

PSYC753 Data Fluency

Ben Whalley and Chris Berry

From the module aims:

The module aims to foster fluency and confidence in the handling, visualisation and communication of quantitative data, alongside skills and techniques for working with larger corpuses of textual and other data. Data visualisation is taught as a foundational technique for exploring and communicating insights from quantitative data. Existing knowledge of linear models is extended with the introduction of the generalised linear model, and a contemporary approach, emphasising prediction and model evaluation is introduced.

In a nutshell: we want to give you the skills to analyse your data as independent researchers, and to give you confidence in working with data which will stand you in good stead in your future careers.

Approach

Psychology students often learn statistics through a “bag of tricks” approach. Workshops might teach how to “do an Anova”, or “how run a multiple regression”. Or you might be given a checklist of things to do when analysing data of a particular type, but without any bigger picture of what we are trying to achieve when we collect and analyse data.

To provide a common thread to our teaching, research methods modules at Plymouth adopt the model for the work of data scientists proposed by Wickham, 2017 (see figure):

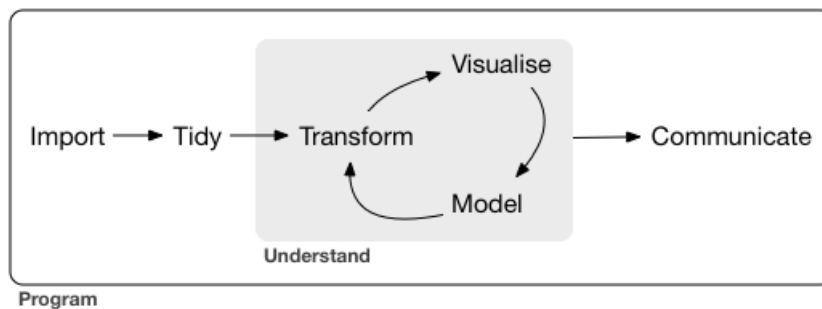


Figure 1: Wickham's model of a data science workflow

In this module we do cover specific skills, but these are embedded within a broader approach to working with data, and integrate it into your own research.

Format of the sessions

We have 16 sessions, which work as follows:

- We avoid extended lectures. This doesn't work well with this subject matter.
- The focus is on learning by doing (this is more like cooking than chemistry).
- In the first hour of each session we will (often) work together.
- In the second hour your work will be self-paced, or in pairs or small groups.
- Activities in the workshops are variable in length, sometimes you will finish early, other times you may be expected to complete the activities outside of class.

In this course we'll learn R as we go, building a little at a time. If some of this is familiar, there will be extension exercises to enhance your knowledge.

The most important thing of all

The most important thing of all is to **practice**.

These materials provide lots of practice tasks. You **NEED** to work through them all to be able to pass the course effectively.

Lab diary/R project archive

It's recommended that you keep a running note of all the work you do in class. This can take the form of a notebook in Word, a blog, or an R script (see first session).

Without a running record of what you have/haven't done it's much harder for teaching staff to help you. The record also allows us to review your progress and make suggestions/improvements.

Access to R

Throughout the module we use R for data processing and analysis.

If you are taking this course at Plymouth University, the easiest way to run the code examples here is to the school's RStudio Server.

- Login to your account on the server [here](#)
- To get an account on the server, or reset a password, contact the Psychology technical office

Why do we use R?

See <https://ajwills72.github.io/rminr/why-r-student.html>