I2DS Tools for Data Science Workshop guidelines

Introduction to Data Science, Hertie School, Berlin

November 21, 2022 | 12-19h

Where, when and how will the workshop take place?

- The workshop will take place at the Hertie School on November 21, between 12 and 19h.
- Every session will be allocated a slot of 40mins. There'll be a break of 10mins between two sessions.
- The exact schedule will be announced a few weeks before the workshop.
- You can attend as many other sessions as you want/can, but we expect you to attend at least two other sessions.

What am I supposed to contribute?

1. A pre-recorded presentation that introduces the topic.
2. A live version of the same presentation at the workshop.
3. Practice materials in form of a GitHub repository.
4. A live tutorial session using the practice materials at the workshop.

Where, how and until when do I upload my workshop materials?

- Please upload your workshop video on Moodle (new assignment at the bottom; see description for more details). Note the file size limit of 100MB!
- For the practice materials, please check out the workshop repo README.md template at https://github.com/intro-to-data-science-22-workshop/00-template-munzert-kermit. Use it for your repo and only adapt the contents (not the structure/headings).
- The deadline for both the video and the GitHub repo with materials is November 14, 11.59pm CET (NO extension possible)

How do we run our workshop session?

1. Be there on time (that is, a couple of minutes early).

2. Have your presentation and tutorial materials ready on a laptop. We'll provide a presentation machine but yours will serve as a backup.

3. At the beginning of your session:

    - Welcome the participants
    - Confirm what the session will be about (ideally, the first slide provides that information and is visible before the session starts)

4. Run your presentation (live, not the video!). Try to spend no more than 15 minutes on it.

5. Reserve 5 minutes for a Q&A.

6. Run the live practice/tutorial session. - How exactly you structure this is at your discretion

    - It is important to keep an eye on the timing. You should come to an end when your session time is over so that the team after you can prepare their session

How should our presentation look like?

- The topics vary in breadth and potential depth. I don't expect you to go very deep.
- You should not assume any prior knowledge among your fellow students (other than the topics and tools that have been covered in class).
- Try to offer a quick yet engaging overview: What is this tool/technique/package good for? How can we use it? What are the key features? Where should students go to learn more?
- The absolute maximum is 15 minutes!

How should live tutorial go?

- The tutorial should be structured around the practice material that you provide for the session.
- You don't have to cover everything in the material though. It could make sense to focus on one aspect and let students engage with it / work on it.
- All students will have access to (and hopefully have downloaded) all materials in advance. You can encourage them to follow along on their machines if that makes sense. Be sure to include code that installs the necessary packages though.
- The absolute maximum is 15 minutes!

What are the pre-recorded videos for, and how to produce them?

- The videos are meant to serve as a long-term learning resource for Future You and others.

- The videos will be hosted on Hertie's Vimeo account.
- We will set up a webpage that you can use to find videos + materials for each session.
- Your video and materials will only be hosted publicly if both session creators consent.
- In the workshop repo at https://github.com/intro-to-data-science-22-workshop/workshop-presentations, it says: "Please try to make your presentations using R Markdown." However, this is not a must. If you feel more comfortable with PowerPoint, Pages, or any other presentation tool this is fine, too. It is the content that matters.
- Same goes for recording software. Choose whatever works best for you. We will keep an eye on video and audio quality though.

How are we supposed to divide labor between groups and team members?

- Groups that have been assigned the same topic prepare separate sessions You can exchange ideas but the sessions should work as standalone units.
- It is up to you to split the work in your team. For the live session it would be ideal if both of you attend. It is absolutely mandatory that at least one of you attends and moderates the session.
- For the prepared exercise materials we will pull the repositories that you created on the workshop GitHub page. You should use the README.md of your repo to summarize the content of your session and to document who contributed to which parts of the unit.

How will the sessions be graded?

- Both the recorded talk and the prepared practice materials will be graded. Each component counts towards 50% of the overall grade. The live session itself will not be graded.
- We will consider the following criteria for grading:
    - Presentation video: Is the presented information correct? Does it adequately cover the topic? Is the presentation interesting and engaging? Is it pedagogically/didactically valuable? Is the style appealing? Is the recording quality good?
    - Tutorial materials: Do the materials adequately introduce the topic? Are they interesting and engaging? Do they actually help to learn to use the tools/packages?