

## Module 5: Multivariate displays

(The attached PDF file has better formatting.)

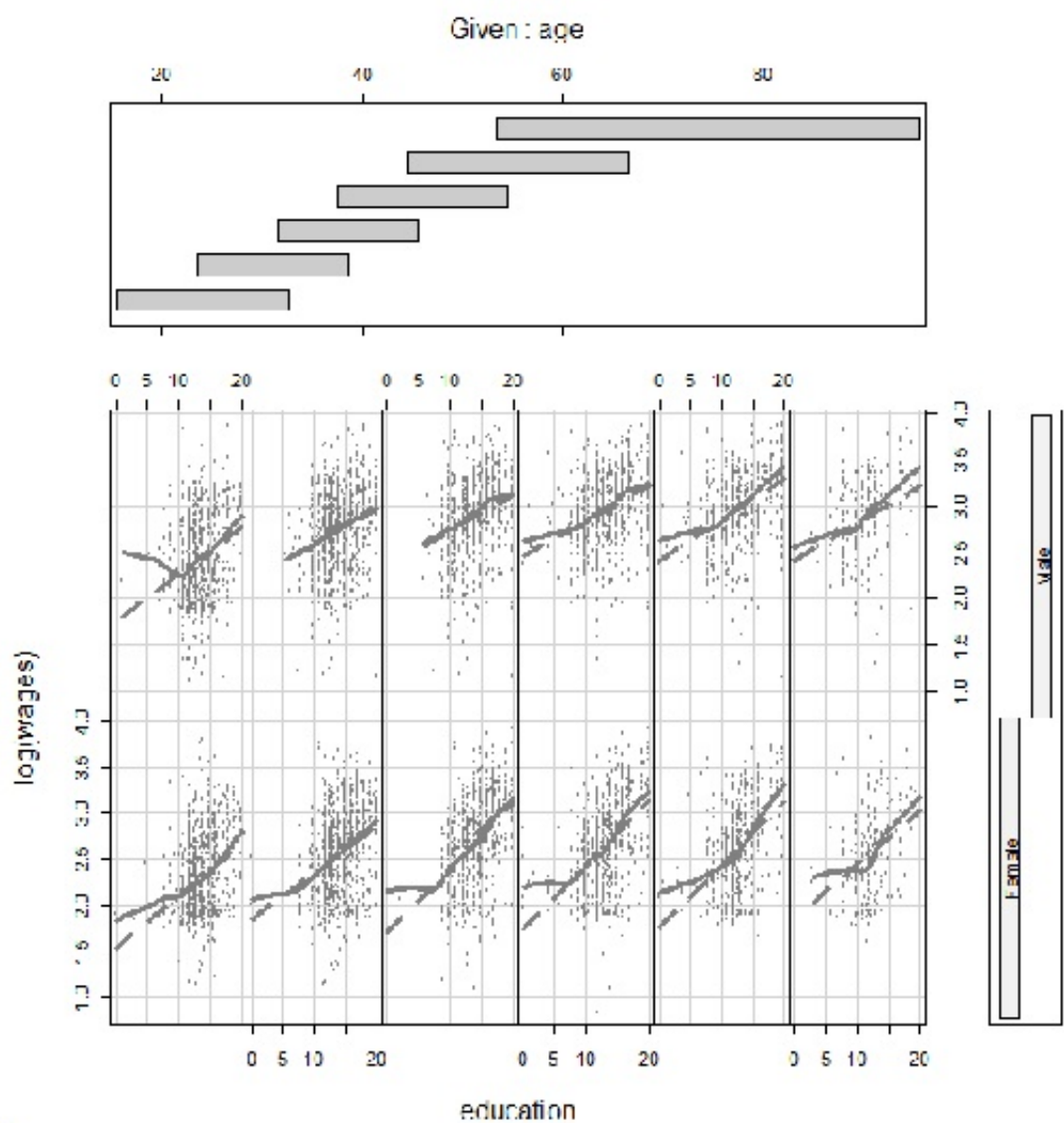
### *Homework assignment: conditioning plots*

This data set has 7,425 observations and 4 variables: age, sex, education, and wages.

- The dummy variable for sex has values 1 = female and 2 = male.
- The ages range from 16 to 95.
- The education is years of schooling (from 0 to 20)
- The vertical axis is the logarithm of wages (2,300 to 49,920 Canadian dollars).

Each graph has two lines.

- The straight (diagonal) line is a simple regression line.
  - The curved line is a lowess curve.
- A. What is the dependent variable? What is the independent variable? What are the conditioning variables?
  - B. How much do the ages overlap? Note that upper bound of one age is the midpoint of the next age and the lower bound of the next age.
  - C. The graph has 12 panes. Explain what the graph in the top row and second column from the left shows. (A one sentence explanation is sufficient.)
  - D. The lowess curves match the regression lines at higher education levels but lie above the regression lines at lower education levels. What does this imply?
  - E. What group has a stronger relation of wages to education level: women age 38 to 54 or women age 23 to 38?
  - F. What group has a stronger relation of wages to education level: women age 38 to 54 or men age 38 to 54?



Given: sex