Cultivating a Data Science Learning Community

Matthew Burton, School of Computing & Information
Gesina Phillips, University Library System
Melissa Ratajeski, Health Sciences Library System

Motivation

- There is a growing demand from faculty, postdocs, students, and staff, for informal data science training outside of disciplinary curricula
- Demand for data science training has outpaced current capacity (both formal and informal)

Context

- Research and teaching across the disciplines is increasingly data and computationally intensive
- While the data and computational needs in the humanities, social sciences, natural sciences, and health sciences might differ, they all need *basic data skills*
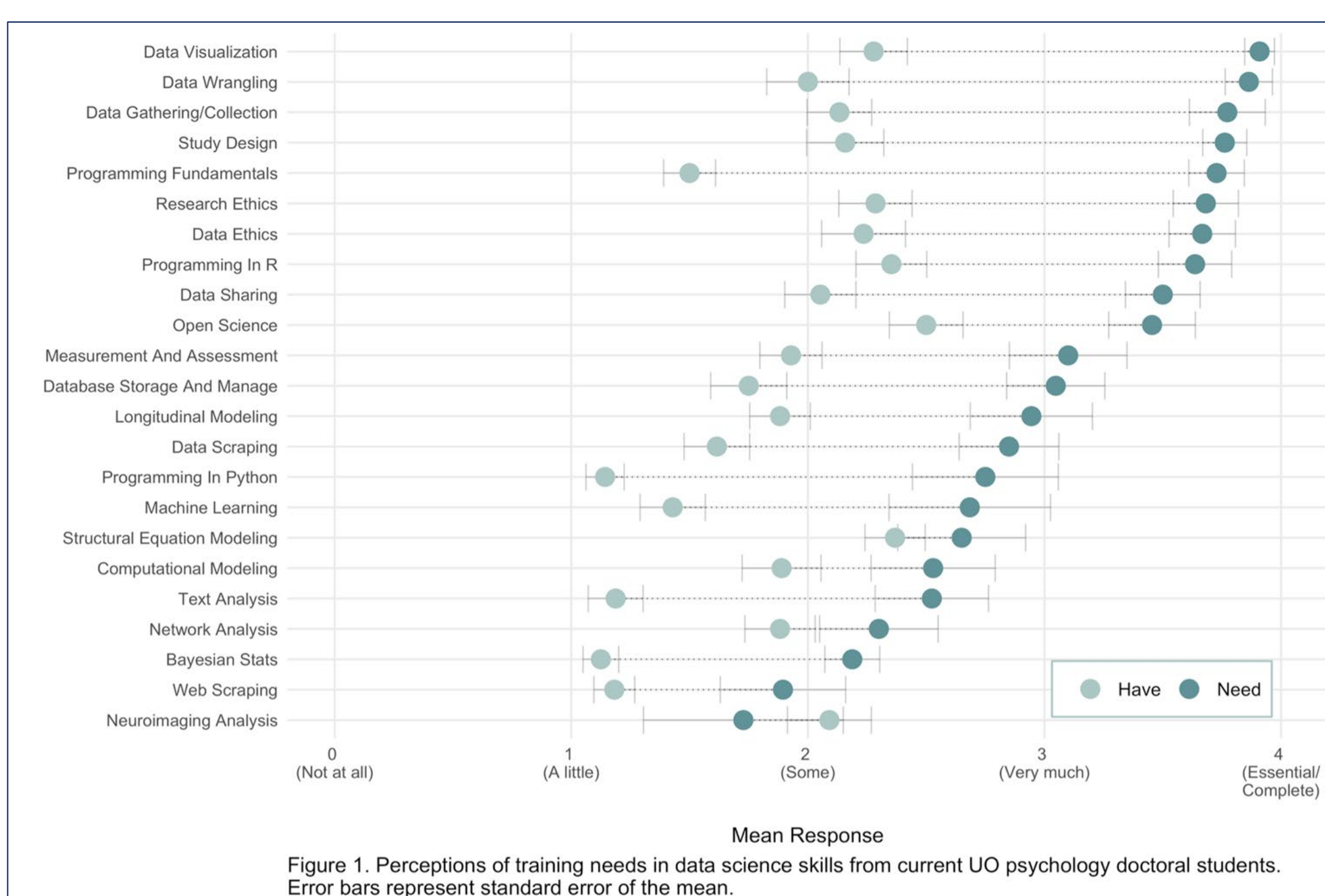

Figure 1. Perceptions of training needs in data science skills from current UO psychology doctoral students. Error bars represent standard error of the mean.
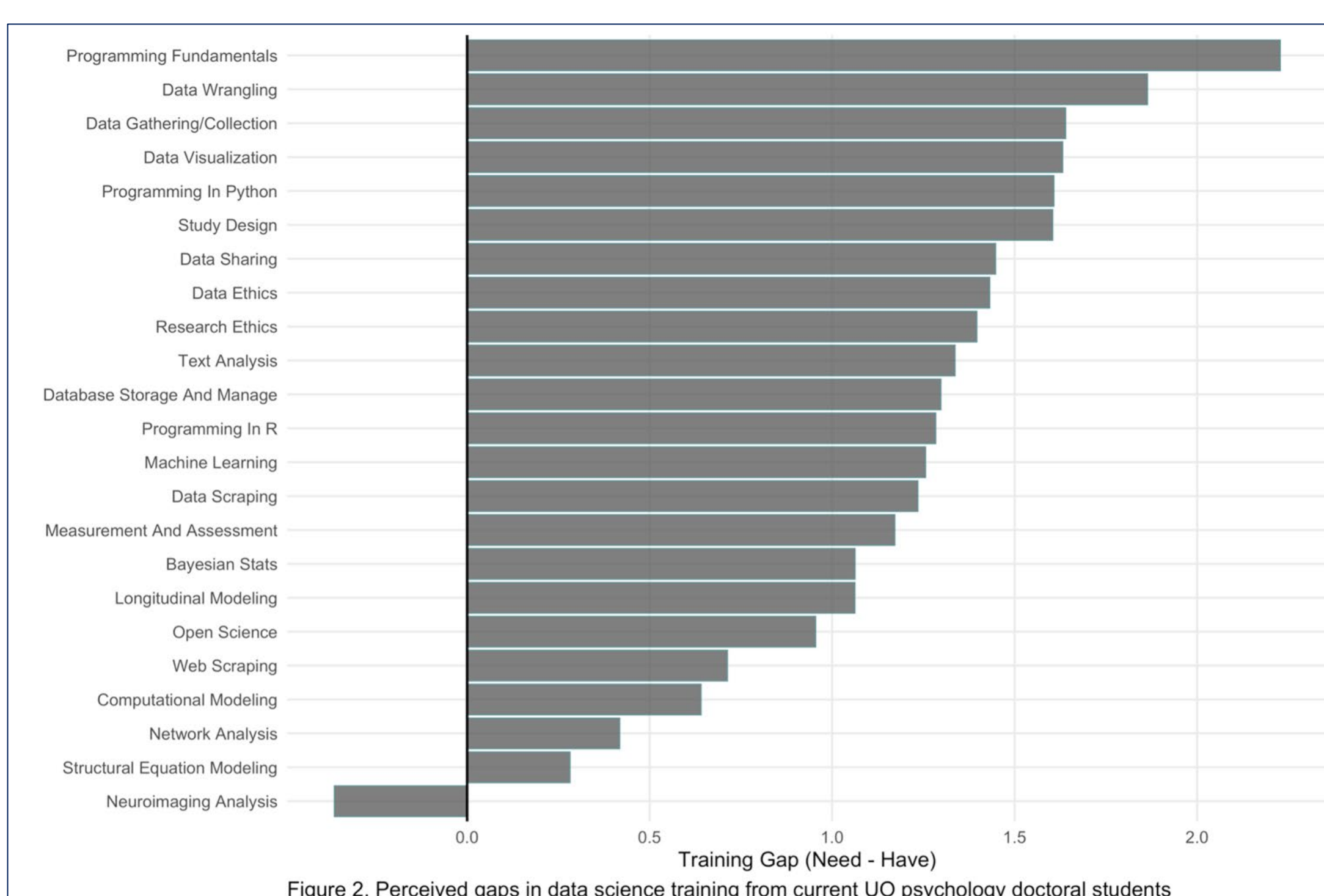

Figure 2. Perceived gaps in data science training from current UO psychology doctoral students

Cultivate a **community of practice** and **learning** around **data science**. Expand **informal data science training** to meet the growing demand for data science expertise at the University of Pittsburgh.

Project Description

- Leverage existing data science expertise at Pitt (gather, update, and create new informal open-access instructional materials)
- Scale up instructional capacity (with new instructors and workshop design)
- Feature the libraries as a hub to create a diverse and inclusive data science learning community.

Project Deliverables

- Environmental scan and gap analysis of existing informal education opportunities
- Small grants for the development of informal data science training materials
- Self-guided workshop materials for personalized, independent learning
- Instructor training materials (i.e. train the trainers) for scaling data science education
- Expanded data science workshop offerings in the Libraries and School of Computing and Information
- Online space for community of practice with code of conduct and guidelines

Potential Impact

- Expand the computational expertise and capacity of Pitt's researchers and educators
- Create a community of practice ready to take full advantage of computational infrastructure
- Foster interdisciplinary and cross-community collaboration in research and teaching

References and/or Acknowledgements

- Image Credits: Brendan Cullen *Data Science Training in Psychology*. https://bcullen.rbind.io/post/2020-03-08-data-science-training-needs-in-grad-school/