STOR 455
STATISTICAL METHODS I

Jan Hannig

Project

- I have decided to assign a project. It will be part of the midterm exam score (10% of the final grade).

- The two in-class midterm score make 40% of the grade.

- Due Thursday, December 2.

- There are three files (info, more info and data)

Comments

- Do not forget to plot the variables and look for patterns.

- When coding the categorical data, remember that you need to use one less variable (SAS will be very unhappy otherwise)

- Feel free to explore! Think about the meaning!

- Remember to use the test statement to test statistical significance.

Grouping of variables

- SAS allows for grouping of variables when running a data selection procedure.

model rootprice= carat {GIA IGI} {E F G H I}

{caratGIA caratIGI}
{caratE caratF caratG caratH caratI}

{GIAE GIAF GIAG GIAH GIAI IGIE IGIF IGIH IGIF IGII}

/selection=stepwise groupnames='carat'
'certfication' 'color'
'carat*certification' 'carat*color'
'certification*color';

Real Estate Example

- A data set containing data from 'recent' sale
- Variables are
  - ID, price, square footage, number of bedrooms, number of bathrooms, ac (1 present, 0 absent), number of cars garage can hold, year built, quality (3 high, 2 medium, 1 low), architectural style, lot size, adjacent to highway (1 yes, 0 no)
  - Which ones are categorical?
- See the SAS code from class