Data examples.

**Arabidopsis nitrogen.csv:** You are investigating the role of nitrogen on plant growth using the model organism *Arabidopsis* using hydroponic tanks. You randomise 4 tanks to one of 2 levels of nitrogen: 02 or 2. Plants are harvested at either Day 10 and 24, dry mass is measured.

**Breast cancer density.csv:** **Cases** (Case = 1) were women who developed breast cancer after their first follow-up mammogram and **Controls** were still breast cancer-free at the time. **Age** (years) at first mammogram, body mass index at first mammogram. The researcher wants to assess whether breast density was a risk factor in this case-control study. She is also interested in the relationship between breast density and age and body mass index. Were these factors correlated? What was the strength of the association?

Does the relationship between breast density and risk differ by subgroups of (i) age, (ii) body mass index and (iii) treatment (tamoxifen (ARM = 2) or placebo (ARM = 1)).

**chloroquine.csv:** You are investigating whether co-administration of chloroquine during sporozoite immunization (RAS) improves T cell activation in response to *Plasmodium berghei* challenge in mouse models. In this experiment, you used 8 cages with 4 mice per cage: mice from 4 cages received chloroquine; in the other 4 cages, no chloroquine was administered. Two mice from each cage were randomly selected to RAS immunization. Mice were challenged with either 10,000 or 50,000 sporozoites. 21 days post-challenge, animals were euthanized and liver and spleen extracted. CD8+ T memory cells specific to *Plasmodium berghei*  were measured using an IFNγ assay. You are interested in estimating the organ specific effects, as well as an overall effect, if appropriate.

**complex formation.csv:** We have developed an assay to measure the formation of platelet-mediated erythrocyte-parasite complexes in the presence of *P. falciparum* infected erythrocytes, expressed as the percentage of erythrocytes (out of 500) that have formed complexes. We tested our assay using infected and uninfected erythrocytes, under 3 different platelet:erythrocyte concentrations (0.02:1, .2:1, and 2:1), and at 2 different incubation times (2 hours and 24 hours). We ran the assay at 5 different periods. We are interested in understanding how concentration and incubation time affect complex formation, and if these effects differ by infection status.

**dark respiration.csv:** This researcher was interested different species of plants differ in dark respiration, and whether older leaves are different to younger leaves wrt dark respiration. In this data set, there were 3 plant species. Each plant contributed 3 leaves of different ages. Dark respiration was measured at multiple points on each leaf.

**diet and diabetic mice.csv:** Are diabetic-prone mice more susceptible to high fat diets compared with non-diabetic-prone mice? In this experiment, two strains (NODk and b10.Br) of baby mice were randomised to either a high fat diet or regular chow. Weights (g) were recorded over a 14 week period. Whilst the high fat diet was expected to increase weight gain compared with regular chow, the hypothesis was that the high fat diet would have more of an impact on weight gain in the NODk mice.

**drought data.csv:** We have introduced a mutation in a tomato strain, and we have tested the effect of the mutation on leaf temperature. Plants of each genotype were randomised to receive either drought treatment or normal watering conditions. After 7 days, we measured leaf temperature of the second youngest leaf. The hypothesis was that the mutated strain would be less affected by drought conditions relative to the wild type strain.

**drought and leaf water retention.csv:** In this mock data set, plants were genetically modified to one of 4 genotypes: AB, Ab, aB or ab. We are interested in understanding the effects on leaf water retention when knocking out gene A and/or gene B. Are these effects additive?

**Leaf\_properties.xlsx:** How does leaf shape vary between plant species? This data set invites you to take a multivariate perspective comparing leaf size and morphology between plant species. How do size and morphology influence leaf robustness (measures of shear strength)?

**mock LWR.csv:** Can one improve drought tolerance in Arabidopsis through gene knockouts? In this experiment, we knocked out compared drought tolerance amongst 4 genotypes: AB (wildtype), Ab (B KO), aB (A KO) , or double knockout ab. We exposed all plants to drought conditions for 7 days, and then measure the percentage leaf water retention in the largest rosette leaf for each plant.

**nematode.csv:** This researcher has identified several genes that may be important for nematode resistance in legume plants. He developed three mutant genotypes in which some of these gene are knocked out. He exposes half of his plants to nematodes (Treatment = Infected), and recorded plant characteristics after a fixed time period. He repeated this experiment over three time periods (Experiment = 1, 2, or 3).

**pea data.csv:** Peas were grown under 5 different growth media which differed in the type of sugar used. The different types of sugar (including a no sugar control) were: control, glucose, fructose, g&f and sucrose. The experimenter recorded the lengths of pea sections. Ten pea section lengths were recorded per treatment.

**photosynthesis.csv:** This was a greenhouse experiment, designed to test the effect of soil temperature on photosynthetic rate. The experiment was conducted on 4 different tables (Position) in the greenhouse. On each table, there were 6 plants; their soils were exposed to either normal (temp=1) or elevated (temp=2) temperature.

**plant growth.csv:** Tree seedlings were planted in a 6 x 6 grid, to assess the effect of six different nutrient doses on plant growth (height). The treatments were arranged so that each treatment was represented in each row and each column of the grid (Latin Square design). The researcher is primarily interested in the Dose effect, but would also like to assess the effect of grid position on growth.

**Plastic Bag Survey Results.xlsx:** This was a survey conducted in the ACT to assess plastic bag consumption amongst grocery shoppers. Shoppers were surveyed at randomly selected Woolworths, Coles and Aldi stores. To select the stores, the ACT was stratified into its five electorate divisions: Kurrajong, Ginninderra, Yerrabi, Murrambidgee and Brindabella. Surveys were undertaken at each selected store on the weekend and during a weekday. There were three stores where surveying was not able to be undertaken on both a weekday and the weekend.

**respiration data.csv:** This experiment was designed to assess whether the temperature effect on dark respiration depends upon plant genotype. Genotypes were labelled A/B. The researchers exposed different leaves of the same plant to 4 different temperatures.

**seed orchard data.csv:** A field trial was planted to compare a seed lot (i.e. a batch of seeds) derived from a seed orchard (SO) with one collected from a routine plantation (P). Eight plots were planted for each seed lot, and these were thinned when the young trees were seven years of age. Tree diameters at breast height (dbh) were measured at 15 years.

**thorndata.csv:** Ecology researchers recorded the density of thorn-like plants (thorndensity) in multiple locations across five regions (site), and measured per hectare consumption of plant material by herbivores (herbivory). Their objective was to model the relationship between thorn density and herbivory.

uc40.csv and uc200.csv: Eighty(80) (or 400) patients with active ulcerative colitis (UC) were recruited and randomised in a 1:1 ratio to receive either infliximab or placebo. A total of 12 (30%) of the 40 patients assigned to placebo achieved a clinical response, compared with 20 (50%) of the 40 patients assigned to infliximab. The same percentages achieved a clinical response in the 400 patient study. The purpose of these data sets is to compare the information carried by having five times as many patients.