STM4PSD – Workshop 11 Solutions

1. (a) Ask for help with this question if you need it.

   (b) Plots 3, 4, 6, 7 and 8 are clearly non-linear. We will therefore reject the validity of the simple linear regression model for the data in these plots. The fitted line in Plot 9 also does not bode well for the simple linear regression model. Note here that points above the line appear to be spread further from the line. Only Plots 1, 2 and 5 do not have any clear violations.

   (c) See below. Matching Plots 1, 2 and 5 is probably hardest (but not impossible).

Figure 1 label	Plot 1	Plot 2	Plot 3	Plot 4	Plot 5	Plot 6	Plot 7	Plot 8	Plot 9
Figure 2 label	D	E	G	B	I	A	C	H	F

   (d) There are clear violations of constant variance for the errors in E which we matched with Plot 2. Closer inspection of Plot 2 reveals that this does, in fact, appear to be a problem. Hence, the residuals versus fits plots aided in detecting a problem that was hard to spot in the original plot.

2. It can sometimes be difficult to decide on whether data resembles sampled data from a normal distribution when just looking at histograms. For example, the first histogram shows some signs of symmetry although the other histograms all seem to exhibit some skew.

3. (a) We can see that the Q-Q plots in the middle top row and bottom right are approximately linear. Therefore, the best guess we be that these are the Q-Q plots corresponding to sampled normal data. This is, in fact, correct.

   (b) The Q-Q plots were much more useful here. When n is not large, the histograms can be very sensitive to choice of bins and small differences in frequencies can look indicative of problems that are not there.

4. (a) Yes, there is a violation — the residuals versus fitted plot shows clear evidence of fanning.

   (b) Yes, with R^2 of approximately 0.8, this indicates a good fit.

   (c) The coefficient is 4.9864, indicating that, for each mm of diameter, the amount of toxin increases by $4.9864\,\mu\mathrm{g}$.

   (d) Yes, as the p-value is (substantially) less than zero.

   (e) $4.9864 \pm .2574 \times 1.992 = (4.473659, 5.499141)$

   (f) We are 95% confident that the average amout of toxin in a mushroom will increase by between $4.473\,\mu\mathrm{g}$ and $5.499\,\mu\mathrm{g}$ for each mm of cap diameter.

   (g) We are 95% confident that the amount of toxin in a mushroom with cap diameter $50\,\mathrm{mm}$ will be between $395.424\,\mu\mathrm{g}$ and $609.230\,\mu\mathrm{g}$