

Sprints 1-3 Feedback

Prof GS

SPRINT 1-2

- Good job following the instructions from the video (in terms of code and updates).
 - You mentioned that you will "We will each work on our own branches and merge to a main branch" but that did not happen in Sprint 3.
 - I'm glad you talked about your different preferences and considerations as a team.
-

SPRINT 3

Git/GitHub

- PLEASE do not do any manual updates! The whole point of that video in Sprint 2 was for you to learn about working in teams in git/github and for you to continue those practices throughout all sprints. Moving forward I want to see you all contributing to the repo, making branches, resolving merge conflicts and working as a team.
- Make sure to commit and push more frequently. The idea behind going through sprint 2 is for you to keep using these principles when working as a team via GitHub. I do not see much of that in your git log.
- Make explicit commit messages
- Make sure everyone is working as a team and contributing equally.
- Keep in mind that nothing is lost when you are using git. If you have any concern first go and revise our materials on git (including the videos provided). Secondly you are always welcomed at our office hours.

README

- Go through your README and make it more readable in terms of markdown.
- Your links are not properly formatted for markdown.
- Please upload your dataset to a gdrive and share the link in the README for the next

assignment.

Feedback on questions:

- In the coming years, are women in certain countries more likely to graduate with a STEM major than others?
 - You mention 'Share of STEM Graduates' and 'Country Name' as target variables but they can't both be your target variables. I would target the share of STEM graduates (as a prediction task) and if you have enough data build a model per country or groups of countries.
- What factor(s) influence women's access to STEM fields?
 - This question is a subquestion to the first one.
 - Come up with another target variable for another model.
- Can we expect the proportion of women entering STEM fields to grow in the future? If yes, what factors will influence this growth?
 - Again, this is the same target variable as your first question, therefore a subquestion.
- You need two more questions with two more different target variables.
- One of your questions can be an unsupervised learning question. This could work really well looking at your dataset.

Feedback on Cleaning:

- Overall your cleaning looks thorough.
- Include more explanations on your cleaning steps and considerations. It's hard to follow all the considerations you took.
- Fix warnings in your code for the next sprint.
- What are your plots telling you? Anything interesting you uncovered? Missing interpretations.
- Include references

Steps moving forward

For the next sprint:

- Use GitHub as a team and push every small change.
- Make sure you are branching and following a strategy.
- Make sure everyone in the team is equally contributing to the repo.
- Revisit your team contract and check if your strategies are working. Feel free to change

them if they are not.

- Add more explanations to your notebook.
- Revise your questions
- Meet up with your assigned TA.
- Assigned TA: Meghan