hours variable
   5. Make relate/non-relate household member variables
   6. make stratification variable of the presence and age of children and the number of adults in adults-only households
   7. Categorize variables and get rid of the ‘NIU’s
   8. Merge ATUS base weights
   9. Check data availability, especially for non-respondents
4. Analysis
   1. Overall trends
      1. Response, contact, cooperation and types of nonresponse
      2. Z-tests/risk differences for periods
   2. Subgroup cross-tabulation
      1. Cross-tabulation interview outcome by explanatory variables
      2. Line charts – by each variables
      3. Cross-tabulation in 2004 for comparing with Abraham 2006
   3. Subgroup difference in two periods comparison 2004-2008 and 2008-2018
      1. 2004-2008
      2. 2008-2018
   4. Logistic regressions for sub-period subgroup analysis
      1. Weighted logistic regression: outcome = factor + year
      2. LR test: the full model with interaction term and the reduced model without the interaction term
      3. Multiple logistic regression: 2004, 2008 and 2016
      4. Case: R’s hours worked and spouse hours worked
      5. Case: Central city vs. Nonmetropolitan
   5. Decision tree

~~Logistic regressions~~

~~Logistic: outcome = year by subgroups~~

~~Logistic: outcome = year\*factors (interaction term only)~~

~~Logistic: outcome = year + factors (without interaction term)~~

~~Logistic: outcome = single factor by year~~

~~Parameter estimates: outcome = single factor by year~~

~~Joint tests: outcome = single factor by year~~

~~LR test: the reduced model without the interaction term (controlling for stratification variables)~~

~~(paste it in the front of SAS program)~~