****

A graph showing different colored dots

Description automatically generated with medium confidence

**Biology 2A: Data Exploration**

**Course Guide**

**2023/2024**

**Course Organiser: Dr Catherine Kidner**

**Course Secretary: Ms Karen Sutherland**

**Course Email:** [**bio.2A@ed.ac.uk**](mailto:bio.2A@ed.ac.uk)

**Course code: BILG08024**

This **Course Guide** provides essential information about the course and schedules for activities. You should read this in conjunction with the **BTO** **Essential Guide**.

More detailed information is provided on the course **Learn** site. Any changes to the course will be communicated there so it is important to check regularly for announcements.

Banner image shows a Multi-trait Manhattan plot for soybean varieties from Bruce et al., (2020)  
Theoretical and Applied Genetics (2020) 133:1967–1976<https://doi.org/10.1007/s00122-020-03569-1>

**If you require this document in an alternative format, please e****mail:** [**bio.2A@ed.ac.uk**](mailto:bio.2A@ed.ac.uk)**.**

# Assessment Information and Deadlines

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Assessment** | **Weighting** | **Date** | **Submission** | **Feedback** |
| Online Quizzes before most classes | 20% | Monday and Friday  classes 2, 3, 4, 5, 6, 9, 10, 12, 13, 14, 15, 16 | On wooclap in class | Immediately after submission |
| Class Test | 25% | Monday 27 November  (Week 11) | In class | Within 3 weeks |
| Group Poster | 40% | Friday 1 December  (Week 11) | Poster session | Within 3 weeks |
| Portfolio notebook | 15% | Friday 1 December | Upload to Learn | Within 3 weeks |
| Mid-course portfolio reflection? |  |  |  |  |

Contents

[Assessment Information and Deadlines 2](#_Toc141363548)

[Course Calendar 4](#_Toc141363549)

[General Information 5](#_Toc141363550)

[BTO Essential Guide 5](#_Toc141363551)

[Graduate attributes 5](#_Toc141363552)

[Responsibilities – yours and ours 6](#_Toc141363553)

[Course Information 8](#_Toc141363554)

[Course Structure and Description 8](#_Toc141363555)

[Learning Outcomes 9](#_Toc141363556)

[Course Learn Site 9](#_Toc141363557)

[Reading/textbook 9](#_Toc141363558)

[Other course information 9](#_Toc141363559)

[Course Staff 10](#_Toc141363560)

[Assessment 11](#_Toc141363561)

[Course Assessment Details 11](#_Toc141363562)

[Extensions and Special Circumstances 12](#_Toc141363563)

# Course Calendar

Classes will take place in the Larch Lecture Theatre in the Nucleus building at the following times each week:

* Monday 14:10 – 17:00
* Friday 14:10 – 17:00

Week 6 is a “Rest Week” and there will be no classes.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Week** | **Class** | **Day** | **Date** | **Title** |
| **1** | 1 | Monday | 18/09/2023 | Introduction |
|  | 2 | Friday | 22/09/2023 | Reading in tricky files |
| **2** | 3 | Monday | 25/09/2023 | Filtering, cleaning and rearranging data |
|  | 4 | Friday | 29/09/2023 | Tidying horrible data |
| **3** | 5 | Monday | 02/10/2023 | Appropriate plots |
|  | 6 | Friday | 06/10/2023 | Different plot types |
| **4** | 7 | Monday | 09/10/2023 | Introduction to project data sets |
|  | 8 | Friday | 13/10/2023 | Hypothesis generation |
| **5** | 9 | Monday | 16/10/2023 | Joining data sets |
|  | 10 | Friday | 20/10/2023 | Working with your data sets |
| **6** |  | Monday | 23/10/2023 | Rest week – no class |
|  |  | Friday | 27/10/2023 | Rest week – no class |
| **7** | 11 | Monday | 30/10/2023 | Population parameters, linear models |
|  | 12 | Friday | 03/11/2023 | Multiple regression, Binary and Poisson data |
| **8** | 13 | Monday | 06/11/2023 | Multivariant analysis |
|  | 14 | Friday | 10/11/2023 | Advanced Linear Regression |
| **9** | 15 | Monday | 13/11/2023 | Clustering |
|  | 16 | Friday | 17/11/2023 | Dimension reduction |
| **10** | 17 | Monday | 20/11/2023 | Analysis survival/ Working with your data sets |
|  | 18 | Friday | 24/11/2023 | Poster Advice/ Working with your data sets |
| **11** | 19 | Monday | 27/11/2023 | Class Test/Poster work |
|  | 20 | Friday | 01/12/2023 | Poster Presentations |

# General Information

## BTO Essential Guide

Refer to the essential guide here for information on:

* Assessment
* Late penalties and extensions
* Special circumstances
* Progression
* Academic Misconduct
* Student Support
* Safety
* Learn
* Student Administration
* Timetabling
* Tier 4 Visas
* University Complaints Procedure

You will find the essential guide as a PDF file online [here](https://www.ed.ac.uk/sites/default/files/atoms/files/essential_guide_2022-2023.pdf).

## Graduate attributes

The University has identified six groups of abilities, listed below, that should be developed as part of the University of Edinburgh training experience, and to enhance your employability as a graduate. These abilities take your skill-base beyond simply academic knowledge and are enhanced at each stage of your degree.

**Knowledge and Understanding:** This course will train you in managing, manipulating and exploring biology data using Python, Pandas and Seaborn. The skills are transferable to dealing with any large datasets and the understanding of data science will allow you to speedily learn other computing approach (such as R), in further years.

**Research and Enquiry:** This course provides an introduction to the process of scientific research and enquiry using large datasets. Support for this is provided in Workshops, which will teach you how to locate and identify primary sources of information (research papers), to cite in your work. This is a vital skill for any scientist. In addition, practical sessions will introduce students to the art of hands-on research and enquiry, and the vital skill of keeping a detailed record of ideas, observations, results and conclusions obtained in the lab.

**Communication:** This is a critical attribute for all scientists – there is no point conducting excellent research if no one else ever hears about it. Students will learn how to communicate thoughts, ideas and discoveries clearly and concisely. Communication skills will be taught in relation to group work and the formal poster presentation.

**Technical and Practical Skills:** Any successful scientist combines intellectual understanding with the technical skills required to carry out research. Ability to analyse data is key to being able to explore biological patterns and tests hypothesis. This course will give you a strong foundation in managing and exploring data which is critical for success in modern bioogy and many other careers.

**Personal Effectiveness:** At University, every student is responsible for their own learning. The job of staff is to provide teaching materials in the form of lectures, practicals, Workshops and online resources; however it is the responsibility of the student to convert these into an understanding of the topics covered. Hence you will need to develop skills in recording information in a way that captures the key points, and will be useful for you at a later date. Workshops and practicals will help you with this. For all University courses you will need to manage your time effectively, which means first getting to your classes, and second being awake enough to get something from them when you are there. Time management will also be crucial as your assignment deadline approaches – to do yourself justice with an assignment like this, it is vital to give yourself sufficient time to complete and proofread your paper, as those written at the last minute never do the writer justice. You are at this University because you have shown tremendous potential, and good time management is one way to ensure that your potential is realised, both at University and in whatever you choose to do afterwards.

**Personal and Intellectual Autonomy:** The single most important skill for a scientist, or anyone working in a field connected to science, is critical thinking. Hence while learning established facts and theories is an essential part of your training, you must also begin to learn how to test and challenge accepted ideas. Theories such as evolution stand strong precisely because they have survived such challenges. Hence we will encourage you to investigate for yourself whatever you are being taught. If something does not appear to make sense, try to find out why – this will deepen your understanding of the topic. This course provides an opportunity to explore data, develop your own opinions about the patterns you see and most importantly, to test these opinons.

## Responsibilities – yours and ours

The process of learning involves a sharing of responsibilities between you and the College of Science and Engineering. This will probably involve a change in approach compared with your previous experience of education, at school for example. The College will provide you with the facilities you need, and support you, but your learning can only be successful if you accept responsibility for it, and manage your own studies effectively.

The College of Science and Engineering is responsible for providing you with the means to learn. We will:

* Help you to learn, by providing lectures and interactive sessions, laboratories, Workshops and other opportunities for learning.
* Advise you on how to make best use of these learning opportunities.
* Help you to plan your academic career at the University, and to prepare you for your life after university.
* Assess your accomplishment in each of your courses: to give you our view of your achievement; to determine whether you may proceed to the next stage; and (in third and subsequent years) to determine the class of degree you will be awarded.
* Be receptive to your comments on how learning opportunities might be improved.

You are responsible for your own learning. You should:

* Decide for yourself how to make use of the various resources (formal teaching, electronic resources, and textbooks) available to you; this will probably depend on your personal preferences and your own circumstances (for example, work commitments or disability).
* Manage your own time effectively, and prioritise competing demands within and beyond your courses.
* Monitor, and reflect on, your own progress during the course by carrying out the self-assessment exercises that are made available to you.
* Develop a self-awareness about your own levels of engagement and professionalism in relation to your courses.
* Consider how you can best make use of the learning opportunities to attain the learning outcomes of the course, and to meet your personal objectives. The “learning outcomes” are the knowledge and skills you should acquire. These are listed individually for each set of lectures and practicals.

# Course Information

## Course Structure and Description

This course will involve interactive class teaching with coding and data analysis. Students will be taught data science using example datasets, and this learning will be reinforced by application to student-selected datasets from a variety of biological disciplines.

This course aims to give students the skills to work with biological datasets to present, summarise and explore patterns in a wide range of datasets using python, pandas and seaborn. We will take the student through guided examples, supported exploration and on to independent work. For each course topic, student will apply the concepts they have learned to complex research datasets. Student groups will choose a 'real world' dataset to work with at the start of the course, with each being taken from a broad range of biological disciplines to ensure that students can select a topic with which to develop and test hypotheses. The course will be taught in workshops with short lectures introducing the topic then group practice with example and complex datasets using Jupyter notebooks.  
  
Topics covered include:

* Cleaning and organising datasets gathered in field or lab or downloaded;
* Summarise data series using descriptive statistics;
* Explore datasets to look for associations and inter-group differences;
* Use appropriate data summaries and visualisations for the presentation of data analyses;
* Collate information from many groups and discuss data analysis results in a biological context.

### What will this course teach you?

We will build on the python you learnt last year to teach you data management and exploration using pandas and seaborn - two python-based modules. You will learn how to enter, clean, and re-arrange data, how to merge and split datasets, how to produce basic summary statistics, how to use a variety of plots to explore the data, how to define groups and test relationships between them, how to use clustering to look for patterns in large datasets, how to reduce dimensions in large datasets using PCA.

### How will we teach you?

We will provide short introductory lectures and walk you through coding to illustrate the topics or each week, but much of the work will be done in groups working through notebooks online. The notebooks will guide you through learning new techniques and give you scope to explore how these techniques are used with real data to address biological questions. Once you have a good grip of the techniques, you will choose a dataset and work as a group to explore it, producing a poster to describe your results.

## Learning Outcomes

On completion of this course, you will be able to:

1. Understand and critically evaluate data collection methods, appreciating ethical considerations, scientific integrity and sources of bias
2. Write computer scripts to load, filter, visualise, and analyse data sets
3. Use appropriate data science techniques to interpret biological data within a chosen area of modern biology
4. Document and justify analytical workflow and present data-driven biological findings

## Course Learn Site

Learn is the Virtual Learning Environment (VLE), accessed via the MyEd portal [www.myed.ed.ac.uk](http://www.myed.ed.ac.uk). It will be used for providing resources, assessments and content relating the course. Course announcements will be sent from Learn to your University email address and will include reminders of assessments and notifications of changes.

## Reading/textbook

The introductory lecture will give pointers for where to get help, but if you like a printed book these are good:

Biological data exploration with Python, pandas and seaborn: Clean, filter, reshape and visualize complex biological datasets using the scientific Python stack.

Martin Jones (2020)

Python Data Science Handbook: Essential Tools for Working with Data

Jake Vanderplas (2022)

Data Science from Scratch 2e: First Principles with Python

Joel Grus (2019)

## Other course information

*Please be sure to be in the lecture theatre and logged on to LEARN at the beginning of each class. The warm up quizzes open at 2:10 and close at 2:20 (2:23).*

*If you have your own laptop, please bring it along to each class. If not, then there are various options to borrow a university laptop to use in these sessions. For information on how to borrow a laptop see* [*here.*](https://www.ed.ac.uk/information-services/library-museum-gallery/using-library/borrowing-a-book/borrowing-laptops) *There will also be a small number of laptops that can be borrowed in the classes. These will be allocated on a first come first served basis.*

*The course infomation is on LEARN and also on the github:*[*https://github.com/ckidner/DExB2.git*](https://github.com/ckidner/DExB2.git)

## Course Staff

The course team is comprised of the staff members below, along with additional tutors and demonstrators for classes.

| **Name** | **Role** | **Email Address** |
| --- | --- | --- |
| Dr Catherine Kidner | Course Organiser | [catherine.kidner@ed.ac.uk](mailto:catherine.kidner@ed.ac.uk) |
| Dr John Curtis | Deputy Course Organiser | j.curtis@ed.ac.uk |
| Ms Karen Sutherland | Course Administrator | [bio.2A@ed.ac.uk](mailto:bio.2A@ed.ac.uk) |
| Dr Alison Cullinane | Portfolio Director | [Alison.Cullinane@ed.ac.uk](mailto:Alison.Cullinane@ed.ac.uk) |
| Dr Edward Wallace | Course Team Member | [edward.wallace@ed.ac.uk](mailto:edward.wallace@ed.ac.uk) |
| Dr Iain McNae | Course Team Member | [iain.mcnae@ed.ac.uk](mailto:iain.mcnae@ed.ac.uk) |
| Dr Andrea Weisse | Course Team Member | [andrea.weisse@ed.ac.uk](mailto:andrea.weisse@ed.ac.uk) |
| Dr Nick Savill | Course Team Member | [nick.savill@ed.ac.uk](mailto:nick.savill@ed.ac.uk) |
| Dr Abigail Cabrelli | Course Team Member | [abigail.cabrelli@ed.ac.uk](mailto:abigail.cabrelli@ed.ac.uk) |
|  |  |  |

**External Examiner**

The External Examiner has the important role of ensuring that our courses meet appropriate standards and are equivalent to those at other universities of similar status to Edinburgh. The External Examiner will scrutinise examples of the in-course assessments and some of your answer scripts. The External Examiner for Biology 2A is Dr Kerry Franklin from the University of Bristol.

# Assessment

## Course Assessment Details

### Class Quizzes (20%)

These are small quizzes, consisting of five multiple-choice questions run via LEARN at the beginning of classes 2, 3, 4, 5, 6, 9, 10, 12, 13, 14, 15, 16

### The Quiz will be open from 2:10 to 2:20 (2:23 for those with adjustments)

You will need to be in the lecture theatre at 2:10 and ready to work. As there are bound to be a few issues with buses/life getting in the way we will only take the top marks for the 12 quizzes.

### Class Test (25%)

The Class Test will be held in class on Monday of Week 11. It will be completed on Noteable and you should bring your own computer. If you miss (or fail) the class test, you will fail the course.

Marks and feedback for the class test will be available within 3 weeks of submission.

### Portfolio (15%)

 You will work on a professional development portfolio throughout the Bio2A course (the portfolio is also a component of your other Year 2 Biological Science).

You will complete weekly reflective tasks related to course content and on the skills matrix which will support to your professional and personal development. You will also complete a final report which is where you will evidence your learning for course selection for next year. There is an optional interim report that you can submit for feedback. It is essential that you pass the portfolio to pass this course. The portfolio is worth 15% of the course and your portfolio work will be graded. You will need to receive at least 40% to pass. The portfolio will focus on learning for the future and the tasks you complete will be helpful for course choices and future career opportunities.  The final report will be due at 16:00 on 22.11.23 (week 10).

You will write short reports, make regular portfolio submissions, attach supporting evidence, and participate in self-evaluation.

Your portfolio criteria include the following:

·       Conveying an understanding of how your learning has progressed and what areas of development you would like to pursue in the future courses

·       Analyse your skills and abilities (or lacking or developing abilities) with respect to course criteria

·       Reflect on how your skills, abilities and awareness have developed over the course of their research project

·       Support claims with evidence that has been collected over time and added periodically to portfolio tasks

·       Making novel connections between different parts of the course and how you see these connections in future courses.

·       Oversight of your emotional development and attributes.

A detailed overview of the marking criteria is available on Learn, in the assessment folder.

**Data Analysis Project (40%)**

After the project datasets have been introduced in Class 7 (9th October) students will pick which to work with and put their first and second choice on LEARN by 9am 10th October.

The course team will assign students to groups all working on the same dataset, and make the group composition available on LEARN by Friday 13th Oct.

In Class 8 (13th Oct), student will work in their group to each pick a question or hypothesis to address. Each student should have their own question, but themes should be linked as a group poster will be produced. Several possible hypotheses will be suggested by staff, but you can pick your own.

Hypotheses must be uploaded to turnitin on LEARN, along with 2-3 sentences justifying their choice by Monday 16th Oct. They will be approved by staff members. If your hypothesis is not approved feedback will be provided and a new choice must made in the next class.

There is class time for data exploration in classes 8, 10, 17, 18, 19. If you finish the notebook work early in other classes that will free up more time to work on this.

In Class 18 (24th November) Groups will decide which of the analyses to include in a poster. Posters can be worked on in Classes 18 and 19 (after Class test).

Posters will be presented on the 1st December. Location to be confirmed.

Posters count for 40% of the course mark.

Assessment will be 25% by peers and 75% by staff using the rubrics below.

WebPA peer contribution assessment will apply to 25% of the mark.

**Poster rubric for peer assessment**

Use these 7 criteria to rate the poster on a scale of 1-5

(1=strongly disagree; 3=neutral; 5=strongly agree).

Enter the result on the sharepoint excel file here.

1. Display attracts viewer’s attention. 1 2 3 4 5

2. Words are easy to read from an appropriate distance (3-5 feet). 1 2 3 4 5

3. Poster is well organized and easy to follow. 1 2 3 4 5

4 . The question being addressed is stated clearly. 1 2 3 4 5

5. Graphics clearly show the results 1 2 3 4 5

6. Conclusions are stated clearly. 1 2 3 4 5

7. Conclusions are supported by results. 1 2 3 4 5

**Poster rubric for staff assessment**

Use these 7 criteria to rate the poster on a scale of 1-5

(1=strongly disagree; 3=neutral; 5=strongly agree).

Enter the result on the sharepoint excel file here.

1. The topic is well introduced 1 2 3 4 5

2. The hypothesis/question is well justified 1 2 3 4 5

3. The method of analysis is clear and appropriate 1 2 3 4 5

4 . The results are put in context 1 2 3 4 5

5. Conclusions are supported by results 1 2 3 4 5

6. Presenter summarised work well 1 2 3 4 5

7. Presenter dealt well with questions on methods 1 2 3 4 5

8. Presenter dealt well with questions on results 1 2 3 4 5

### Passing the course

To pass the course you must meet the learning objectives by:

* obtaining a course mark of 40% or more overall,
* scoring 40% or more for the Portfolio, and
* scoring 40% or more in the Class Test.

**If you fail the course**, you will have the opportunity to undertake replacement work during the summer in order to pass.

## Extensions and Special Circumstances

**Extensions:**

Extensions are not available for the assessments in this course due to the nature of the submissions.

**Special Circumstances:**

If there are exceptional circumstances outside of your control that mean you miss a number of classes and/or essential pieces of coursework, you can apply for special circumstances through the [Extensions and Special Circumstances](https://www.ed.ac.uk/student-administration/extensions-special-circumstances) service.