# Stat 6021: Guided Question Set 3

A glass bottle manufacturing company has recorded data on the average number of defects per 10,000 bottles due to stones (small pieces of rock embedded in the bottle wall) and the number of weeks since the last furnace overhaul. Download the dataset "defects.txt".

1. Plot *defects*, the average number of defects per 10,000 bottles, against *weeks*, the number of weeks since the last furnace overhaul. Comment on the appearance of the plot. Do any assumptions or conditions for simple linear regression appear to be violated? If so, which ones?

2. Create a residual plot. Describe the appearance of the graph of residuals versus fitted values, and comment if assumptions are not met for simple linear regression.

3. Based on your answers to parts 1 and 2, do we need to transform at least one of the variables?

4. One of your classmates says that since she is not sure if the variance is constant, she should use the Box-Cox method to see if the response variable should be transformed first. Do you agree with her idea? Briefly explain.

5. Regardless of your answer to part 4, use R to produce a plot of the profile log-likelihoods for the parameter, $\lambda$, of the Box-Cox power transformation. What transformation, if any, would you apply to the response variable? Briefly explain.

6. Apply the transformation you specified in part 5. Then fit another simple linear regression model and produce the residual plot to assess if the assumptions are met.

7. Create an ACF plot of the residuals. Comment if assumptions are not met for simple linear regression.

8. Create a QQ plot of the residuals. Comment if assumptions are not met for simple linear regression.