# Data exploration, regression, GLM and GAM course

Highland Statistics Ltd www.highstat.comn

# Exercise 1: Data exploration Loyn data

## *Data description*

Forest bird densities were measured in 56 forest patches in south-eastern Victoria, Australia. The aim of the study was to relate bird densities to six habitat variables; size of the forest patch, distance to the nearest patch, distance to the nearest larger patch, mean altitude of the patch, year of isolation by clearing, and an index of stock grazing history (1 = light, 5 = intensive). The variables are given in Table 1.

The data are in the file loyn.txt. Note the spelling! Each row in the text file represents one forest patch.

**Table 1. Variables in the file loyn.txt.**

| Variable | Description | Type |
|---|---|---|
| ABUND | Density of birds in a forest patch | Continuous response variable |
| AREA | Size of the forest patch | Continuous explanatory variable |
| DIST | Distance to the nearest patch | Continuous explanatory variable |
| LDIST | Distance to the nearest larger patch | Continuous explanatory variable |
| ALT | Mean altitude of the patch | Continuous explanatory variable |
| YR.ISOL | Year of isolation by clearance | Continuous explanatory variable |
| GRAZE | Index of stocking grazing intensity | Nominal (ordinal) explanatory variable with levels 1 (light) to 5 (intensive) |

## *References*

These data were originally analysed in Loyn (1987), and again in Quinn and Keough (2002).


- Loyn RH (1987) Effects of patch area and habitat on bird abundances, species numbers and tree health in fragmented Victorian forests. In: Nature Conservation: the role of remnants of native vegetation (Saunders DA, Arnold GW, Burbidge AA, Hopkins AJM, eds), pp. 65-77. Surrey Beatty & Sons, Chipping Norton, NSW
- Quinn GP Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press


## *Underlying question and task*

The variable ABUND is the density of birds in 56 forest patches. The explanatory variables are size of the forest patches (AREA), distance to the nearest forest patch (DIST), distance to the nearest larger forest patch (LDIST), year of isolation of the patch (YR.ISOL), agricultural grazing intensity at each patch (GRAZE) and altitude (ALT). The underlying aim of the research is to find a relationship between bird densities and the explanatory variables.

The task for the moment is to apply a data exploration.