First data analysis example

A study is performed to assess the effect of irrigation and variety on the yield of corn. Irrigation is set at two different levels, high and low. There are four varieties of corn, A, B, C, and D.

The experiment used 8 different fields: 4 are randomly assigned to have high irrigation and the others receive low irrigation.

Within each field there are 4 different plots. One of each variety is assigned to each of the four plots within each field.

There are a total of 32 observations of the yield of corn in bushels per acre.

1. Describe how to make graphs of the data that will be descriptive about the primary scientific questions.

2. What other graphs may be useful to demonstrate important features of the data?

3. Select an initial statistical model for the data that predicts corn yield.

4. What interactions should be included in the model? What evidence (graphical, hypothesis testing, experimental design) favors this conclusion?

5. Which factors are better modeled as fixed effects and which are better modeled as random? Justify your responses.

6. Address the scientific questions of interest:

   (a) Is high or low irrigation better for larger crop yield? Does the answer depend on the variety? Estimate the size of important effects and provide error estimates.

   (b) Which variety or varieties yield the most? Does the answer depend on the irrigation level? Estimate the size of important effects and provide error estimates.

   (c) How should variation among fields be modeled so that its impact on questions of interest is controlled?