Statistics and software resources

Sophie Lee

Statistics with R

R is a very powerful statistical software with a huge (very helpful!) online community. Because of this, there are lots of free resources to learn R so that you can apply the methods shown in this module.

Here are a list of free resources that are available when getting to grips with R:

- R for data science: comprehensive introduction to R for data analysis, including worked examples and exercises for more practice
- R for the rest of us: A stats-free introduction to R coding (they also have tons of useful tips posted on their blog)
- Introduction to R with Tidyverse: materials from a short course I teach going from absolute beginner to summarising, visualising and reporting data
- Regressions with R: materials from my Regressions with R short course with worked examples of linear and poisson regression

Datacamp courses

Some Datacamp courses covering R and statistics:

- Introduction to Statistics with R
- Introduction to regressions in R
- Hypothesis testing in R
- GLMs in R
- Intermediate regression in R
- Hierarchical and mixed models in R: Where observations are not independent
- GAMs in R: For modelling nonlinear relationships

Statistics with Python

Python is fast becoming one of the most popular coding language and a go-to for data science. Python tends to be used more for machine learning than statistics, but can do everything we covered in the module. It is also a freely available software and has a growing number of free resources available to help learn.

- Python for Data Analysis: Comprehensive ebook introducing the python language
- Python for data science: Python's answer to R4DS, introduction to coding in python

Datacamp courses

- Introduction to statistics with python
- Exploratory data analysis with python
- Introduction to data science with python
- Hypothesis testing in python
- Introduction to regression with python
- Introduction to linear models with python
- GLMs with python