ST2304 - Statistical Modelling for Biologists/Biotechnologists

Bob O'Hara

bob.ohara@ntnu.no

This week we will...

Start the course

- admin
- try to motivate you
- overview of the course

Start learning R

- introduction
- hands-on work

Administration Matters

(we will deal with these in more detail later)

- Reference Group
- Blackboard
- web page: https://www.math.ntnu.no/emner/ST2304/2021v/

How the Course Will Run

Online: Blackboard + Zoom

Will (try to) pre-record lectures & put online at the start of the week

Mainly group work during lectures

Assessment

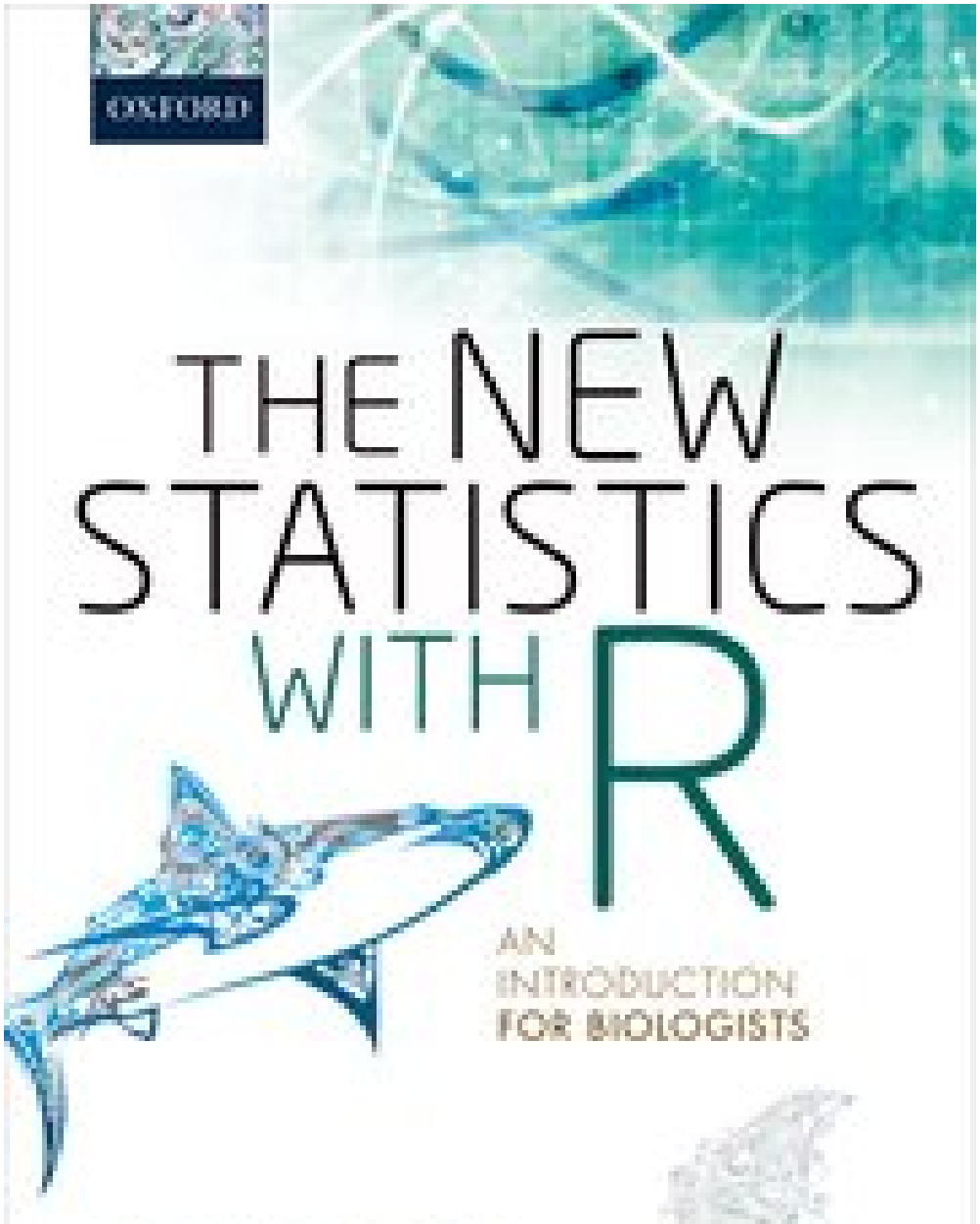
Complete 8 exercise sets (of about 10)

- do in groups
- pass/fail
- first couple of weeks won't count
  - we will tell you when they start to count

A Home Examination

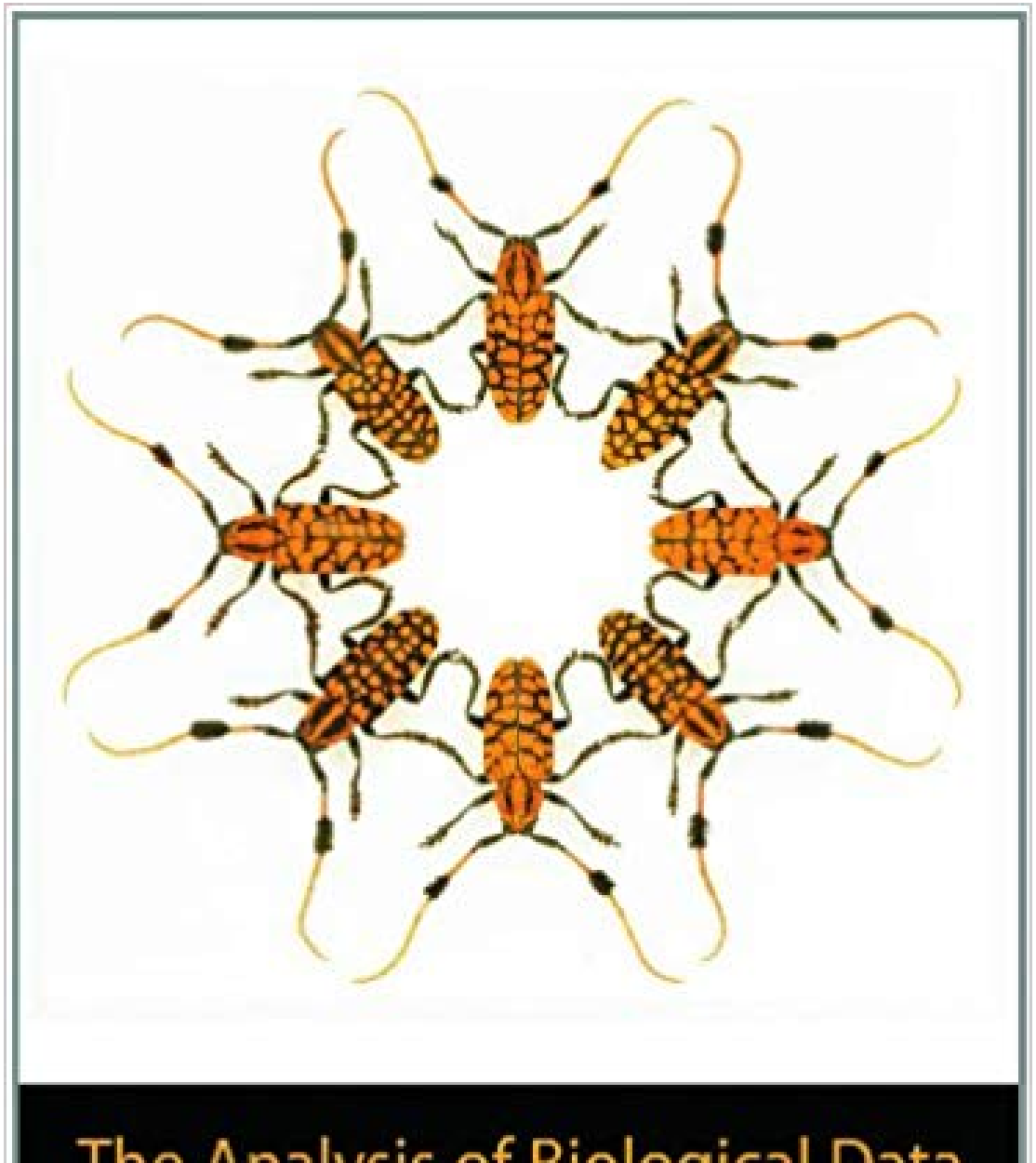
Text Books

New Statistics with R - Andy Hector



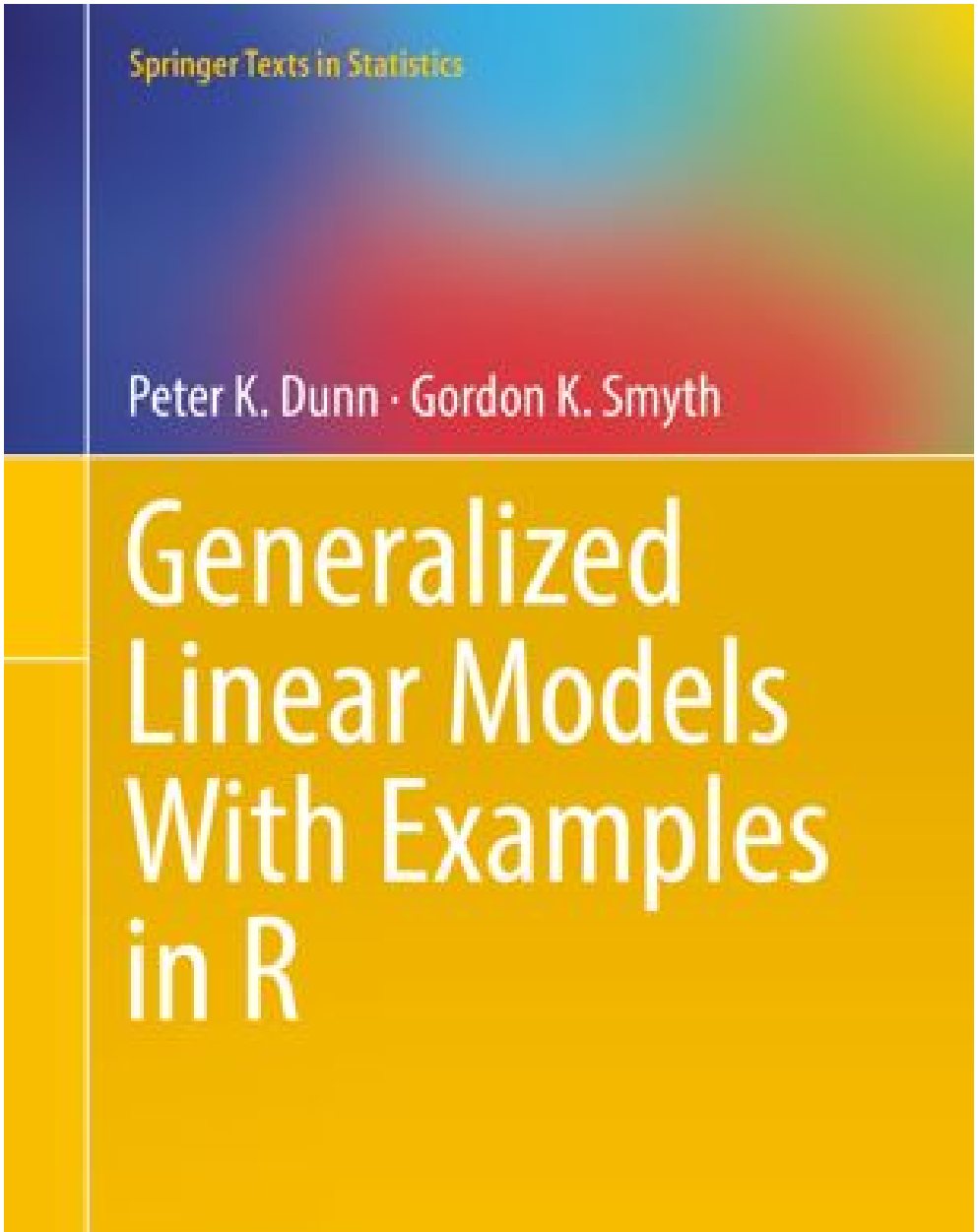
Text Books

The Analysis of Biological Data - Whitlock & Schluter



Text Books

Generalized Linear Models With Examples in R - Dunn & Smyth



Recap: why do we we use statistics in biology

What do you remember from ST0103?

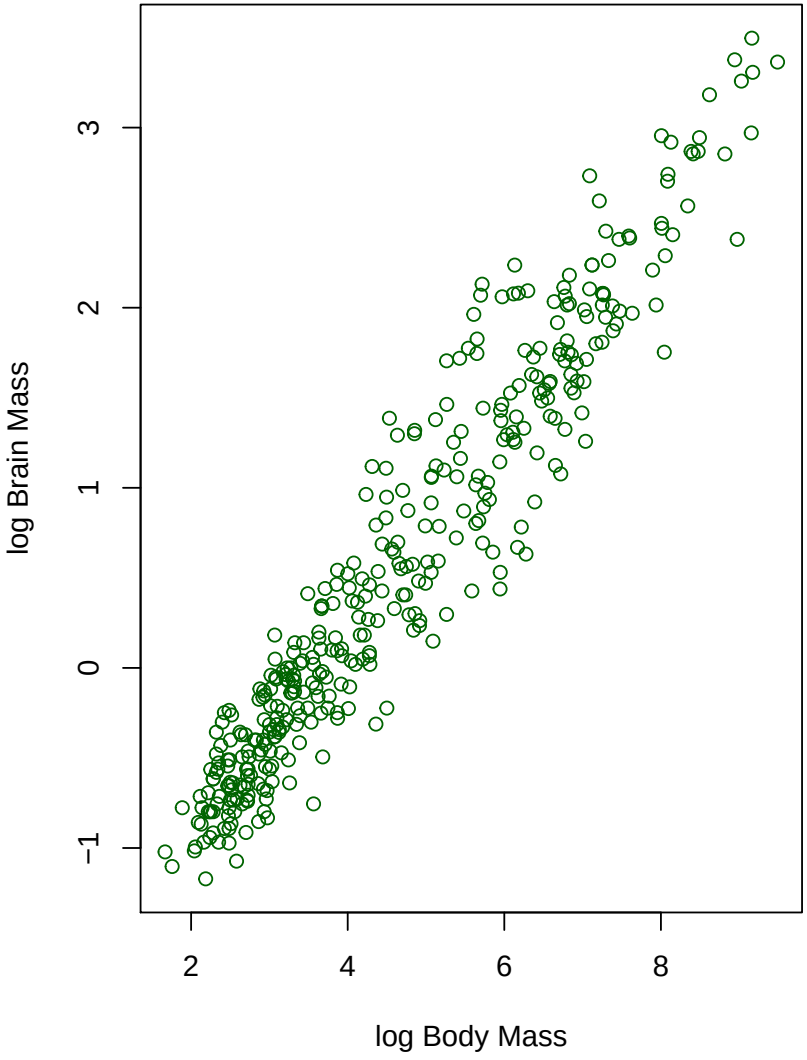
- discuss in your groups, and come up with 3 topics you learned about, and for each give an example where they are used in biology (or biotechnology!)
- be prepared to explain your choice to the whole class

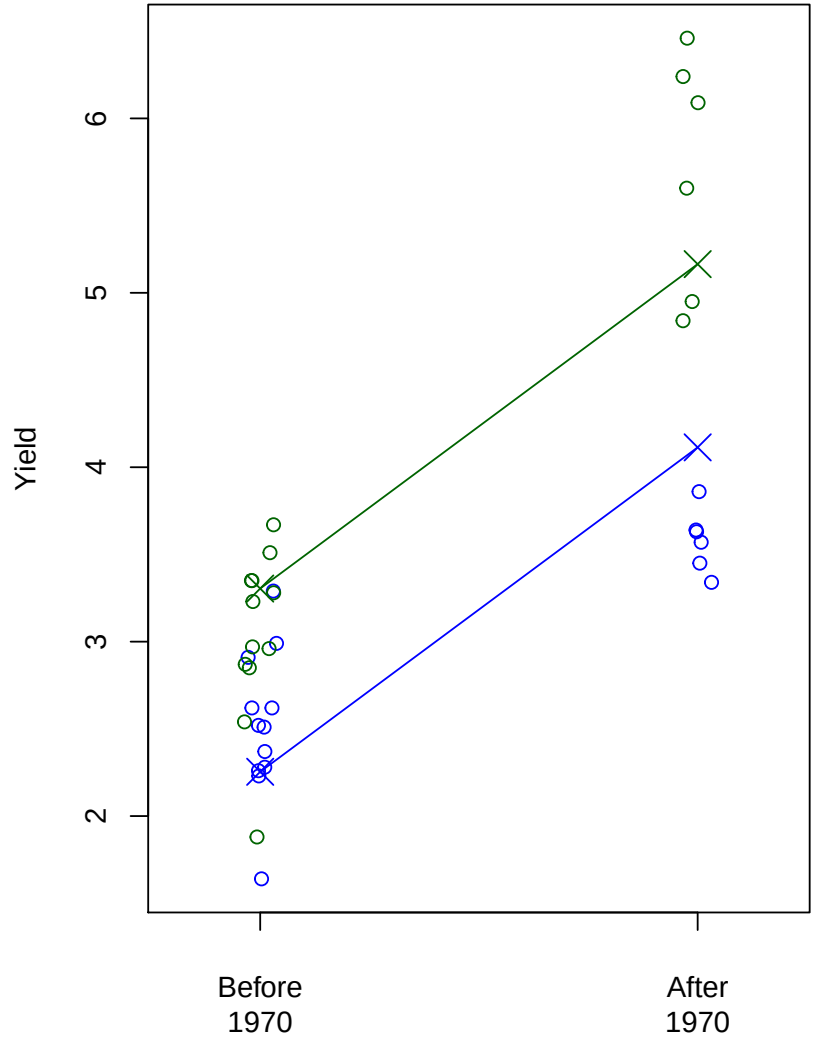
What we are aiming for

We want you to be able to analyse your own data (and understand what you are doing!)

- fit the right models to data
- assess if the model is any good
- compare models and decide which is 'best'
- interpret the models

Types of model I: Linear models (regression, ANOVA)

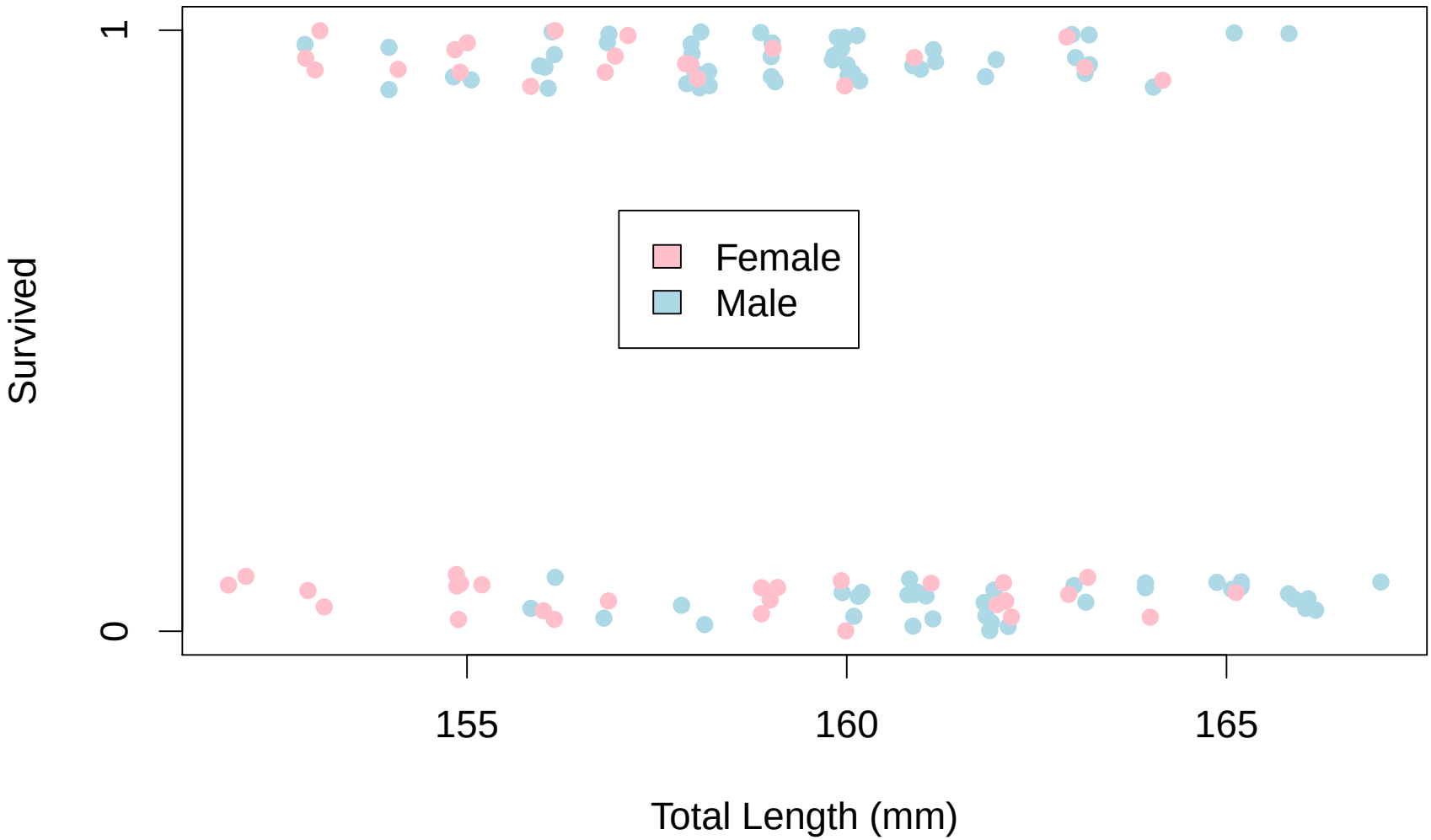




Types of model I: Generalised Linear models

When things aren't normal

- binary (e.g. survive/died)
- counts (e.g. how many sparrows are there?)



Likelihoood

The statistical framework to do this

Likelihood = probability of the data

- means we can write everything as probabilities

R

The stats package we will use

- free, most commonly used
- more shortly

Course Structure

Weeks 1-3: Likelihood and R

- statistical theory, and programming
- the background you will need to understand what follows, and to do it

Weeks 4-10: Linear models (regression, ANOVA)

- fitting straight lines

Weeks 11-13: Generalised Linear models

- fitting straight lines to different types of data

How we will teach

Active Learning

Group work

Problem solving