Course outline weeks 7–11
27th April–29th May

BIOSCI220

Students are expected to be able to

- Identify the following types of variables in the dataset
  - discrete
  - continuous
- Write their own R code to summarise each variable using an appropriate plot. Specifically be able to produce the following plots
  - boxplot
  - scatter plot
  - histogram
  - barplot
- Represent their data as a clean data frame in R and manipulate it accordingly. Use of [& : etc.
- Communicate a summary of the dataset accurately and concisely
- Communicate any limitations relevant to the dataset

Short pre-lecture quiz (not assessed)

Week 7 (beginning 27th April): Experimental design

Learning objectives

By the end of this lab student should be able to

- Identify the independent and dependent variable in an experiment
- Describe the setup of the following experimental designs
  - independent measures
  - repeated measures
  - matched pairs
- Identify the type of design used in the given case studies
- Discuss the advantages and disadvantages of different design types
- Critique experimental designs used in the the given case studies
- Design an experiment that follows good experimental design principles

Mini lectures

- Randomization
- Replication
- Local Control (blocking etc.)
- Lab tutorial explained (group presentation)

Assessed quiz on CANVAS worth 1% of your final grade (Practise quiz here)

Lab—peer assessed worth 6% of your final grade

- In groups discuss the experiments covered in these articles. Identify the following.
  - explanatory variable
  - response variable
  - treatment
- Discuss the experimental design. For example, is randomisation, blinding, placebo, or matched pair design, used in the experiment?
- Discuss potential flaws of the experiments. For example are there any confounding or interacting variables?
- Each group will then be paired with another. In turns—in each pair—one group will present thoughts on the studies design, results, and flaws to the other and be assessed anonymously according to these (**TODO link assessment sheet**) criteria.

Week 8 (beginning 4th May): Visualising and analysing multivariate data

Learning objectives

By the end of this lab student should be able to

- Discuss the aims and motivations of Multidimensional Scaling (MDS) and its relevance in biology
- List and summarise the three main types of MDS:
  - classical MDS
  - metric MDS
  - non-metric MDS
- Write R code to create an MDS plot appropriate to the given dataset
- Interpret an MDS plot
- Explain the aims and motivation behind Principal Component Analysis (PCA) and its relevance in biology
- Write R code to carry out PCA
- Interpret the effectively communicate the output of a PCA

Mini lectures

- Intro to MDS and scaling
- When the distances are Euclidean (PCA & bread)
- Carrying out and drawing inference PCA in ‘R‘
- Lab tutorial explained (PCA cheatsheet)

Assessed quiz on CANVAS worth 1% of your final grade (Practise quiz here)

Lab—worth 6% of your final grade

Labs 9–11 will be assessed via a one page Executive Summary due 11.59pm on the 31st May 2020.

This Executive Summary will be worth 18% of your final grade and is in addition to he weekly CANVAS quiz. During lab time you are free to work through the material provided and your final report. You may work in groups, however, the final report must be your **own** work. Any plagiarism will automatically result in 0% for the report. Your Executive Summary should be no more than one A4 page. If should concisely effectively communicate your hypothesis, the statistical analysis undertaken, and your findings. Here are some guidelines to follow when writing your executive summary

- It should **not** contain any statistical terminology that would only be properly understood by a statistician
- Recall there are a set of main messages that you should report from your analysis, the reader doesn't need to know about all the work you carried out
- A brief outline of an Executive Summary should follow the sections listed below
  - Introduction: a one or two sentence description of the data and the purpose of the analysis
  - Methods: important **non technical** information for the reader about the analysis carried out
  - Report findings and the strength of evidence for them
  - Quantification: how reliable/generalisable are those findings
  - Summary: a one or two sentence summary of the major findings

Week 9 (beginning 11th May): Hypothesis testing

Learning objectives

By the end of this lab student should be able to

- List appropriate questions posed by the biological questions and outline an appropriate hypothesis test that would answer it
- Describe the aims of the following hypothesis tests
  - one-sample t-test
  - two-sample t-test
  - randomization test
- List the aims of hypothesis testing and write out the appropriate null and alternative hypothesis using statistical notation
- Write R code to carry out an hypothesis test using the appropriate variables in their dataset. Specifically write R code to carry out
  - one-sample t-test
  - two-sample t-test
  - randomization test
- Interpret and communicate the findings of an hypothesis test accurately and concisely
- List the limitations of the hypothesis in relation to the questions posed by the data

Mini lectures

- Hypotheses, why?
- Differences in mean
- Randomization tests
- Lab tutorial explained (executive summary)

Assessed quiz on CANVAS worth 1% of your final grade (Practise quiz here)

Week 10 (beginning 18th May): Introduction to linear modelling

Learning objectives

By the end of this lab student should be able to

- Develop a biologically relevant question of interest from the dataset and identify the following types of variables in the dataset
  - response variable
  - explanatory variables
- Express their question of interest accurately and concisely
- Carry out and interpret tests for the existence of relationships between explanatory variables and the response in a linear model
- Write R code to fit a linear model with a single continuous explanatory variable
- Write R code to fit a linear model with a continuous explanatory variable and a factor explanatory variable
- Interpret estimated effects with reference to confidence intervals from linear regression models. Specifically the interpretation of
  - the intercept
  - the effect of a factor
  - the effect of a one-unit increase in a numeric variable
  - the effect of an x-unit increase in a numeric variable
- Critique the fitted model

Mini lectures

- ANOVA $\equiv$ regression
- Explanatory variables and the response
- Fitting and interpreting linear models in 'R'

Assessed quiz on CANVAS worth 1% of your final grade (Practise quiz here)

Week 11 (beginning 25th May): Modelling II

Learning objectives

By the end of this lab student should be able to

- Write R code to fit a linear model with interaction terms in the explanatory variables
- Interpret estimated effects with reference to confidence intervals from linear regression models. Specifically the interpretation of
  - main effects in a model with an interaction
  - the effect of one variable when others are included in the model
- Explain why you may want to include interaction effects in a linear model
- Describe the differences between the operators : and * in R model-fitting formulae
- Critique the fitted model

Mini lectures

- Multiple explanatory variables
- Interactions
- Model diagnostics

Assessed quiz on CANVAS worth 1% of your final grade (Practise quiz here)