

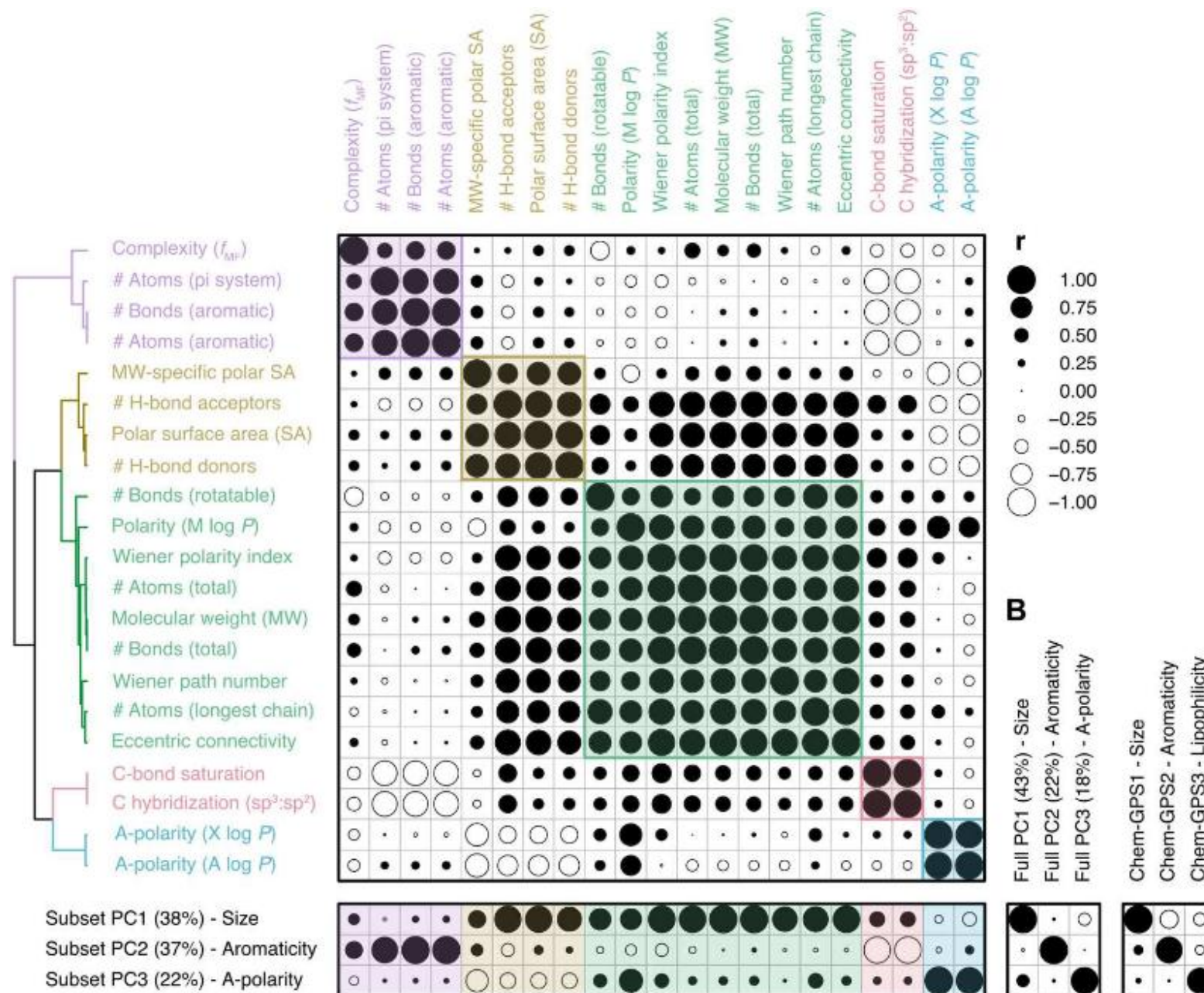


DSA103 Advanced Chemical Data Science

Lecture 21: Data Visualization Recap & Starting Projects

Prof. Meredith C. Schuman

meredithchristine.schuman@uzh.ch



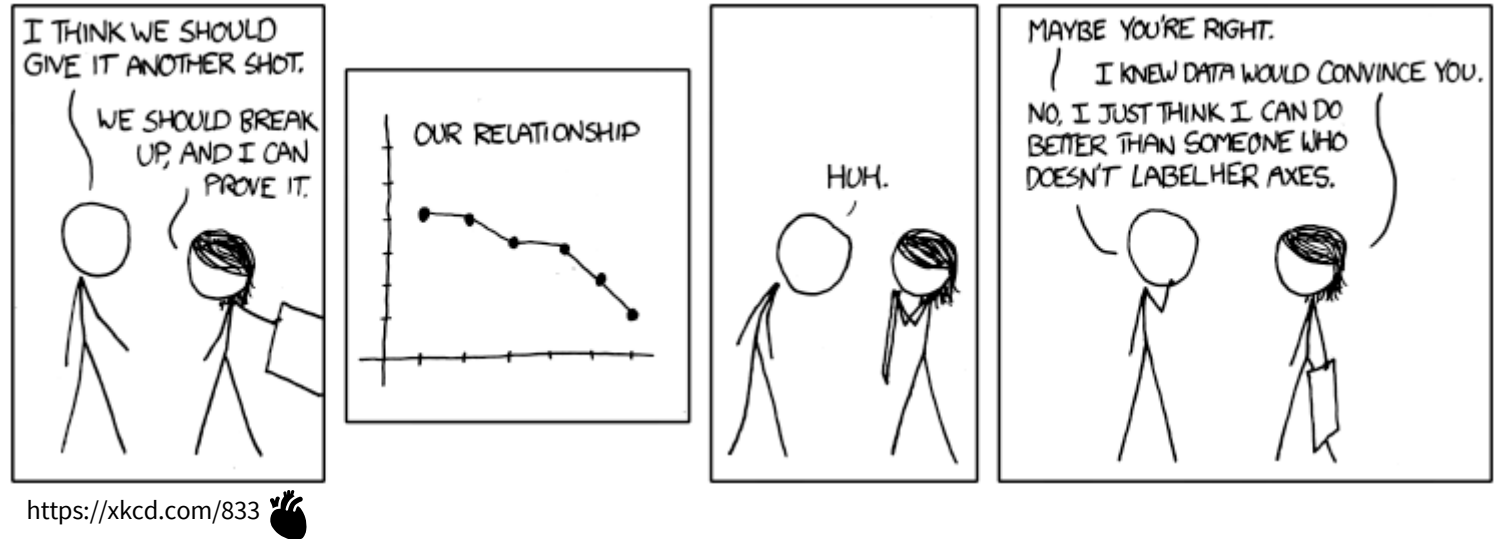
DSA103 revised overview

Block	Topic	Instructor(s)
17.-26.09. (sessions 1-4)	Computational thinking with Python: review and practice	Merry
01.-03.10. (sessions 5-6)	Version control: review and practice	Johannes
08.-10.10. (sessions 7-8)	Exploring data	
15.-17.10. (sessions 9-10)	Generating and reporting on data	
22.-29.10. (sessions 11-13)	Exploring life science data	
31.10. (session 14)	Recap / Q&A	Both
05.-07.11. (sessions 15-16)	Data wrangling	Johannes
12.11. (session 17)	Statistics (review)	
14.-19.11. (sessions 18-19)	Data visualization	
21.11. (session 20)	Recap / Q&A	
26.11-03.12. (sessions 21-23)	Putting it together: small projects	Merry
05.12. (session 24)	Working time (no presence); projects due by end of day (23:59)	
10.12. (session 25)	Outlook: machine learning, vibe coding	
12.12. (session 26)	Mock exam	Both
17.12. (session 27)	Recap / Q&A	
19.12. (session 28)	Exam	

Data visualization review

https://github.com/schoergj/DSA103/tree/main/exercises/lecture_18_visualisation

- Appropriate plot selection
- Units
- Color palette
- Uncertainty indication
- Exporting figures
- Labels, annotations, legends
- Scales
- Multiple axes
- Distributions
- Multi-panel figures
- (Executable publications)



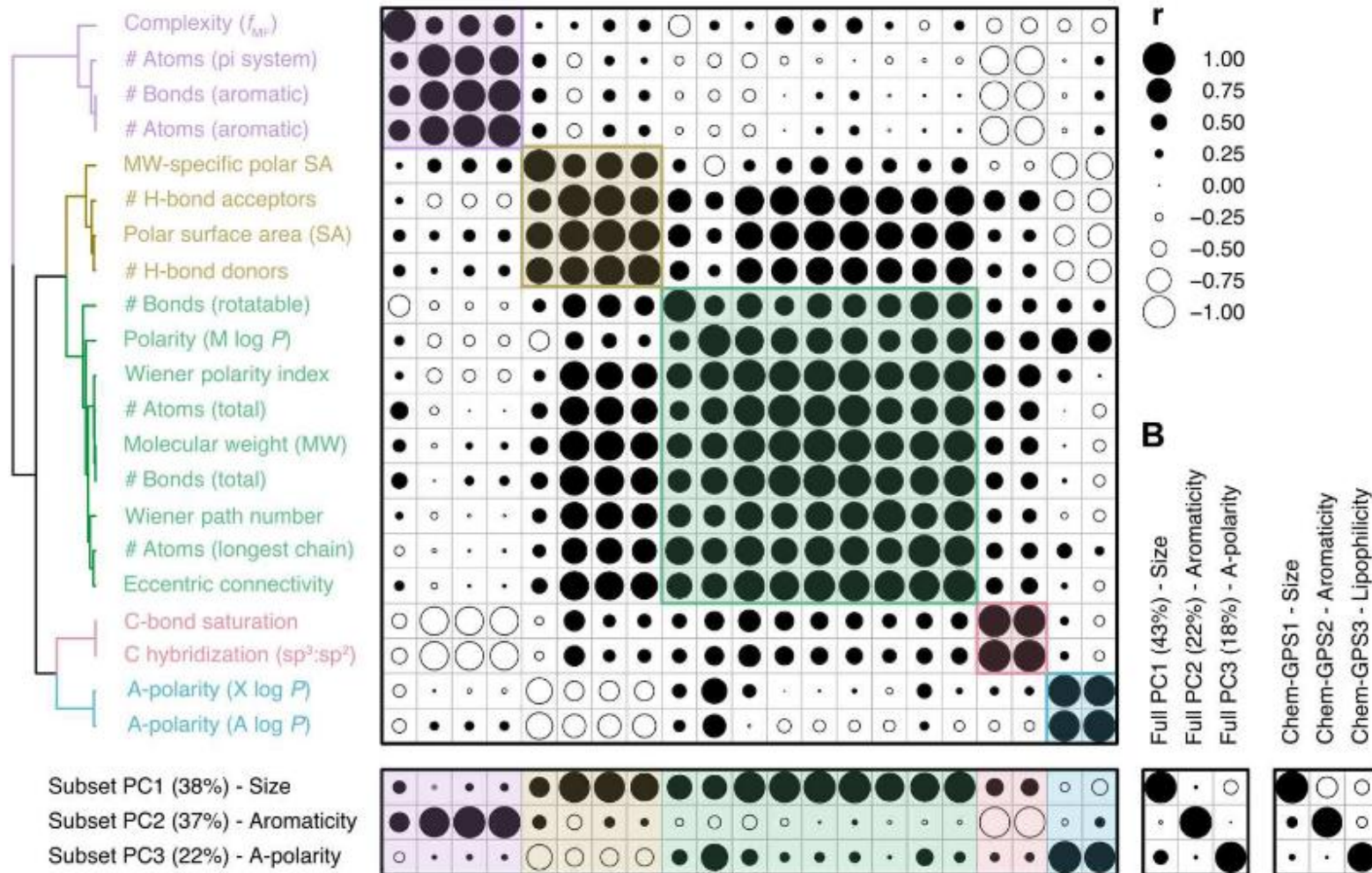
Starting on projects

Make a new repository for your project with the contents of the folder main/projects.

Then:

- Read the paper and summarize key points for yourself: fill in “ReadingScientificPapers”
- Make your documentation and data management plan: fill in “DocumentationDataManagement Template”

I’m here for questions. Bring filled-out documents on Friday! We’ll discuss and start the next steps.



Project assignment

To hand in by 23:59 on Fri Dec 5 by sharing your repository:

- Filled-out sheets “ReadingScientificPapers” (for Walker et al. 2023) and “DocumentationDataManagementTemplate” (for your workflow)
- A one-page data story on your project (I’ll introduce this Friday Nov 28)
- Your code in Python*

*Note that Walker *et al.* has associated source data and code – have a look at it! Their code is in R, so you won’t be able to directly adopt it in Python, but it will help you see exactly what they did.

Starting on projects

Make a new repository for your project with the contents of the folder main/projects.

Then:

- Read the paper and summarize key points for yourself: fill in “ReadingScientificPapers”
- Make your documentation and data management plan: fill in “DocumentationDataManagement Template”

I’m here for questions. Bring filled-out documents on Friday! We’ll discuss and start the next steps.

