

Data Considerations

- Sample size (do we have enough power?)
- Multicollinearity and Singularity.
- Errors of prediction are normally distributed: Skewness, Nonlinearity, Homoscedasticity.
- The influence of outliers and influential cases.

How big should our N be?

- For testing the regression $N \geq 50 + 8k$ (where k = no. of predictors)
- For testing individual predictors $N \geq 104 + k$
- Usually we want to test both so calculate both equations and choose the one that produces the largest number.
- For example, with $k=6$ we require $N=98$ to test the regression and $N=110$ to test the individual predictors. We select 110.