

## Exploratory data analysis – Barley data

**Goal:** Can you predict the protein content in barley? How does mutant and/ or location (growth) influence the protein content?

**Filename:** barley.ltx

**Number of measurements:** 388

**Mutant** (388 × 1): There are three barley mutants (1 = N+ - Normal, 2 = P – Protein mutants, 3 = C – Carbohydrate mutants and the rest (0) are other).

**Growth** (388 × 1): The barley has been grown in three main locations (1 = Greenhouse, 2 = Outdoor and 3 = Field).

**Protein** (388 × 1): Protein content as measured by the Kjeldahl method. (Validation set indicated by missing values)

**NIR** (388 × 1050): NIR spectra of the milled barley samples.

**Wavelength** (1 × 1050): Wavelength in nm.

**Object names:** Indicate sample number, location (G/ O/ F) and mutant (N+/ P/ C).

### Description:

388 different barley samples have been made by milling several barley seeds taken from different barley mutants and locations. The outdoor location means that the barley has been grown in pots outside of the greenhouse, compared to the greenhouse, where they have been grown in pots inside. The field is, of course, plants grown (less controlled) in the field. In addition to measuring the NIR spectra on these barley samples, the protein content in the milled samples have also been determined by the Kjeldahl method.

### Acknowledgement:

These data have kindly been made available by Lars Munck and Birthe Møller Jespersen.