

UNIVERSITY OF NORTH TEXAS
Department of Geography

GEOG 5800
Research Design and Geographic Applications
Dr. Murray Rice

Data Screening Checklist

It is important to consider all of the issues highlighted below prior to data analysis. If you look for issues later when you have problems with your analysis, you could be tempted to choose solutions based on how they can help you generate “better” results (and that wouldn’t be honest). Good statistics textbooks will give you more advice on what to do in each of these areas.

Inspect univariate descriptive statistics for accuracy

- Out-of-range values
- Plausible means and standard deviations
- Coefficient of variation
- Univariate outliers

Evaluate amount and distribution of missing data, and deal with problems

Check pairwise plots for nonlinearity and heteroscedascity

Identify and deal with non-normal variables

- Check skewness and kurtosis
- Transform variables (if needed)
- Check results of transformations (if done)

Identify and deal with multivariate outliers

- Variables causing multivariate outliers – can you identify a single variable?
- Description of multivariate outliers – anything in common?

Evaluate variables for multicollinearity and singularity

After Tabachnick and Fidel, Chapter 4, Using Multivariate Statistics