

Oct 11, 2022 Inter-Agency R User Group Meeting notes

Introductions: No new people.

Upcoming events and announcements:

R Govys Meetup: October 20 at 1500. <https://www.meetup.com/rgovys/events/288502391/>

Donna: Abstract submission for the Symposium on Data Science and Statistics (SDSS) is open - <https://ww2.amstat.org/meetings/sdss/2023/submitanabstract.cfm>

Peter Meyer: U of Virginia invites projects for its data science masters program. I did it last year, advising ("sponsoring") such a project and will "pitch" similar projects for this year. The students used R's ranger package. (Details at bottom of these notes.)

R Conference for Gov and Public Sector in December <https://rstats.ai/gov/#about>

Agency Updates

NOAA Fisheries (Eli Holmes and Emily Markowitz)

- Eli Holmes and Em Markowitz taught a class on using GitHub for collaborating in a team on a project. All material online. <https://rverse-tutorials.github.io/WSDS-StLouis/>
- NOAA Fisheries Openscapes fall cohorts started. 160 staff across all our science centers <https://nmfs-openscapes.github.io/>
- quarto_titlepages a Quarto extension for creating custom title pages and covers for Quarto output https://nmfs-opensci.github.io/quarto_titlepages/

BLS (Brandon Kopp):

- Reporting for BLS, Brandon Kopp says they have open slots for data scientists. It took a while to work it out.
- Using <https://www.dataquest.io/> for training
- Data Analyst in R: <https://www.dataquest.io/path/data-analyst-r/> AND Data Scientist in Python: <https://www.dataquest.io/path/data-scientist/>

CBO (Kevin Perese):

- Accessing databases with R or SAS.
- Python learning group. Going through Harvard CS 50 class together
- "Transparency" sessions: covering naming practices, shared vocabulary, version control

FDA (Tomas Drgon)

- Scientific computing days were last month, R and Python were covered

Census Bureau (Yves, Wendy Martinez, Bruce Webster, Chuck Coleman)

- Transitioning to open source when possible. "True North" initiative. Window of transition is 3.5 years
- <https://www.census.gov/newsroom/blogs/research-matters/2022/10/technology-transformation.html>
- Bruce: attended two presentations for True North, one for CenSas and one a week later for R Users Group (Jessica Klein's group) - the later presentation was softer on the timeline, they know they need to coordinate with division chiefs on the timeline
- Other national statistical agencies are making the same transition to open source
- 2023 Data Science Training Program launched 10/3, applications until 10/21 - two tracks; Data Science Generalist and Machine Learning specialty
- Joint Program of Survey Methodology (Census has been a big supporter in the past, but maybe less so in the future(?)) has also recently added a Data Science track as well - program between Maryland and Michigan.

(Just cut/edit the notes below as appropriate. – peter)

The Univ of Virginia School of Data Science invites proposals of data science projects for their masters students to take on for their "capstone" project, analogous to a masters thesis. I "pitched" such projects to students last year and my students did a project related to my research using R's ranger package. It was rewarding. The code should be useful in the future for research on this topic (mapping between classifications of industry and occupation).

I will again pitch such a project, this Thursday evening. If a student accepts, I become a "sponsor," and my agency too. My project had only public data so we didn't need an MOU.

If you are interested in having students consider your topic, there might be time to get something together. The key administrator there is Claudia Scholz, <cws3v@virginia.edu>.

Project often go from Oct to April and the students graduate in May. They are having another call in Feb '23 for projects. I don't know the timing for those cases.

I can help connect people to UVa if you are interested. – meyer.peter@bls.gov

Brandon Kopp reports having similar arrangements with Georgia Tech students.

We discuss: These are NOT internships; that distinction is administratively important. These students do not follow our instructions, we just advise. They don't have agency phone numbers or email addresses or workspace. They are really students. A project *could* involve

confidential agency data or source code, but then an MOU (memorandum of understanding) or contract will be required. UVa has also a contract form of its own for situations like that.