

A Practical Checklist: Preparing Data for Analysis

1. Run frequencies on each variable
2. Perform cross tabulations
 - there shouldn't be a pregnant male or a 50 yr old pediatric patient
3. Check the range of the data
 - Minimum value
 - Maximum value
 - Do these results make sense for this variable? (there shouldn't be a 20 ft. tall patient)
4. Check the dates fall within correct time frames
 - patients' admit dates should be after patients' birth dates
 - are dates within the time frame of the survey?
5. Check the number of missing values for each variable
 - are these missing values expected?
 - are there more than the expected number of missing values?
 - could there be another problem?
 - accidental loss of data when importing or exporting data from the database to the analysis software
 - **unknown/don't know responses are handled differently than missing values**, check the unknown or don't know responses aren't set to missing values by mistake
6. If data are recombined to create a new variable, follow these same steps for the newly created variable.