

Supplemental Materials

Defining an Analysis: A Study of Client-Facing Data Scientists

Abigail Mosca , Shannon Robinson , Meredith Clarke, Rebecca Redelmeier, Sebastian Coates, Dylan Cashman , Remco Chang 

Tufts University

1. Interview Script

In our design of the interview script, we seek to learn about data scientists' practices through questions organized into four areas:

- (1) **Characteristics of data scientists:** how data scientists themselves describe their role and what activities they assume,
- (2) **Interactions between data scientists and clients:** what tools and techniques do data scientists employ to interact with clients,
- (3) **Interactions between data scientists and data analysis:** what tasks do data scientists perform for analyses, and
- (4) **Nature of data scientists' workflow:** how do data scientists themselves describe their workflow with clients.

The full interview script provided below is organized by these areas.

For **Characteristics of data scientists** we used the following questions:

- Could you please describe your role to us?
- Can you take us through an example of a recent typical meeting with a client?
 - What obstacles can come up during a meeting?
 - How do you navigate those obstacles?
- What advice would you give to someone interviewing for your role?
- What do you need to be expert in to excel at this role?

For **Interactions between data scientists and clients** interviewees were asked specifically about gathering questions from clients and presenting analytic results to clients. Interview questions addressing this were:

- What are the 5 most frequently asked questions from clients?
- What is the first series of questions you ask a client when starting a new project?
- What deliverables do you give to the client?
- What questions do you ask yourself when deciding what to present to a client?
 - How do you decide what fields in a dataset are of interest to a client?
- How do clients provide feedback?

- What do you do with their feedback?

For **Interactions between data scientists and data analysis** questions were dependent on the interviewee, as some performed analytics themselves while others acted more as a conduit between clients and an analyst or analytic team. Interview questions for this area included:

- How do you interact with data scientists? What role do they play with your clients? (In cases where the interviewee works with an analytic team instead of performing analyses themselves)
- Do you ever need to hand off you analysis to someone else to do the analysis? (If the interviewee does analyses themselves)

Finally, for **Nature of data scientists' workflows** questions aimed to investigate if data scientists view projects with clients as cyclical or linear. Questions addressing this area were:

- How do you know when you have met the clients' needs?
- Do you see your projects with clients as cyclical?
- How do you know when a project is finished?
- How do you determine if a project was successful?

All interviews ended by asking interviewees if they had any additional information they felt might be of use to us, or if they had any questions for the interviewer.