

Interview Script

We will start the interview with some general questions about the participants' data analysis process.

- Can you talk about your usual data analysis process?
- What tools/environments do you use?
 - Are there any special cases where you use something else?

Visualization tool users

- Why do you prefer using a *<tool>* vs a scripting environment? (*Only if they express a clear preference for tools vs scripting?*)
- Do you record/track your analysis session?
 - Yes
 - How?
 - What do you capture?
 - Why?
 - Do you revisit the logs often?
 - Do you use the logs during publications?
 - Challenges?
 - No
 - Why?
 - Would you like to? (*Only if they indicate that its lack of means for doing so*)
 - What would be the purpose of such logs?
 - What would you like to capture?
- Do you have certain recurring parts in your analysis? (e.g. data cleaning/wrangling)
 - Do you see any challenges in these steps being automated?
- Do you switch between multiple tools during your analysis sessions (including pen/paper)?
 - Describe?
 - What are the challenges in doing so?

Script users

- Why do you prefer scripting over visualization tools?
- Do you think that certain tasks like selecting complex patterns would be easier to do in a visualization system?
- Would you prefer such interactive visualizations embedded in your current working environment or as separate tools that you can switch to?
- *<Switch to the above questionnaire when relevant>*

Common questions

- Do you have datasets which are updated from time to time, e.g. when new data comes in?

- Do you have to re-run your analysis/scripts when this happens?
- Do you have to modify your scripts/analysis frequently to account for variation in the data? Can you give an example of such a scenario?

After the above questions, I will do an introduction of our project <refer interview_slides.key> and a live demo using the Cluster dataset on both the tool and Google Colab.

- *<General feedback>* What do you think about such a technique to capture and share workflows?
- What did you find confusing during the demo?
- *<To avoid leading questions and refresh a feature in the minds of the user, I will show the feature in tool + notebook and ask for their thoughts on it.>* What are your thoughts on *<feature>* do you see this being adapted to your analysis process?
 - History/provenance
 - Semantic selections
 - Capturing reusable workflows
 - Compare view
 - Review Mechanism
 - Bridging between tools
- Does our approach of creation of workflows match your mental model of creating workflows*<refer to previous answer if appropriate>*?
- If you could capture recurring parts of the analysis as a workflow you can reuse:
 - Would you want to create a workflow from your previous analysis session? OR Would you want to create the workflow from scratch like a recipe?
- Anything else you particularly like or dislike about this approach?