WORLD BANK GROUP

Panel: Practical Steps to Credibility and Transparency in Research: Who Should Do What?

Guadalupe Bedoya

Development Impact Evaluation (DIME)

Team

Amy Dolinger

Luiza Cardoso

Thomas Escande

Camila Ayala

i2i DIME

TRANSFORM DEVELOPMENT

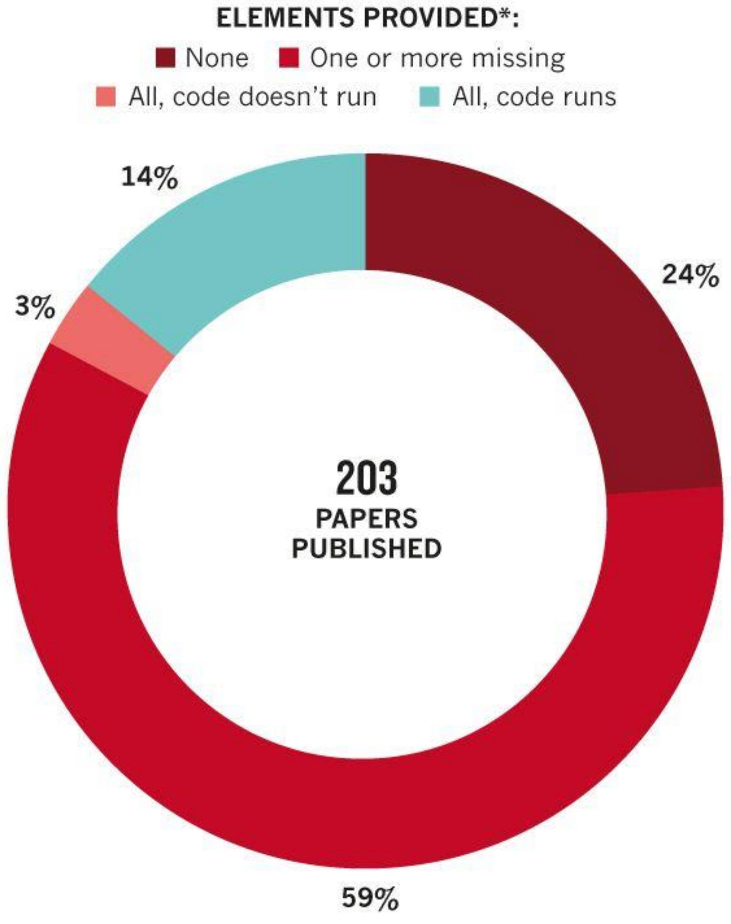
Transparency, Reproducibility and Credibility: A Research Symposium

September 10, 2019

Replication Problem

REPLICATION RARELY POSSIBLE

An analysis of 203 economics papers found that fewer than one in seven supplied the materials needed for replication.



*The elements assessed were raw data, raw code, estimation data and estimation code.

©nature

"We do not think that spending significantly more time would have boosted our success — usually **the code called for data or variables had not been supplied**."

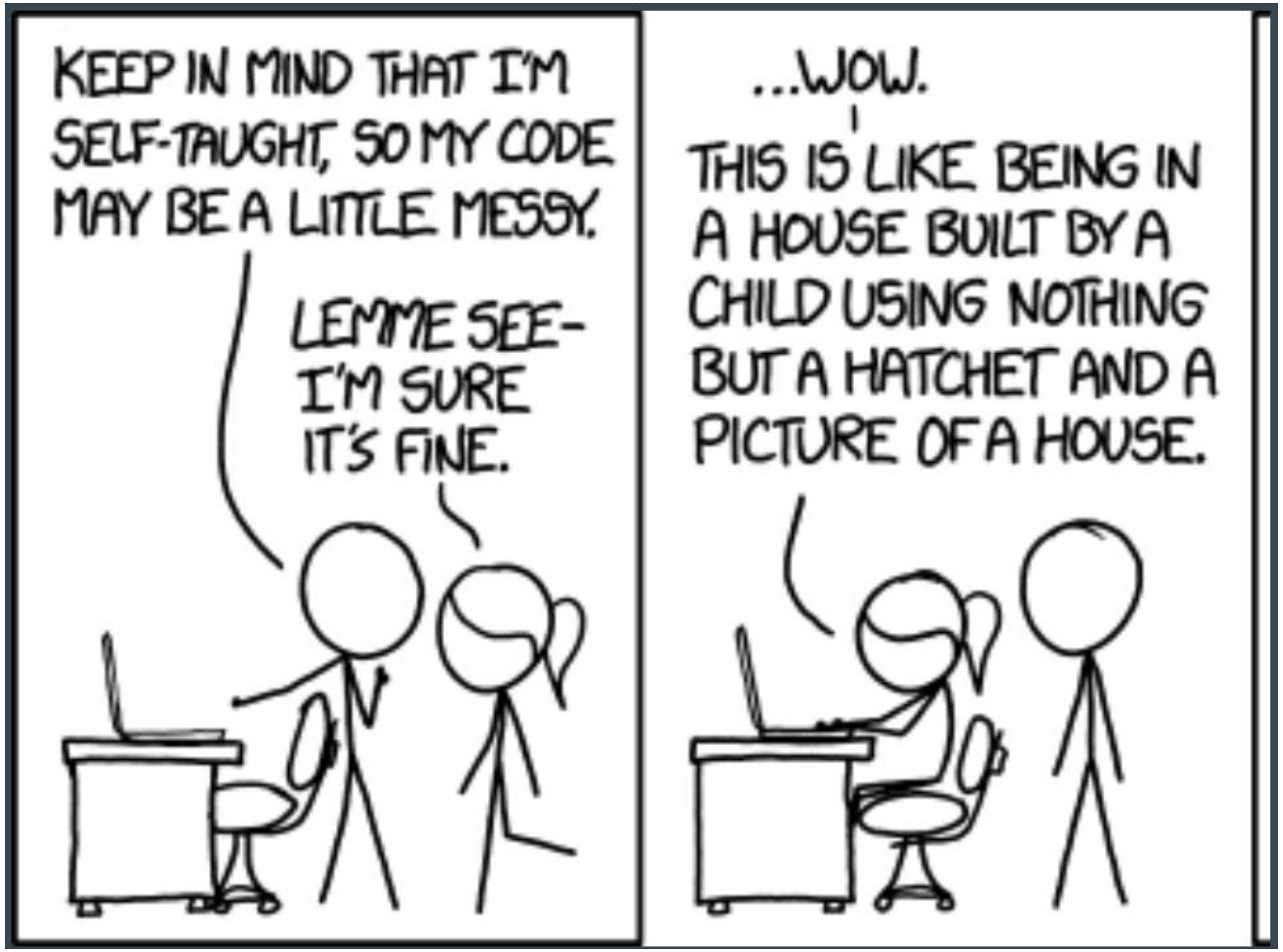
Results consistent with Literature. "A study of the articles published in the *Journal of Money, Credit and Banking* found that only 35.7% of articles met data-archive requirements, and **only 20% of studies could be replicated using the information available in the archive**"

Source: Gertler et al. 2017

Source: https://www.nature.com/articles/d41586-018-02108-9

Focus: Code and data analysis practices

- An immense amount of work is required **to write and document code that is replicable**.
- Many best practices already developed by **software engineers/computer scientists.**
- Few economists or other empirical researchers have **formal training in the science of writing code**.


KEEP IN MIND THAT I'M SELF-TAUGHT, SO MY CODE MAY BE A LITTLE MESSY.
LEMME SEE- I'M SURE IT'S FINE.
...WOW.
THIS IS LIKE BEING IN A HOUSE BUILT BY A CHILD USING NOTHING BUT A HATCHET AND A PICTURE OF A HOUSE.

Source: http://geocenter.github.io/StataTraining/part2/

Main Questions

- What are some of the coding and data analysis practices used by researchers and how they compare with (select) “best” practices?
- What are some of the main constraints researchers identify to adopt these practices, and potential options to address these constraints?

Inputs: survey and learning process

Survey on select practices

- Principal investigators in top development economics organizations (IDEAs ranking).*
- Research assistants (DIME subsample)

DIME analytics

- Learning to implement these practices with internal support from DIME Analytics.
- Public materials available on the DIME wiki – Reproducible Research.

*Notes: *The top 15 institutions were included in the PI survey, except NBER due to duplications. For the World Bank, only DECRG and DIME were included. For universities and think thanks, the email survey was sent to professors/researchers regardless of whether they did development economics or empirical research (some replies indicate that targeted respondents thought it was on for development economists due to the presentation of the survey).*

Survey Statistics (PI and RA samples)

Survey PI profile – all PI samples

Indicator	Unit	Value
Experience conducting empirical work, simulations, or experimental work	Years	14.4
Coauthors in current projects (average per project)	No	2.7
RAs in current projects (total)	No	2.5
RAs per project	No	1.2
Projects in data analysis/journal submission phase in last 2 years	No	4.4
PI does most of the coding	%	42%
PI often reviews the code		27%
PI sometimes reviews the code		20%
PI never reviews the code		6%
Total PI respondents	No	99

Total DIME RA respondents	**No**	**25**

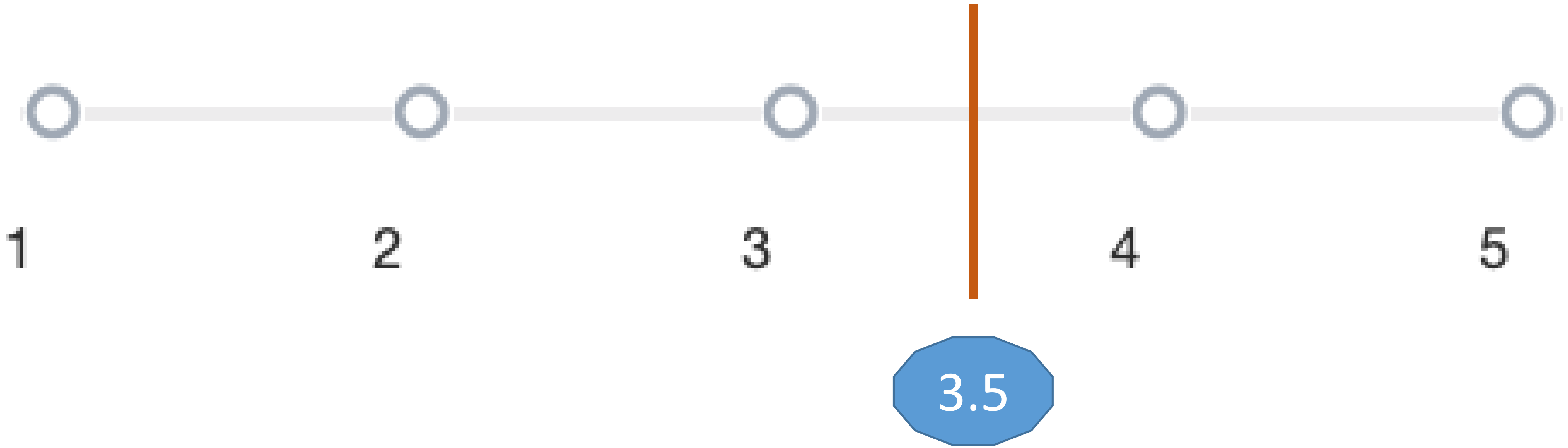
<u>Response rates</u>

Total: 15% / DIME PI: 71% / DIME RA: 63%

How prepared do you think you and your teams are for complying with new AEA Data Policy to improve reproducibility and transparency?

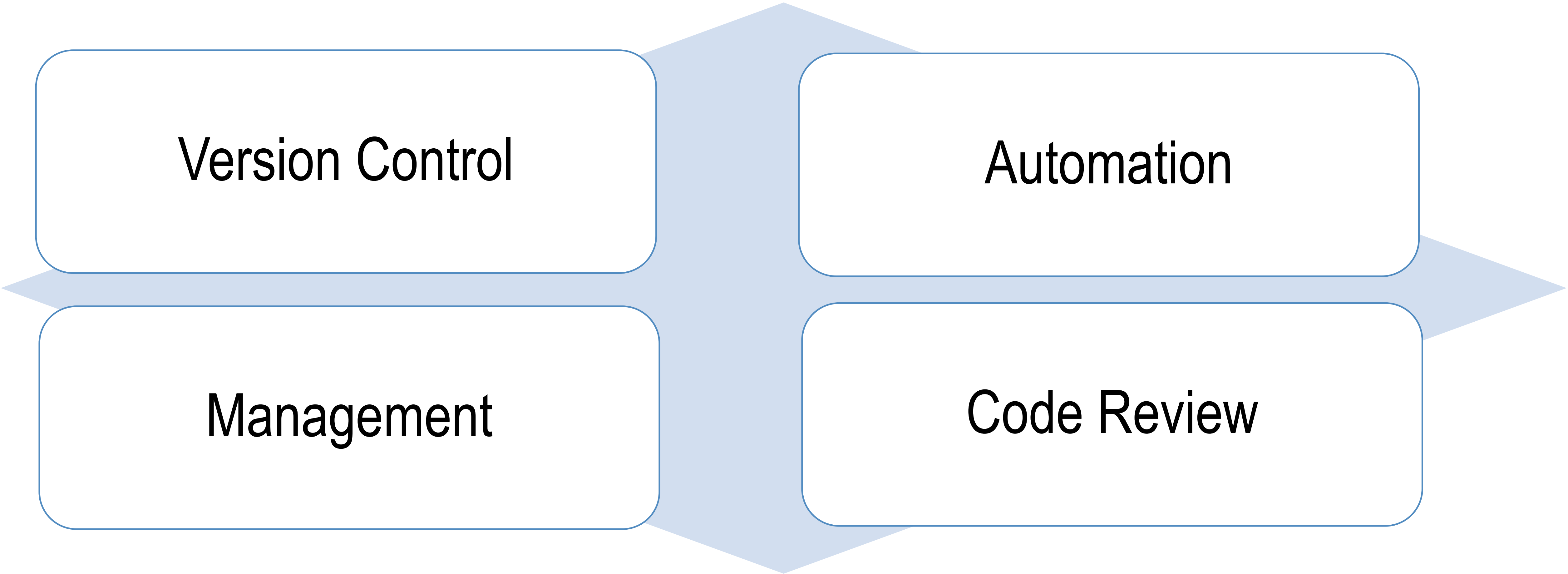
(scale of 1 to 5, 5 being most prepared)

PI Results



Select code and data analysis practices

There is not ONE preferred system or tool, but these few practices are related to concepts that are well-known to facilitate reproducibility and reduce the probability of errors (and may also increase productivity)



Version Control

Automation

Management

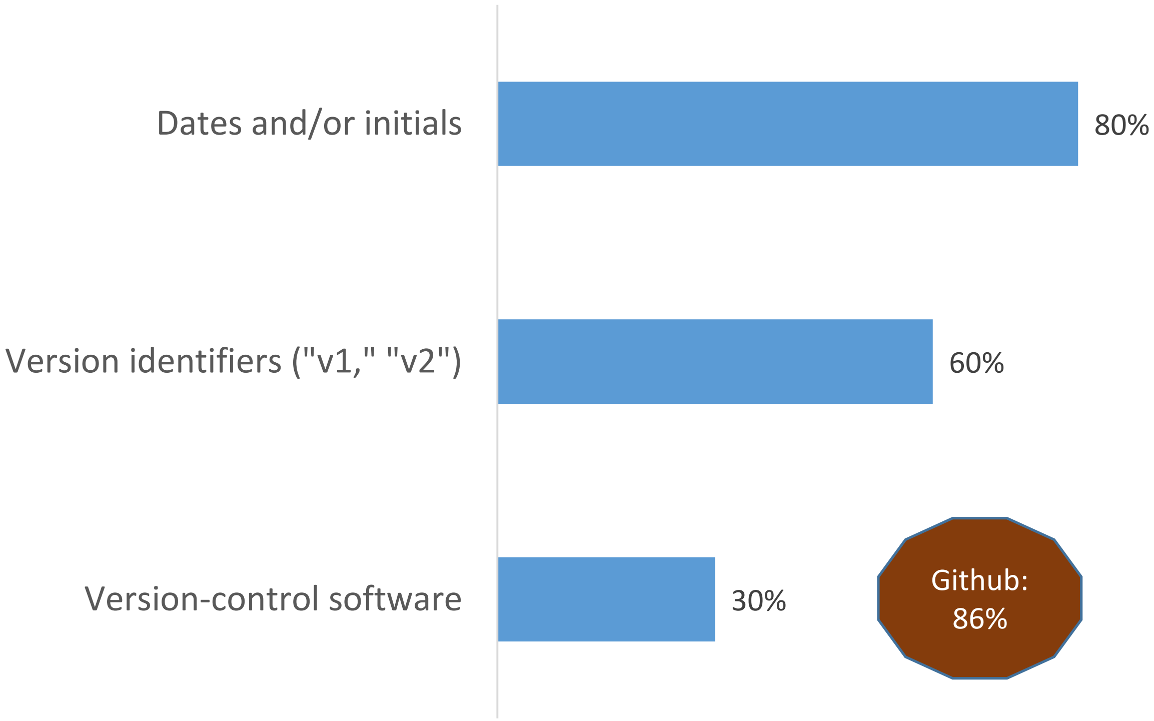
Code Review

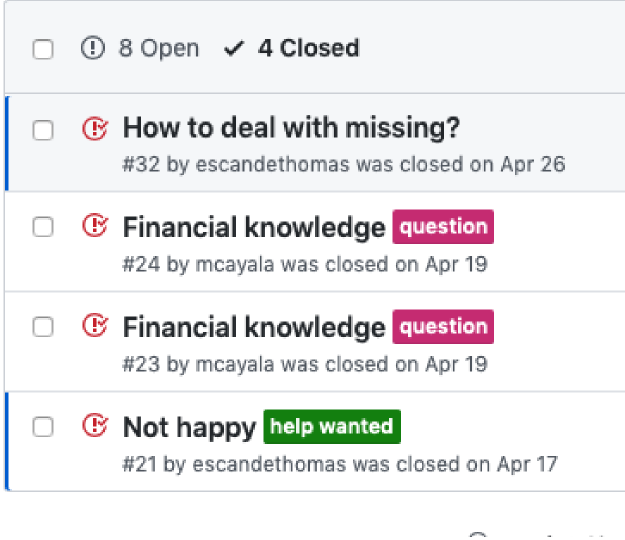
Notes: These practices were selected based on [training and materials from DIME analytics]() and also benefited greatly from [”Code and Data for the Social Science: A Practitioner's Guide” by Matthew Gentzkow and Jesse Shapiro]().

Version Control

What strategies/tools do you use to track different versions of your data, code or output files?

Percentage of PIs who use tool/practice





□ ⓘ 8 Open ✓ 4 Closed

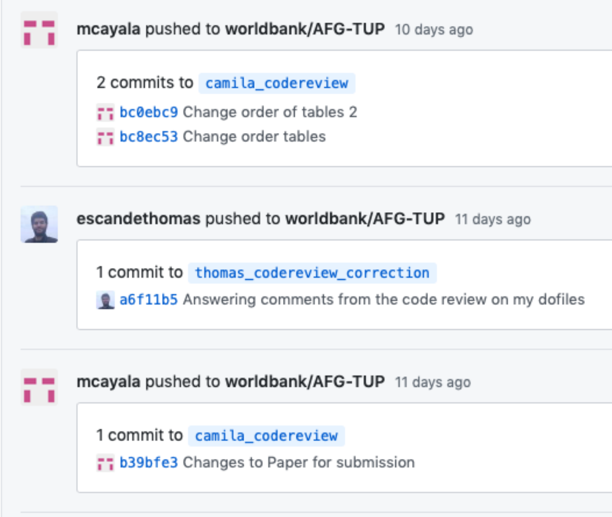
□ How to deal with missing?
#32 by escandethomas was closed on Apr 26

□ Financial knowledge question
#24 by mcayala was closed on Apr 19

□ Financial knowledge question
#23 by mcayala was closed on Apr 19

□ Not happy help wanted
#21 by escandethomas was closed on Apr 17

GitHub



mcayala pushed to worldbank/AFG-TUP 10 days ago

2 commits to camila_codereview
bc0ebc9 Change order of tables 2
bc8ec53 Change order tables

escandethomas pushed to worldbank/AFG-TUP 11 days ago

1 commit to thomas_codereview_correction
a6f11b5 Answering comments from the code review on my dofiles

mcayala pushed to worldbank/AFG-TUP 11 days ago

1 commit to camila_codereview
b39bfe3 Changes to Paper for submission

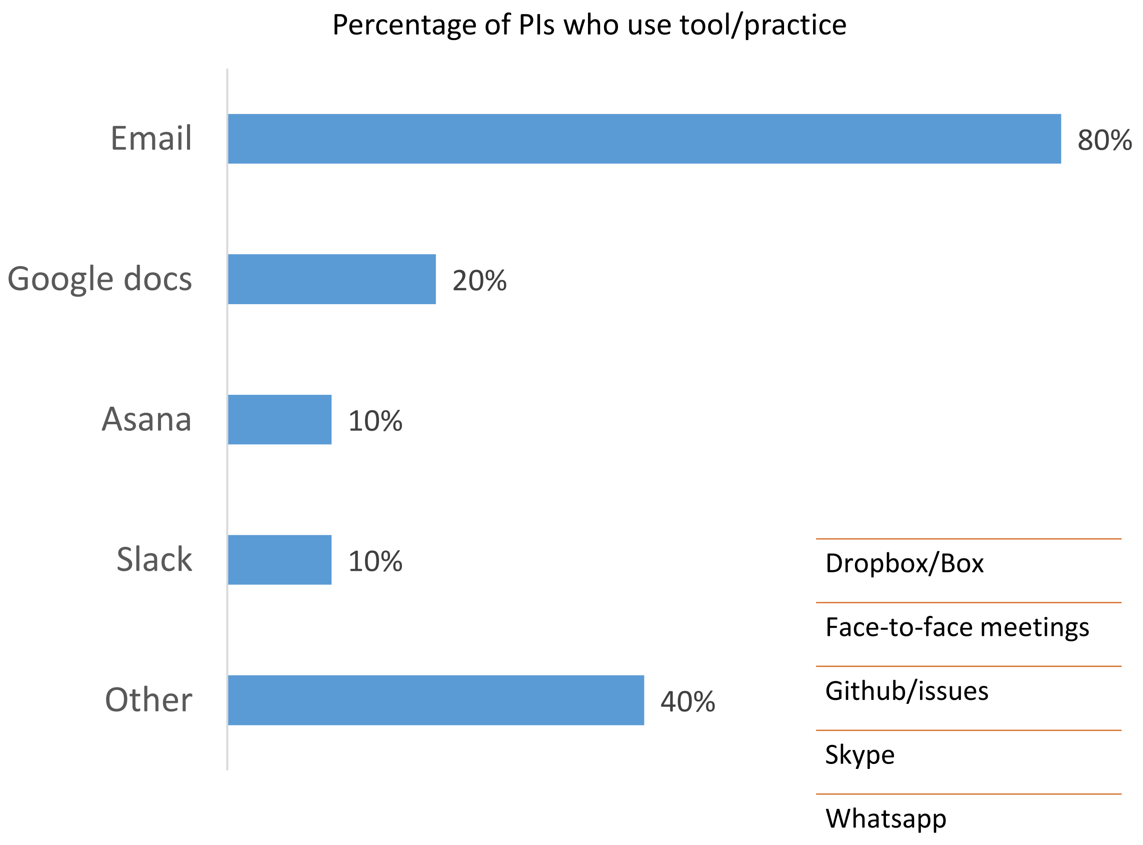
- ✓ Store code and data under version control
- ✓ Run the whole directory before chaining it back in

Gentzkow & Shapiro

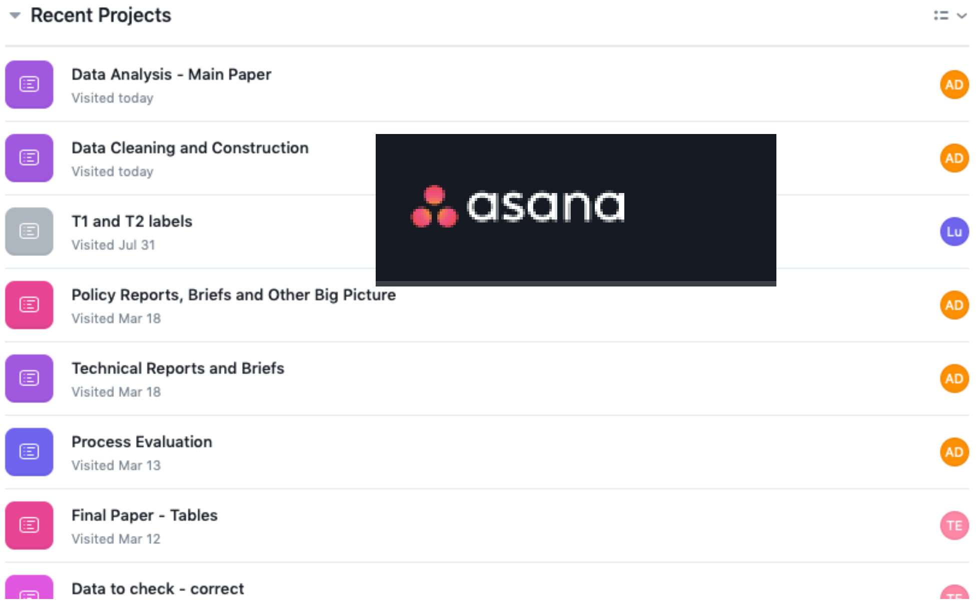
Version-control software: *Systems designed by software engineers to manage changes to files (e.g., documents or programs) in a logical and user-friendly manner, facilitating collaborative work, with the ability to easily recall previous versions later and track changes by every contributor, helping prevent conflicts. Examples of version-control software include Git/Github and Subversion.*

Management of coding and data analysis tasks

How do you/your team commonly manage the tasks for data analysis and coding?
Mark all that apply



- Dropbox/Box
- Face-to-face meetings
- Github/issues
- Skype
- Whatsapp



Recent Projects

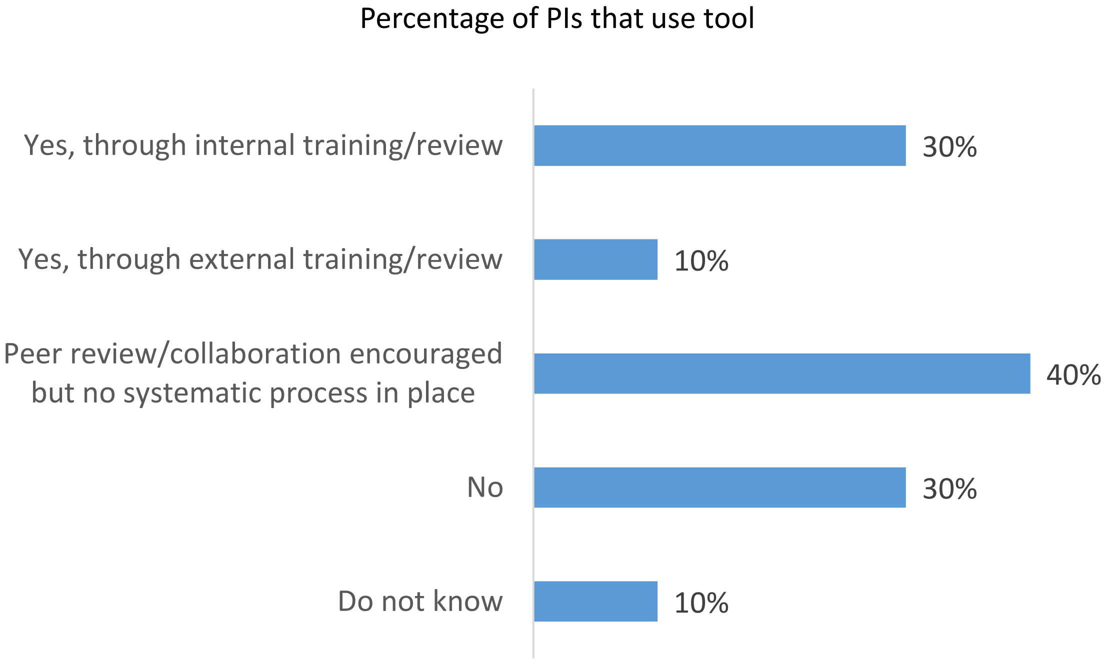
Data Analysis - Main Paper Visited today	AD
Data Cleaning and Construction Visited today	AD
T1 and T2 labels Visited Jul 31	Lu
Policy Reports, Briefs and Other Big Picture Visited Mar 18	AD
Technical Reports and Briefs Visited Mar 18	AD
Process Evaluation Visited Mar 13	AD
Final Paper - Tables Visited Mar 12	TE
Data to check - correct	TE

- ✓ Use a management system to manage tasks
- ✓ Email is not a task management system

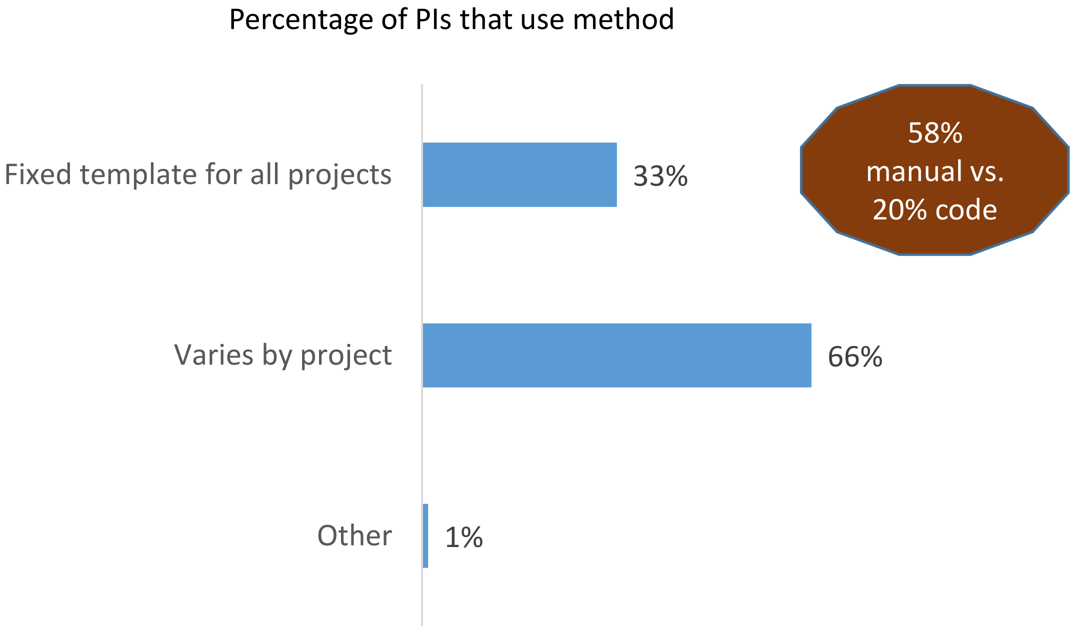
Gentzkow & Shapiro

Automation and abstraction

Is there a process within your team for improving code and coding practices of RAs through better use of abstraction?
Mark all that apply



How is the BASIC structure of the directories defined or modified?
Mark all that apply



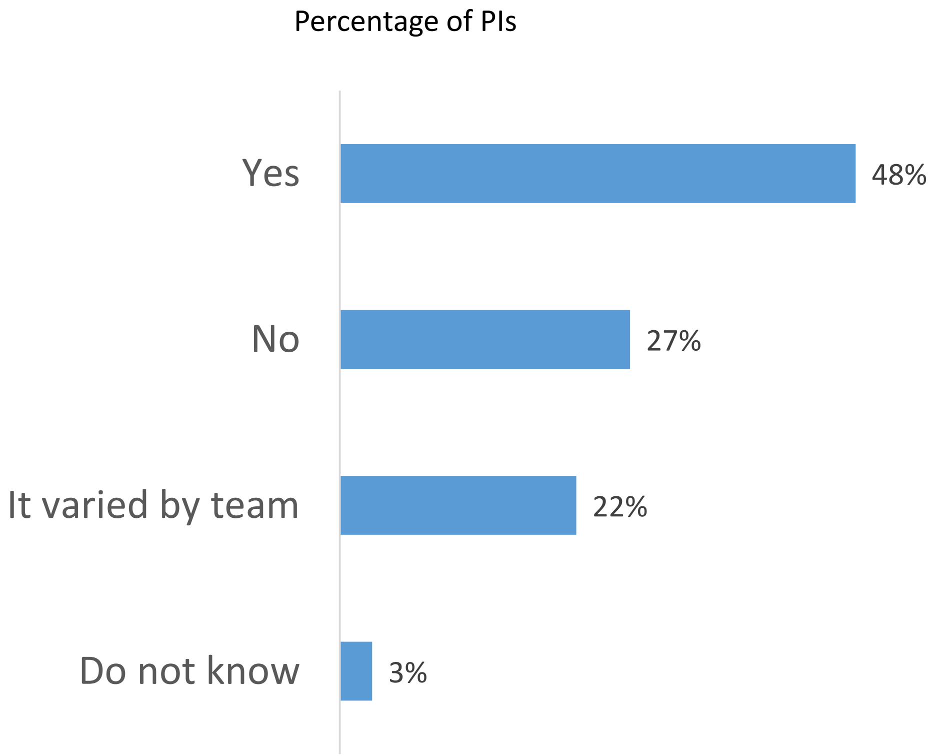
- ✓ Automate everything that can be automated
- ✓ Write a single script that executes all code from beginning to end
- ✓ Abstract to eliminate redundancy and improve clarity

Gentzkow & Shapiro

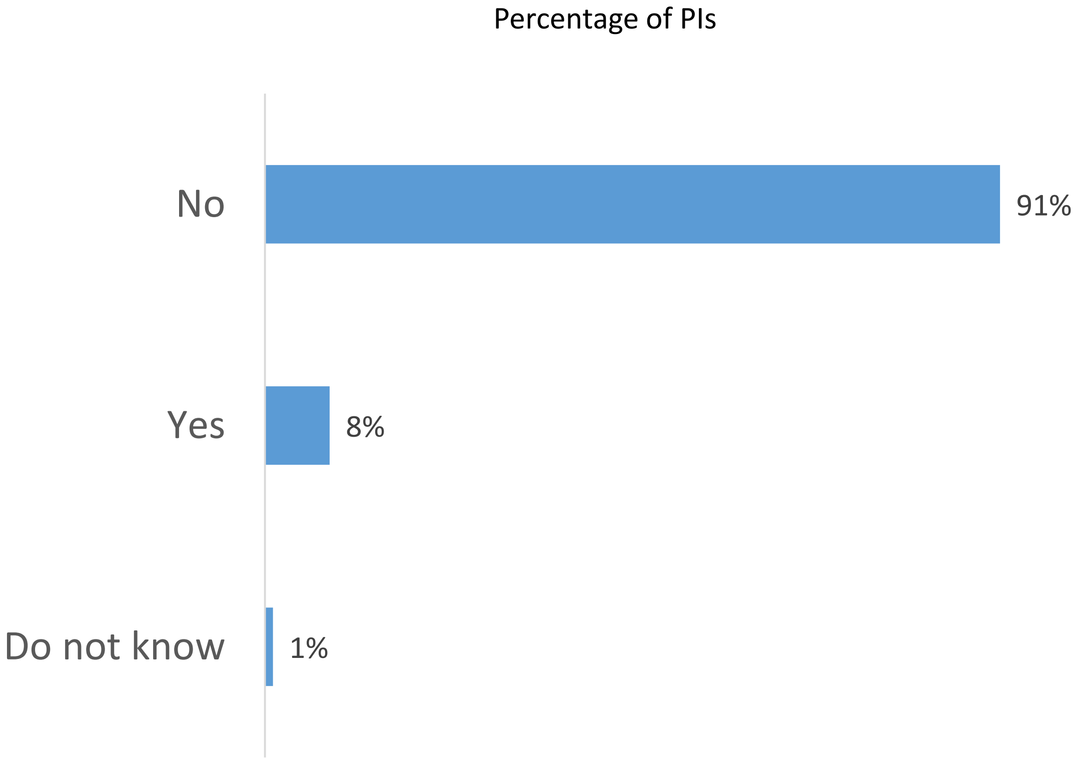
Code automation and abstraction: *Automation is key for writing good code. It allows a higher abstraction level–turning specific instances of something into a general-purpose tool. It helps efficiency and clarity reducing redundancy and errors. Examples include developing user-written commands, programs or loops for repetitive operations, or automating the setup for any team member to open codes and run them, as well as produce outputs (such as using Latex).*

Code Review

Have you performed code review internally within your team (for your last 2 projects)?

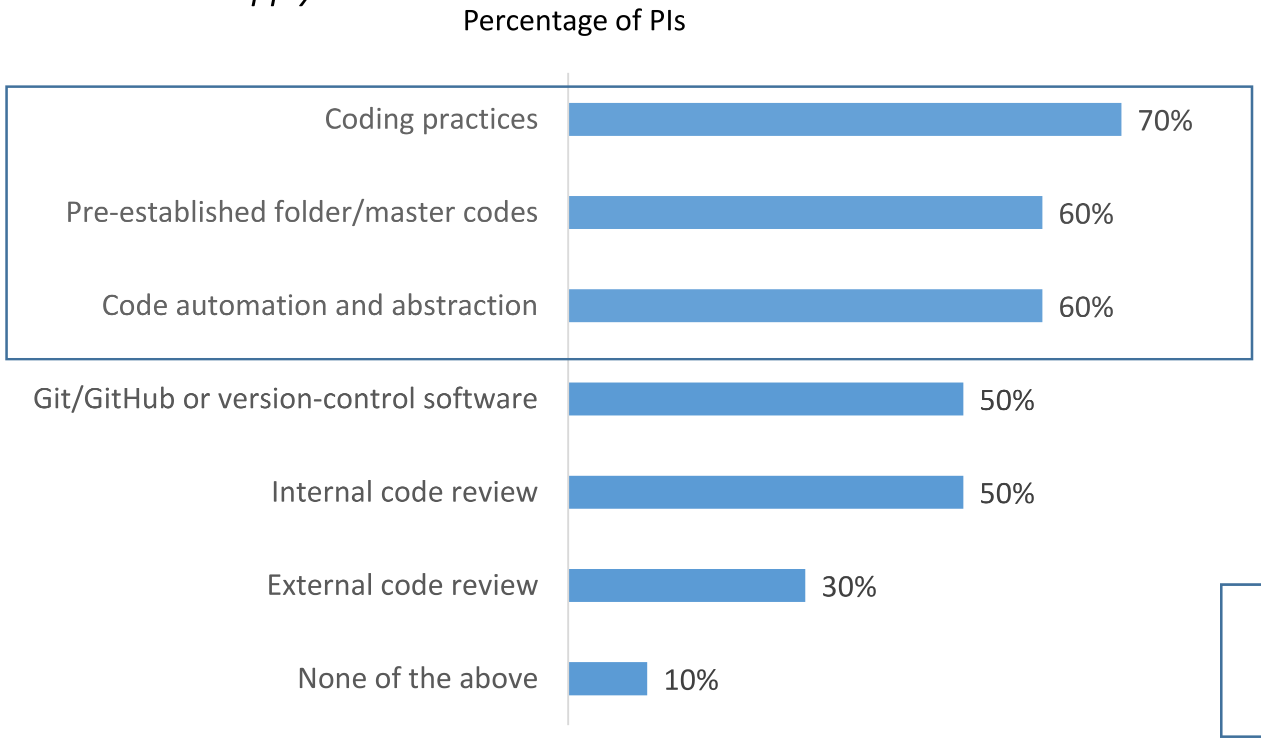


Have you hired or accessed independent code reviewers in your organization or externally (for your last 2 projects)?

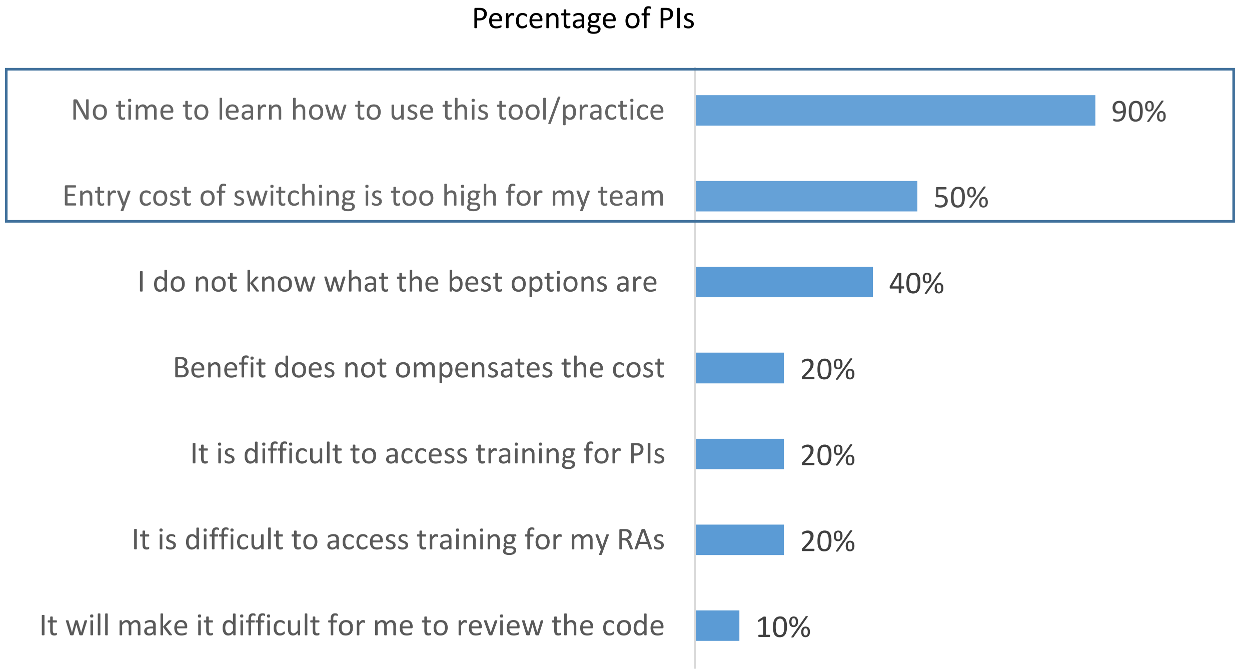


Perceived benefits and constraints

Would you benefit from more/better use of the following software/tools/practices?
Mark all that apply

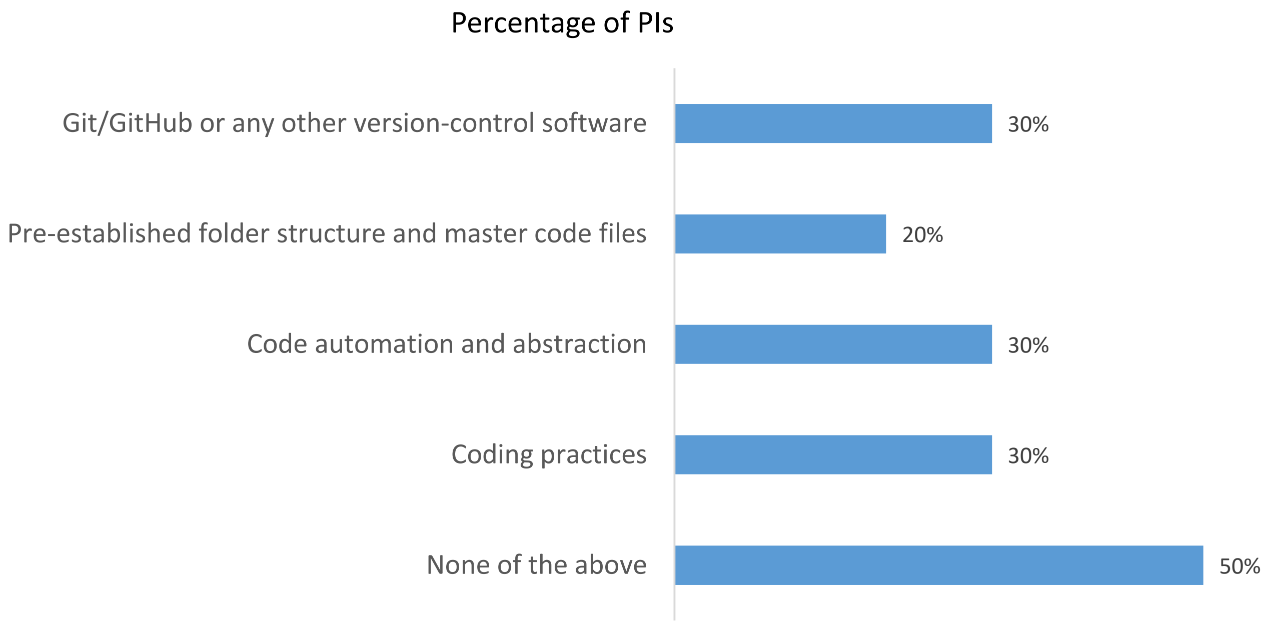


What do you think are the main constraints to increasing the take-up and usage of these tools within your team?
Mark all that apply



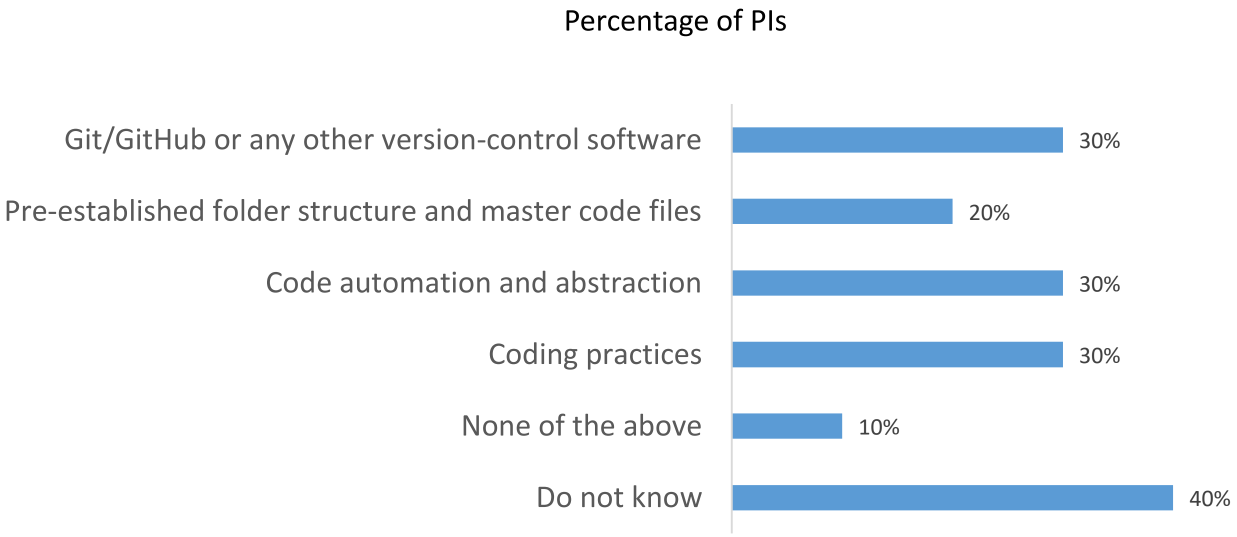
Have you ever taken a training on (or self-taught using guides or other external resources) any of the following software or code management tools?

Mark all that apply



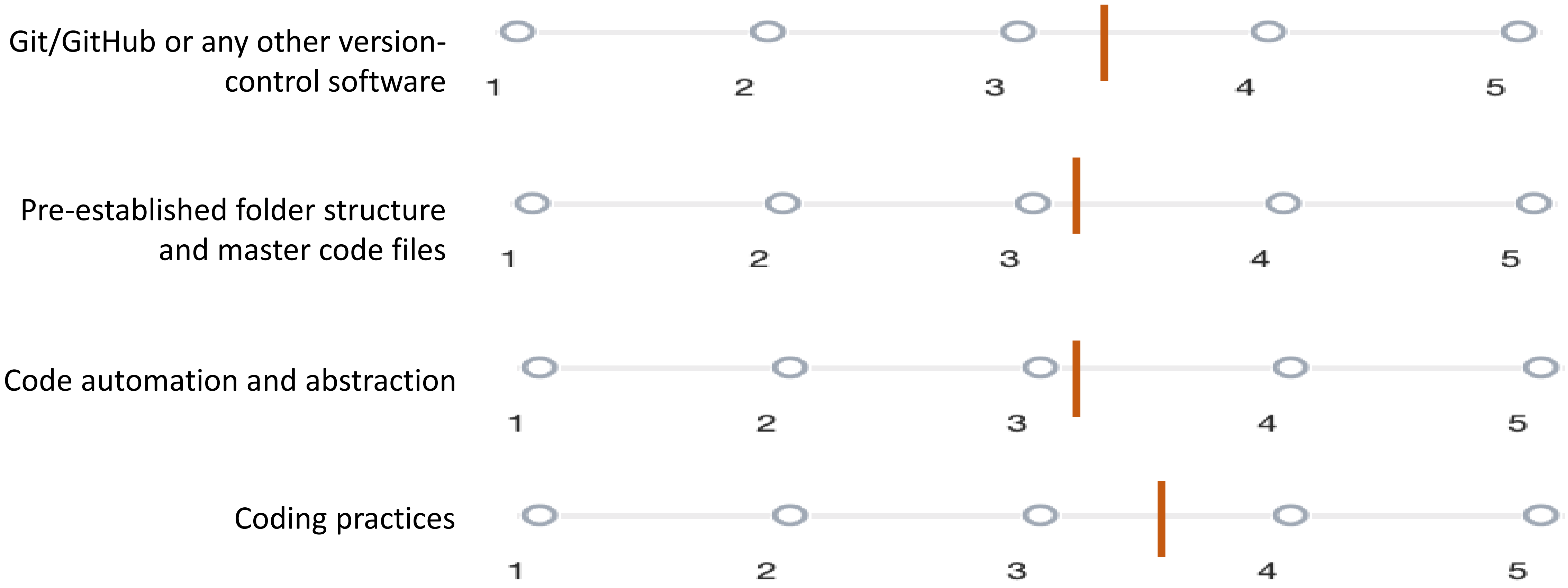
To the best of your knowledge, have any of your RAs/research managers ever taken a training on (or self-taught using guides or other external resources) any of the following software or code management tools?

Mark all that apply



How easily could you find trainings or resources within your organization?

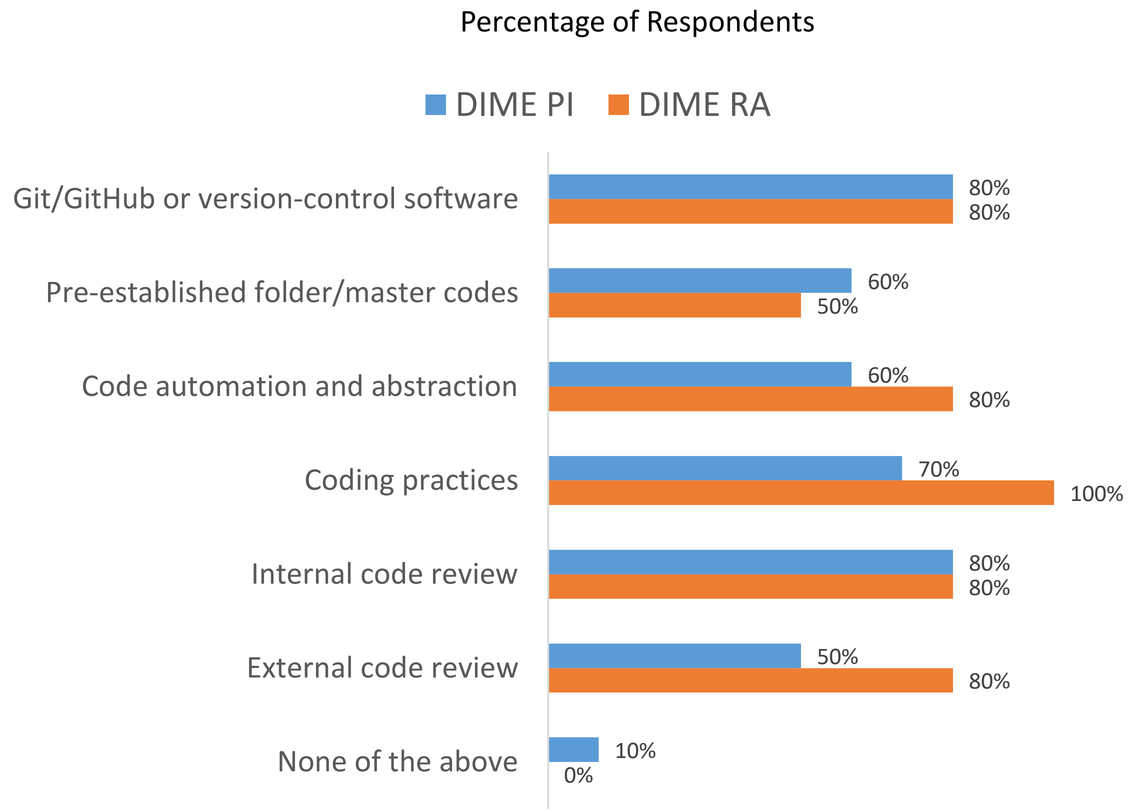
(scale of 1 to 5, 5 being very easy)



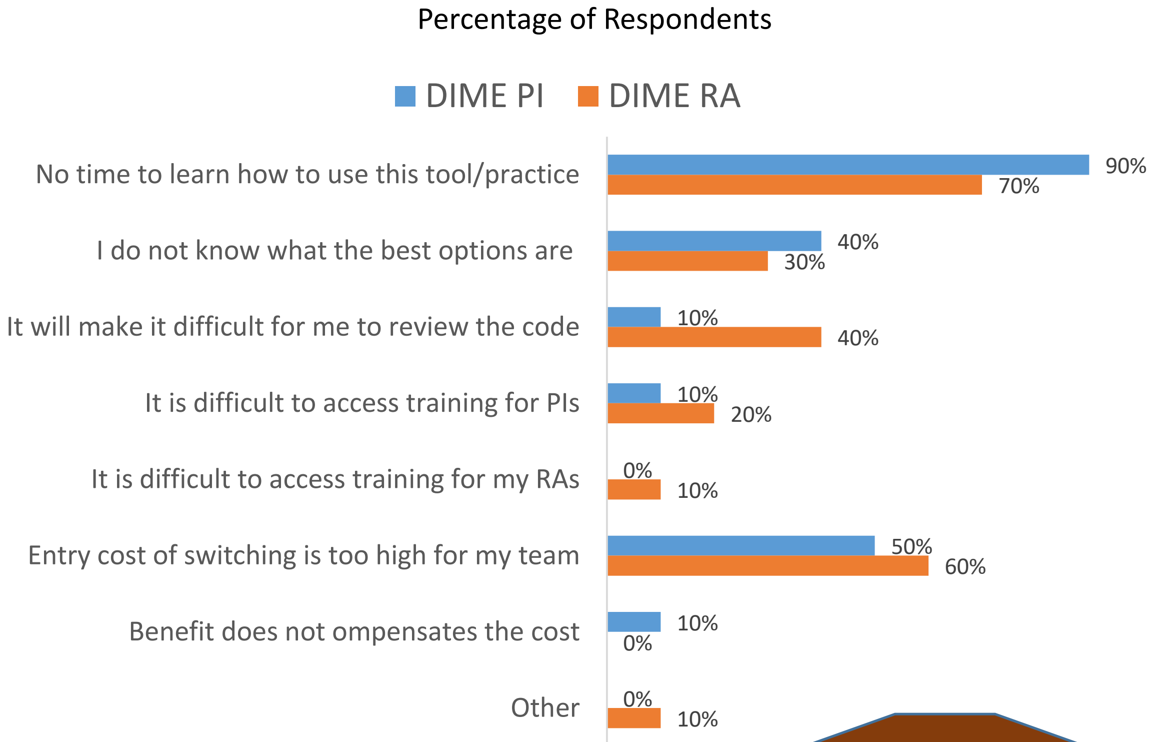
Mean for PIs

Perceived benefits and constraints when training is not a constraint DIME samples

Would you or your team benefit from more/better use of the following software/tools?
Mark all that apply



What do you think are the main constraints to increasing the take-up and usage of these tools within your team?
Mark all that apply



Response rates (DIME)

PI: 71% RA: 63%

70% of RAs did not receive academic training on these

From respondents

“In each project, one co-author is responsible of supervising the work of the RAs. To each RA working in a project, we provide a set of protocols and sample code that they should use. Finally, before publication, we arrange that an external/independent RA review all the steps up to creating the data files for producing results and replicates (without looking at the original code) the results.”

“I have a formula for doing these projects based on my experience as a professional programmer. It does not involve these tools.”

Some thoughts to conclude…

- ✓ Embedding these practices in the process from start can help improve reproducibility and transparency… but also enhances productivity/efficiency.
- ✓ Big entry-cost for adoption, including time to learn and setup the process (protocols and standards are key). But not everybody in the team needs to learn every element.
- ✓ In the long-term, role of academic institutions on state-of-the-art coding/ reproducibility skills for students who are preparing for empirical/simulations/experimental work.
- ✓ Role of other institutions/teams that can provide training and other facilitator tools in the transition (the sponsors of this workshop!)
- ✓ Overcoming barriers to changing habits and norms seems critical.

Thank you!

P.S.: For documentation on the survey and resources on some of the tools and practices listed in the presentation, visit https://github.com/worldbank/ReplicableResearch