«Exercise Human-Centered Data Science»
Session #3

November 17th, 2020

Agenda

1. A1 + R1 ✔️
2. **A2: Wikipedia Pageview Analysis**
   - (a) What was challenging regarding A2?
   - (b) Take the project of a fellow student and try to run it!
3. *Recap Lecture #2:*
   - *What are important concepts and terminology introduced in lecture #2?*
4. A3: ORES and Bias (next programming assignment)
5. Outlook

A1 + R1 ✓

→ Points for A1/R1 are in Whiteboard

Takeaways

- Setup of environment is difficult (especially with Windows 10 and python 3.9.) for some people.
- Dealing with GitHub, jupyter notebooks was challenging for some people.
- Using GitHub, and git and markdown is new for some people.

A1 + R1 ✓

→ Points for A1/R1 are in Whiteboard

Takeaways

- Setup of environment is difficult (especially with Windows 10 and python 3.9.) for some people.
- Dealing with GitHub, jupyter notebooks was challenging for some people.
- Using GitHub, and git and markdown is new for some people.

Please use the issues:

E.g. What is meant by "repo root"?
This might be an interesting question for more than one person!

A1 + R1 ✓

→ Points for A1/R1 are in Whiteboard

Takeaways

- Setup of environment is difficult (especially with Windows 10 and python 3.9.) for some people.
- Dealing with GitHub, jupyter notebooks was challenging for some people.
- Using GitHub, and git and markdown is new for some people.

?

Please use the issues:

E.g. What is meant by "repo root"?
This might be an interesting question for more than one person!

A2: Wikipedia Pageview Analysis

What was challenging regarding A2?

WHAT?

- How long did it take you? (too easy or too difficult or just too much work?)
- Any challenges or limitations?

A2: Wikipedia Pageview Analysis

What was challenging regarding A2?

WHAT?

- How long did it take you? (too easy or too difficult or just too much work?)
- Any challenges or limitations?

HOW?

- Breakout rooms of 4 people.
- https://pad.spline.inf.fu-berlin.de/3YBuVd3YRT
- One person per group gives a short summary afterwards.

A2: Wikipedia Pageview Analysis

Take the project of a fellow student and try to run it!

- What worked well, what did not work?
- What was difficult to understand from your perspective?
- What advice would you give that other person for the next project?

A2: Wikipedia Pageview Analysis

Take the project of a fellow student and try to run it!

- What worked well, what did not work?
- What was difficult to understand from your perspective?
- What advice would you give that other person for the next project?

HOW?

Breakout rooms of 2 or 3 people.

Find groups first:
https://flinga.fi/s/FAYZXQ6

- 👥 → Put your name here: red = did not hand in yet, green = handed in already
- Take notes for yourself and answer the questions on the left.
- One person per group gives a short summary afterwards.

Recap Lecture #2

optional

What are important concepts and terminology introduced in lecture #2?

- Also besides Reproducibility and Replicability - (R2)
- Please take X Minutes and write down for yourself.
- Round robin.

- Results maybe here: https://flinga.fi/s/FAYZXQ6

A3: ORES and Bias (next programming assignment)

→ **GitHub**

A3: ORES and Bias (next programming assignment)

ORES in short: ORES is a web service and application programming interface that provides real-time predictions on edit quality and article quality [60]. The system was originally developed and deployed in order to make machine prediction technology more available to volunteer tool developers who support quality control processes on Wikipedia. By providing open access to transparently developed machine prediction services, the developers of ORES intend to enable a broader set of stakeholders to become more engaged in developing processes/technology to support quality control work on Wikipedia [30, 31].

Supported by an engineering team at the Wikimedia Foundation [65], ORES has been online since July 2015 and currently supports 42 different languages of Wikipedia (e.g., English, Bengali, Arabic, German, Swahili, etc.). ORES has become the underpinning of an entire suite of tools to support quality control, newcomer socialization, and task routing [61, 29].

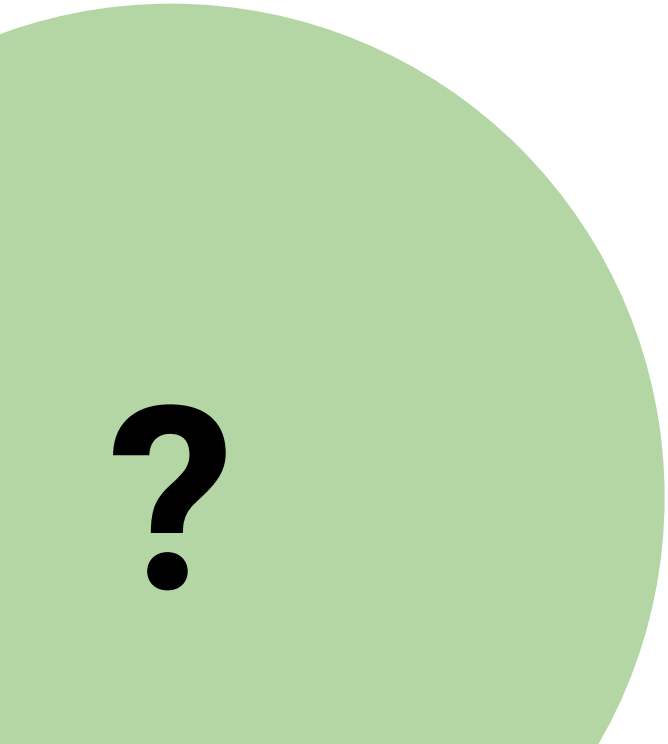
Freie Universität Berlin

Outlook

Recap (yesterdays) BIAS lecture #3

Discuss AMS paper (R3)

Questions



4. Next Assignment