Assignment 6 — due March 4, 2019

1. An experiment was conducted to study the effect of factors that might affect the sweetness of wine measured in grams of sugar per liter of juice. Two experimental factors were of interest: the fertilization of the plants, and type of pruning method used. The fertilizer treatments consisted of: (1) amendment with nitrogen only (NO); or (2) amendment with nitrogen and phosphorus (NP). The pruning methods are defined in terms of the number of buds left. In this experiment pruning was done to leave either 50, 45, 40, or 30 buds. Pruning to 50 buds is least aggressive; pruning to 30 is most aggressive.

   The experiment was conducted in three vineyards at a large winery. We will simply label these as the East, West, and Central vineyard. Within each vineyard, two rows of vines were chosen at random, and each one was randomly assigned to receive one of the fertilizer treatments. Within each of the rows, four trunks were chosen, and then each of the trunks was randomly assigned to receive one of the four pruning methods. Finally, at harvest time, two grapes were sampled from each trunk (call them Grape 1 and Grape 2) and the sweetness of each grape was determined. The data are in file: `grapes.csv`.

   (a) Make a suitable plot to determine whether there is evidence of an interaction between fertilization and pruning method.

   (b) Analyze these data in `R`: make the appropriate tests for interaction and for main effects. (Test the main effects even if the interaction is significant.)

   (c) Check residuals, and comment.

   (d) Suppose that the experimenters had instead decided to focus on the NP fertilizer treatment and a 40-bud pruning regime, using the East vineyard only. Further, suppose they observed 8 trunks (placed in 2 rows) and 3 grapes per trunk. Imagine calculating the mean sweetness from the resulting 24 grapes. Estimate the variance of this mean.

2. The manufacturing of semiconductor (electronic) components often involves a number of steps, including "ion implantation" and "furnace annealing" applied to units called "wafers." An experiment was conducted to evaluate aspects of these two steps. First, for the ion implantation step, two different experimental factors were considered: the energy level used (either High or Low), and the "quantity" of ions (either High or Low) directed toward the target. Using standard notation, we describe the combinations of these two factors at their various levels as: EQ, Eq, eQ, and eq.

   The experiment was conducted as follows. Twelve wafers were selected from a large supply of wafers and divided into 4 equal-sized batches. At random, each batch was assigned one of the treatment combinations EQ, Eq, eQ, or eq. All wafers in a batch were treated at the same time.
   The next step was furnace annealing. For the furnace annealing step, 3 different temperatures were studied: 750 C, 800 C, and 850 C. In this step, the original 12 wafers were divided into 3 annealing batches of size 4. This was done carefully so that, within each annealing batch, each of the 4 ion treatments appeared: EQ, Eq, eQ, and eq. The 3 annealing batches were randomly assigned to the three temperatures. All wafers in an annealing batch were treated at the same time.

   After this, each wafer was measured, and for our purposes we will focus on a quality measure which ranges from 0 to 100; 100 represents the highest quality. The entire experiment was performed twice, once in July (J) and once in August (A). In each month a new set of wafers was used, and the randomization of batches was redone. The data are in file `ion.csv`.

   (a) Write down a model for this experiment, defining all terms in the model, and summarize the assumptions that accompany the model.

   (b) Construct a useful plot to see whether there is evidence of a three-way interaction between Energy, Quantity and Temperature. You may want to create new columns `energy` and `quantity` to separate the two aspects of the ion implantation step. For this, use the `substring` function in `R`.

   (c) Using `R` to fit these data, calculate (and interpret) p-values for each of the effects and interactions.

   (d) Comment on the results of part (c) in light of your plot in part (b).

   (e) Construct suitable residual plot(s) and comment.