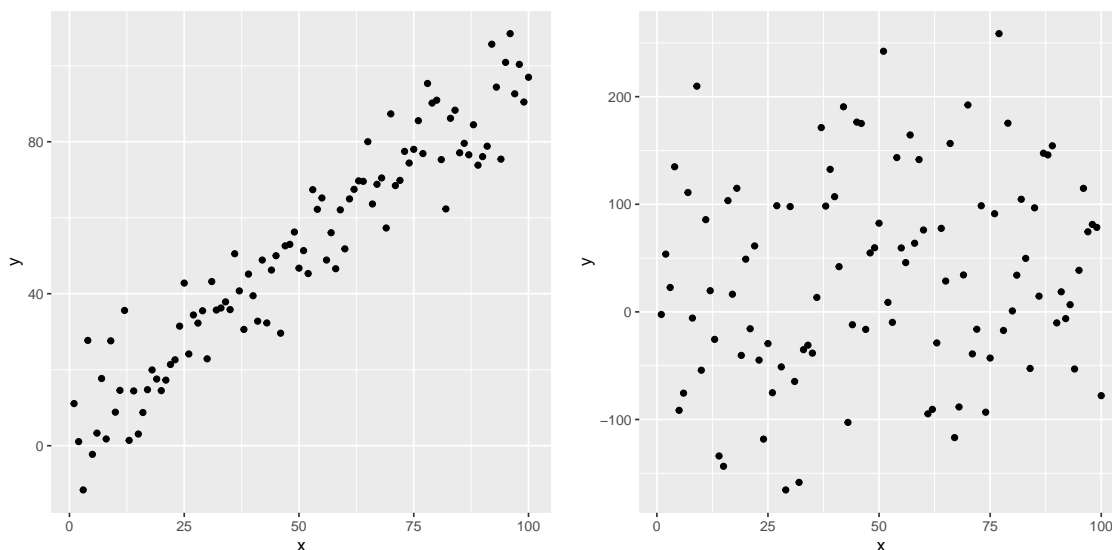


MA117 - WORKSHEET 11

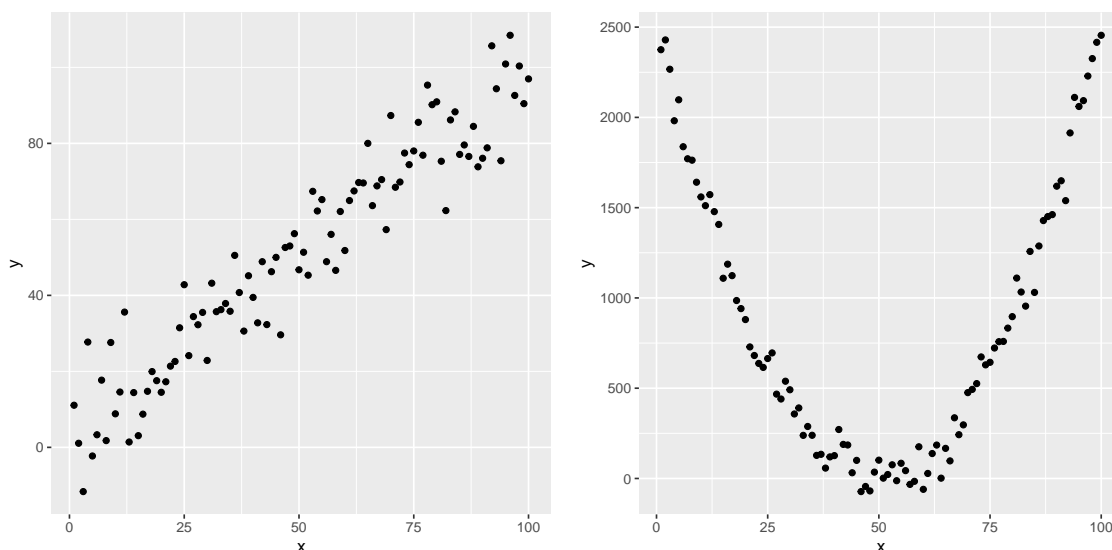
REGRESSIONS

March 18, 2021 - Week 3, Thursday

Problem 1. Determine if the correlation of x and y higher in the scatterplot on the left or on the right.



Problem 2. Determine if the correlation of x and y higher in the scatterplot on the left or on the right.



Problem 3. In each of the following situations, determine what the slope and y -intercept of the best fit line which relates women's salaries to men's salaries.

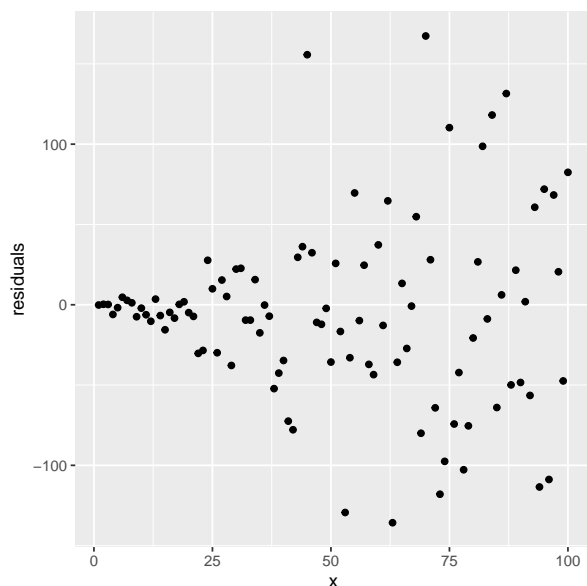
- (a) At a certain firm, men make 25% more than women do in any given position.

(b) At a certain firm, men make \$5000 more than women do in any given position.

Problem 4. Suppose we fit a regression line to predict the number of incidents of skin cancer per 1000 people from the number of sunny days in a year. For 2019, we predicted the incidence of skin cancer to be 1.5 per 1000 people, and the residual was 0.5. Does the regression line overestimate or underestimate the incidence of skin cancer in 2019?

Problem 5. Starbucks lists the calorie content of food menu items but not carbohydrate content. Suppose we gather data on some Starbucks food items and construct a best fit line to predict the carbohydrate content (in grams) using the calorie content. What are the *units* for the slope of the best fit line? Give a physical interpretation of this slope.

Problem 6. The plot below represents the *residuals* after we fit a least squares regression to some data. Should we have concerns about applying least squares regression to this data?



Problem 7. In each of the following plots, there is an outlier. Does it have *high leverage*? If so, is it *influential*?

