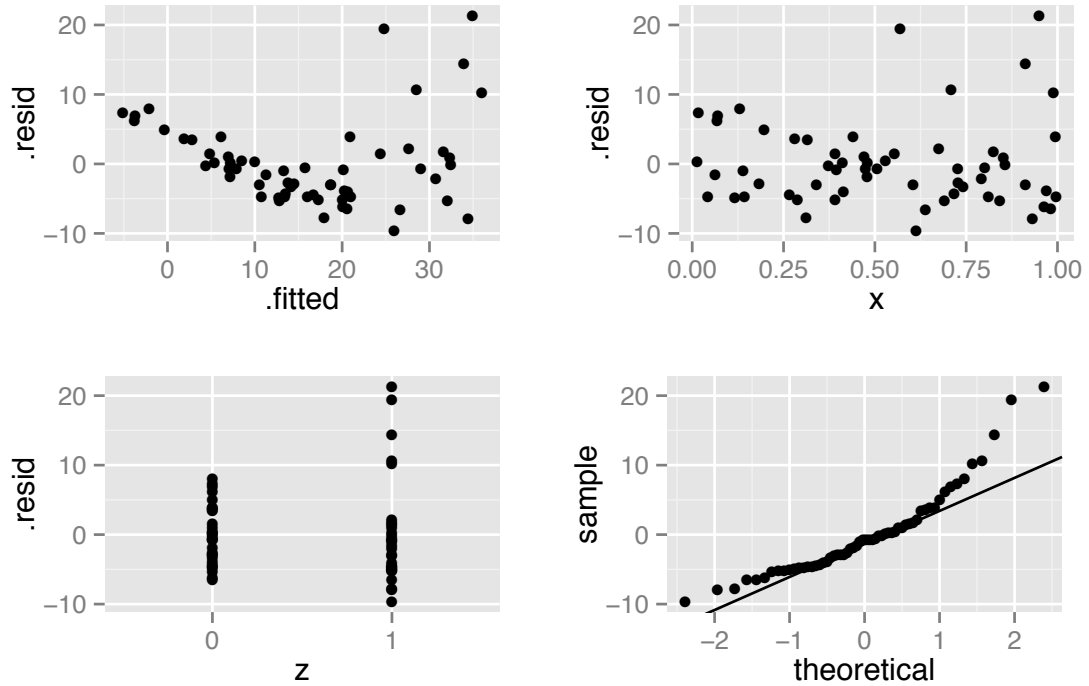


2. (a) The following residual plots come from a regression of the form:

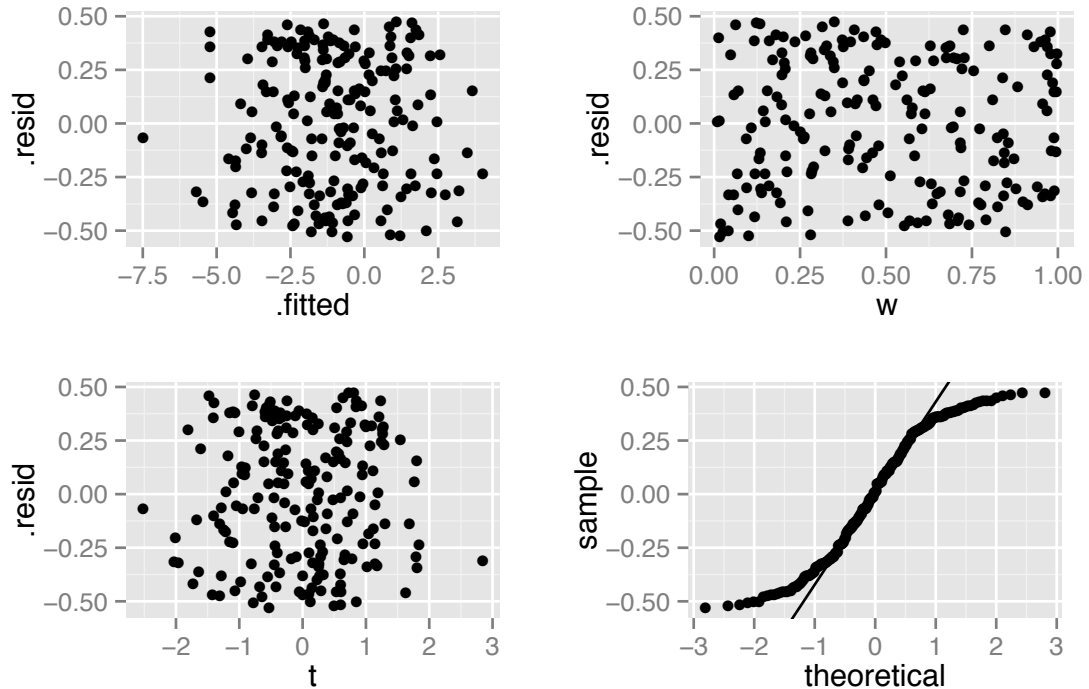
$$y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + \epsilon_i \quad i = 1, \dots, 60$$



- i. Name the assumption that appears to be violated. (2)
  
- ii. Describe the evidence you see in the plots for the violation. (1)
  
- iii. What are the consequences of proceeding with inference ignoring the violation? (1)
  
- iv. How would you suggest proceeding? (1)

(b) The following residual plots come from a regression of the form:

$$y_i = \beta_0 + \beta_1 w_i + \beta_2 t_i + \epsilon_i \quad i = 1, \dots, 200$$



- i. Name the assumption that appears to be violated. (2)
  
- ii. Describe the evidence you see in the plots for the violation. (1)
  
- iii. What are the consequences of proceeding with inference ignoring the violation? (1)
  
- iv. How would you suggest proceeding? (1)

(c) A client has run diagnostics on a regression analysis and identified a single observation with very high leverage, but she admits she doesn't know what leverage is or how to proceed.

i. What does *high leverage* mean? (1)

ii. Sketch a scatterplot that includes a point that has **high leverage** but is **not influential**. (Make sure you label your axes, clearly identify the point of interest, and label any fitted lines you add) (1)

iii. Sketch a scatterplot that includes a point that has **high leverage** and is **influential**. (Make sure you label your axes, clearly identify the point of interest, and label any fitted lines you add) (1)

iv. How would you advise your client to proceed?

(2)