

SMI105 Data Visualisation

Dr. Calum Webb, Sheffield Methods Institute

2024-09-24

This module introduces students to data visualisation and has three complementary parts: the theory of data visualisation, the practical doing of it, and a critical awareness of both. The theoretical part involves looking at the principles of data visualisation best practice. The practical part involves learning a programming language to generate your own data visualisations, using the free and open-source software called R. The critical part involves acquiring an awareness of the roles that data and its visualisation play in society, and reflecting on that to inform your practice.

Learning Outcomes: What will you learn?

This module introduces students to data visualisation and has three complementary parts: the theory of data visualisation, the practical doing of it, and a critical awareness of both. The theoretical part involves looking at the principles of data visualisation best practice. The practical part involves learning a programming language to generate your own data visualisations, using the free and open-source software called R. The critical part involves acquiring an awareness of the roles that data and its visualisation play in society, and reflecting on that to inform your practice.

By the end of the module, you'll be able to demonstrate:

- Competence in the design of visual displays of quantitative data;
- Competence to technically produce data visualisations;
- The capacity to integrate data visualisations within a publication;
- An awareness of social and cultural factors that impact different people's experiences of data visualisation.

As students come to this module from a wide range of backgrounds, we don't assume any existing knowledge about the principles or practice of data visualisation or programming. This is fine, and we'll work on this basis.

It is worth noting that this module isn't just an exercise in acquiring programming skills. Understanding the role of data visualisation in society is just as important as being able to generate data visualisations yourself.

Teaching: How will I be taught and what do I need to do each week?

I would recommend setting up recurring events in your Google calendar at the start of term so you know when and where each session is.

| | | Hours |
|----------------------------------|---|-------|
| Lectures & Activities | Each week will include a lecture and learning activities designed to provide an accessible introduction to the topic being covered with a focus on data visualisation principles. There may be recommended reading for each week. | ~10 |
| Data Lab Sessions | After each lecture we will work on a supported exercise in R in a group environment, learning how to apply the data visualisation methods we have learned about in R with Rstudio. | ~15 |

Course outline: What will be covered each week & student responsibilities?

This is a 10-credit module and therefore it is expected it will require 100 study hours per student. Formal teaching hours for this module will take 25 hours (lectures and workshops). Therefore, as a rough guide, this means you should expect to devote around 5 hours of study time per week to this module during a 15-week semester.

You will also be assigned readings and tasks each week that will accompany the lectures and workshops to help you understand the week's material. These tasks are described in the module outline, and in the handouts that will be distributed during the workshops. These handouts will also be posted on Blackboard after the workshops.

| Week | Lecture | Workshop |
|------|--|---|
| 1 | Overview and introduction | Introduction to ggplot2 |
| 2 | Interpreting a data visualisation | Bar charts |
| 3 | Data storytelling | Scatterplots |
| 4 | Chart types | Box plots, density curves, histograms, etc. |
| 5 | Visual design | Customising graphs |
| 6 | Reading week - No lecture | Reading week - No workshop |
| 7 | Publishing a data visualisation | Starting Maps |
| 8 | Maps | More maps |
| 9 | Guest lecture | Data manipulation |
| 10 | Best practice for whom? | What's new in data visualisation? |
| 11 | Wrapping up and assessment preparation | Assessment preparation |

Tasks

Each week, there's a handful of tasks that we'd like you to complete. In general, these consist of some reading, some coding, and for you to keep your eyes open to different visualisations that you find and then reflect on how well you think they work.

You can find details of your tasks each week in the week-by-week plan in the second half of this document. You will see they are listed either as core tasks or supplementary tasks.

- Core tasks: We will draw on these in the following lecture or workshop, so you are required to complete these.
- Supplementary tasks: This is work that is relevant to the following lecture and or workshop so, if you do it, it will help your understanding.

Most weeks we ask that you submit your response to the core weekly tasks on Blackboard. Please do this as it enables the teaching team to give general feedback at the start of the following lecture or workshop.

If you're having difficulty keeping up with the work, please let us know as soon as you can, and we can work together to make sure that everything gets sorted.

Reading and Learning Resources

Knowledge about the theory and practice of data visualisation is based on long-established notions of 'best practice' and, therefore, there are many resources associated with these parts of what we are studying. However, the arguments that call for more critical awareness of the theory and practice of data visualisation are more recent. As a result, resources that put this critical awareness into practice are still emergent. You will find the resources listed in the **Reading and Resource List** grouped in a way that reflects this:

- Practical resources informed by 'best practice'
- The theory behind 'best practice'
- Practical resources informed from a critical perspective
- The theory behind the critical perspective

- Practical resources for working with R
- Keeping up to date

Assessment

This module consists of two assessments, due at midday (12 noon) on the dates stated on the Blackboard Assessment Pages.

The first counts for 30% of your final grade, while the second assessment counts for 70% of your final grade.

The first assessment mainly assesses your ability to use software to draw specific graphs: you'll be given a dataset and a list of graphs to draw. You'll also be asked to reflect on what the graphs you've drawn show.

In the second and final assessment you'll be asked to reflect on a published article that contains a data visualisation, to assess the data visualisation's strengths and weaknesses, and consider how it might be improved. Then, remaining with the same topic, you'll be asked to tell a relevant story by creating your own data visualisation(s), and then explain the decisions you made about which data and which type of visualisation you thought were most appropriate for the audience.

The assessments will all be posted on Blackboard substantially in advance of the relevant deadlines.

Word limits & late penalties

Please note that, where assessments include a word count, the word count is a maximum rather than a target. (We do not use a 10% rule that says it is acceptable to be 10% above or below this word limit). Bibliographies are not included in word counts, but in-text citations are. Coursework must always be submitted online through Blackboard/Turnitin and by no later than 12:00 noon on the day of the deadline. Any unauthorised late submissions after midday on the day of the deadline will incur a penalty of 5%. An additional 5% penalty will be added at noon each working day following the deadline, up to 5 working days late. After that, a mark of zero will be awarded.

Submission

All written coursework must be submitted electronically.

Marked coursework will be returned within 3 working weeks.

The pass mark for this module is 40% overall.

Any change to assessment arrangements will be announced on Blackboard.

For more details on these policies, please refer to the SMI Undergraduate Handbook, which is linked to from this module's Blackboard page, under the heading Undergraduate Handbook.

Reassessment

In the case of students who fail the assessment of this module, repeat assessment will be by a similar set of projects but with different data and source material.

Any reassessment coursework will be due by a date in August 2024 that will be set nearer the time. Further instructions will be communicated to those who fail the module. Resit candidates must consult the module Blackboard website for further information, up until the time of the reassessment, not merely in semester time.

Resit marks will be capped at 40%.

Use of Unfair Means in the Assessment Process

Because coursework, unlike examinations, is not invigilated, the University lays down general rules so that everyone is clear about what is acceptable practice. These rules are set out formally in the regulations for

non-invigilated examinations in Part I of the University Calendar. Further information can be found here: [Unfair Means – Advice to Students](#).

What constitutes unfair means? The basic principle underlying the preparation of any piece of academic work is that the work submitted must be your own work. Plagiarism, submitting bought or commissioned work, double submission (or self-plagiarism), collusion and fabrication of results are not allowed because they violate this principle. Rules about these forms of cheating apply to all assessed and non-assessed work. More details can be found in the student handbook.

What happens if I use unfair means? Any form of unfair means is treated as a serious academic offence and action may be taken under the Discipline Regulations. Where unfair means is found to have been used, the University may impose penalties ranging from awarding a grade of zero for the assignment through to expulsion from the University in extremely serious cases.

Detection of Unfair Means The University subscribes to a national plagiarism detection service called Turnitin which helps academic staff identify the original source of material submitted by students. It is also a resource, which can help tutors to advise students on ways of improving their referencing techniques. Your work will be checked by this service and results will be assessed by the marker.

Generative AI and Unfair Means Usage of generative AI constitutes unfair means when content that was created using generative AI technologies is presented as your own work. This means when generative AI has been used to create a part, or the entirety, of an assessment submission, either,

- a) Without proper or adequate acknowledgement, description, and/or evidence in the event that the use of generative AI is appropriate for the module learning objective, or
- b) Where generative AI has been used to gain an unfair advantage in demonstrating that learning objectives have been met or in meeting the marking criteria for an assessment, unless some application of generative AI as an assistive or enhanceive technology is explicitly stated as a learning outcome in the module handbook.

In the first instance, a good rule to follow is: ‘if you copied output from a generative AI model directly into your assignment, this probably constitutes unfair means’. In addition, if you know there is output from a generative AI tool that you included in your assessment but did not acknowledge, describe, and evidence this in your assessment cover sheet, this would be considered unfair means.

In the second instance, as a general rule, if your use of generative AI is replacing, rather than reinforcing, your learning, this probably constitutes unfair means.

Generative AI outputs should not be used as sources for assessment and you should never cite anything from a generative AI tool. This is because content generated by AI tools is not reliable and is usually non-recoverable and non-reproducible at a later date, so it cannot be retrieved from a link or citation.

Illness, Coursework Extensions and Extenuating Circumstances

SMI’s policy on coursework extensions and extenuating circumstances can be found in the SMI Postgraduate Handbook. If you wish to apply for a coursework extension, you should contact SMI in the first instance (smi@sheffield.ac.uk), in advance of the coursework deadline, and include the relevant Extenuating Circumstances Form plus medical certification if appropriate. Your application will then be considered, and the outcome communicated to you as soon as possible.

Please note that module leaders cannot grant extensions.

Student Feedback

There will be feedback on your understanding of this module during the course of module delivery. Students will receive individual written feedback on each formal assessment.

In addition, the module will use a Piazza discussion page where students can seek feedback on individual issues both from the teaching staff and from fellow students. Full details of this page will be posted on Blackboard.

Feedback on your understanding as the module progresses as well as on all the elements of module assessment can be obtained at any time by using consultation and feedback hours.

Module Evaluation

We always welcome your comments and feedback as the module progresses, and informal evaluation of this kind will take place throughout the module. Students will also be invited to complete an anonymised module evaluation form towards the end of the module. These will be reviewed by SMI's Director of Learning and Teaching and the Module leader. A summary will then be considered by SMI's Staff-Student Committee and the SMI Learning & Teaching Committee. Any changes made to the module as an outcome of student evaluation will be communicated to new/future cohorts of students and in the following year's module Blackboard pages.