

ASDS Code Camp

Day 5: How to write up and report

Martyn Egan (Teaching Fellow)

September 6 - 10, 2021

Recap

In the last few days we looked at:

- ▶ Some basics of working in R
- ▶ Some basics of working in Python
- ▶ Some general good practices to follow when doing data science
 - ▶ Organising your work
 - ▶ Being consistent in things like object naming

Today's class (the final class!)

How to write up and report

- ▶ For your own projects
- ▶ For doing your class assignments

This class is intended to cover the basics and help orient your future learning.

We will look at:

- ▶ R Markdown
- ▶ Jupyter Notebook
- ▶ \LaTeX

R Markdown

R Markdown is a way of combining your code, its results, and your commentary in a single document. It can then be used to create outputs in a variety of formats, including PDFs, Word files, and HTML documents. R Markdown can be used to produce reports, presentations, dashboards and more.

Activity

Using R Markdown in R Studio is relatively straightforward. We will work on turning our 'diamonds' research from yesterday into a HTML document by creating a new R Markdown file (.Rmd).

Jupyter Notebook

- ▶ Jupyter stands for **J**ulia, **p**ython and **r**, the three languages it was initially designed to work with (although it now supports more)
- ▶ Like R Markdown, it provides a format for combining code, results and commentary in a single document
- ▶ Jupyter runs through your web browser
- ▶ It contains two components: a kernel (to run code) and a dashboard (to navigate files)

Activity

You already have Jupyter Notebook installed as part of your Anaconda configuration. We will work on turning our script file from Wednesday into a notebook.

L^AT_EX

L^AT_EX is a free to use software system for typesetting documents. Unlike word processing software such as MS Word - which uses a 'What You See Is What You Get' (WYSIWYG) interface, L^AT_EX uses plain text and markup language. As with R and Python, L^AT_EX has a steep learning curve. So why use it?

- ▶ It's free ...But so is Libre Office
- ▶ It makes very professional looking documents ...But with a bit of work you can imitate the 'L^AT_EX style' in WYSIWYG word processors
- ▶ It has become ubiquitous in many quantitative academic fields
- ▶ Many data science applications rely on L^AT_EX to create outputs

Why you need L^AT_EX

- ▶ R Markdown uses L^AT_EX 'behind the scenes' to create PDFs. You will therefore need some version of L^AT_EX installed on your system if you want to create PDFs through R/R Studio.
- ▶ Because R Markdown files can be difficult to debug, in Stats I and II we will be using an online version of L^AT_EX called Overleaf (<https://www.overleaf.com>) for your assignments. Overleaf is an online, collaborative L^AT_EX editor.

How do I install L^AT_EX?

- ▶ Go to <https://www.latex-project.org/get/>
- ▶ Follow the instructions for your system
- ▶ Note: you will also need to download an **editor** (a bit like an IDE) to organise your work. There are many free options available. This presentation was written using TeXstudio.

Some useful L^AT_EX links

- ▶ https://www.overleaf.com/learn/latex/Learn_LaTeX_in_30_minutes
- ▶ [https://www.overleaf.com/learn/latex/Free_online_introduction_to_LaTeX_\(part_1\)](https://www.overleaf.com/learn/latex/Free_online_introduction_to_LaTeX_(part_1))
- ▶ <https://latex-tutorial.com/>
- ▶ <http://www.docs.is.ed.ac.uk/skills/documents/3722/3722-2014.pdf>

Including R output into Latex

https://www.overleaf.com/learn/latex/Code_listing

https://en.wikibooks.org/wiki/LaTeX/Source_Code_Listings

Input and format figures from R

https://www.overleaf.com/learn/latex/Inserting_Images

Save figures from R to PDF

<http://www.sthda.com/english/wiki/creating-and-saving-graphs-r-base-graphs>

Any Questions?

This concludes Code Camp. Well done on making it through. If you have any additional queries, don't be afraid to get in touch!