Data Overview

The study contained observations of different species from 31 agricultural studies conducted in the USA and Europe. The provided dataset is heavily simplified from the original dataset found on Dryad. This was to make the data processing easier to follow. Below are the variables we will be using in this lesson with information on how they were derived.

Study

This variable is the number assigned to each of the agricultural studies. Some studies have multiple data points. We will need to account for multiple data points coming from one study in our analysis.

Taxonomic Group

This variable is the taxonomic group that was studied in each of the agricultural studies. These groups include vertebrates, invertebrates and plants.

Crop Type

This variable is the type of crops being studied by each of the agricultural studies. These included mixed, cereal, pasture/meadow, and fruit/vegetable crops.

Local Intensity

This variable is the local management effect size on species richness as a log response ratio :

$$LR_M = \ln(yL/yH)$$

where yL is the mean of species richness in low-intensity farms and yH is the mean of species richness in high intensity farms. This effectively normalizes the data so that we can compare data between studies. This allows us to see if there is a proportional difference between mean levels of species richness in low and high intensity farms.

Landscape Complexity

This variable is the Fisher's Z value which is the distribution of Pearson's correlation coefficient transformed to be normally distributed for landscape factor values from the agricultural studies. This variable allows us to compare the linear correlation between multiple sets of data because it is a method of normalization. The researchers defined the landscape factor as percent natural area, semi-natural area and woodlands as per cent non-crop area. They also included the diversity of habitat types (measured as the Shannon's Index) in this factor. This effect size variable was calculated as the correlation coefficients (R) that related y (richness) to the measure of landscape complexity. They then standardized the coefficients to Fisher's Z as :

$$Z_L = 0.5 \times \ln(1 + R_L)/(1 - R_L)$$

where Z_L is Fisher's Z and R_L is the correlation coefficient of y versus landscape complexity.