Stat 6021: Homework Set 5 Solutions

2. (a) The scatterplot of y against x suggests the relationship is not linear. Looking at the residual plot, the constant variance of error term assumption is not met, as the vertical variation of the residuals increases. I would transform the response variable first. Doing so will enable us to stabilize the variance. Once that is achieved, we may also achieve linearity. If linearity is still not achieved, then I would transform the predictor. Transforming the predictor does not affect the variance of the error terms.

   (b) I agree. Based on the output, we should choose $\lambda = 0$, and apply a log transformation on the response variable.

   (c) $\hat{y^*} = 1.5079 - 0.4499x$, where $y^* = \ln(y)$. The predicted concentration of a solution is multiplied by $\exp(-0.44993) = 0.6376728$ when time increases by one unit. The predicted concentration of a solution is $\exp(1.5079) = 4.5172$ when time is 0.