Class Activity: Approaches to Missing Data

STAT 340: Applied Regression

Consider the following data set, extracted from the 1974 *Motor Trend* US magazine, which summarizes fuel consumption and 10 characteristics of automobile design, as well as performance for 32 automobiles manufactured between 1973 and 1974. The variables are as follows:

- `mpg`: miles/(US) gallon
- `cyl`: number of cylinders
- `disp`: displacement (cu. in.)
- `hp`: gross horsepower
- `drat`: rear axle ratio
- `wt`: weight (1000 lbs)
- `qsec`: 1/4 mile time
- `vs`: engine (0=V-shaped, 1=straight)
- `am`: transmission (0=automatic, 1=manual)
- `gear`: number of forward gears
- `carb`: number of carburetors

Each group will consider a different data set (`mtcars1`, `mtcars2`, or `mtcars3`) with missing data which I will assign at random. For each data set, implement the techniques for dealing with missing data that we have discussed, or assess the scenarios in which they are problematic to implement. Also comment on whether the method would be appropriate based on the missing data mechanism.