

Maximizing Student Development in Data and Information Science-Related Disciplines for Biomedical PhD Trainees

DHHS-NIH-National Institute of General Medical Science
Award T32 GM135748-02S1

PI: Nick Duffield, Texas A&M Institute of Data Science duffieldng@tamu.edu

Presenter: Christi Retzer, Texas A&M Institute of Data Science, ceretzer@tamu.edu

<https://training.tamids.tamu.edu/biomedds>

training.tamids.tamu.edu/biomedds



TEXAS A&M
Institute of
Data Science

BIOMEDICAL DATA SCIENCE ONLINE TRAINING PROGRAM

Project Objectives

- **Curriculum**

- Develop new curriculum of exportable and shareable training modules in Data Science integrating exposition, practice and biomedical context

- **Delivery**

- Deliver curriculum through monthly training events to a broad audience of current trainees across 6 biomedical graduate training programs at Texas A&M University and other institutions.

- **Evaluation**

- Evaluate effectiveness of training and outcomes of the learning objectives to inform future amendments to the training curriculum content and delivery methods.

training.tamids.tamu.edu/biomedds



TEXAS A&M

Institute of
Data Science

BIOMEDICAL DATA SCIENCE ONLINE TRAINING PROGRAM

Biomedical Data Science Online Training Program

- **Schedule**

- 12 monthly online 2-hour sessions (3 recorded, 9 live)

- **Topic Groups**

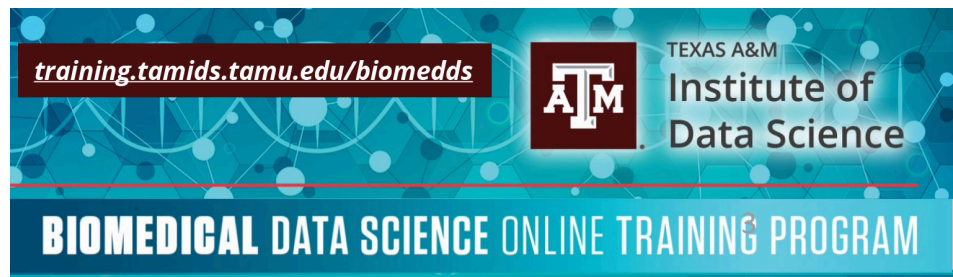
- R language primer and data manipulation (3 sessions)
- Statistical learning methods (4 sessions)
 - Regression, clustering, classification, dimension reduction
- Principles and practice of FAIR Data (2 sessions)
- Data privacy, policy, ethics and equity (2 sessions)
- Cloud computing and big data analytics (1 session)

- **Instructors**

- Faculty SMEs

- **Format and materials**

- R notebooks and presentations in domain context
- Quizzes, exercises & hands-on computation practice
- Recorded videos available post-session



Audience and Reach

- **Study Field and Experience**

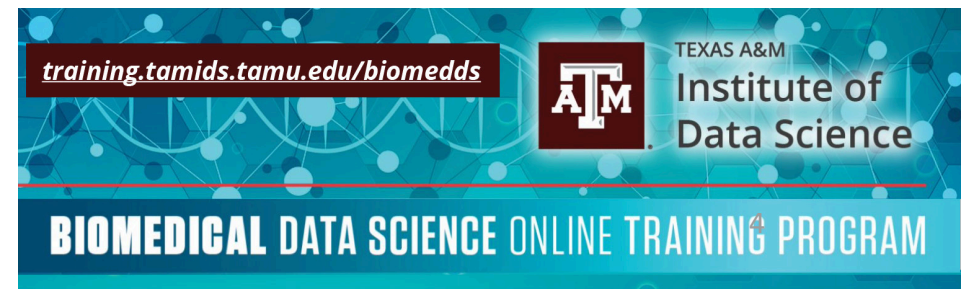
- Primarily at graduate students in the biomedical sciences
- All others interested in the field were welcome to join

- **Institutional Scope**

- Texas A&M Biomedical Graduate Programs
- Texas A&M Institute of Data Science, Colleges of Engineering, Science, & AgriLife
- Peer programs to Texas A&M Superfund Research Centers

- **Attendance**

- Total 97 participants in 9 live sessions
 - 68 TAMU (of which 12 IMSD), 29 non-TAMU



Program Evaluation: Participant Experience

- **Program Evaluation**

- Conducted by Texas A&M Education Research Center

- **Participant Confidence**

- On average, participants perceived their confidence to apply the learning objectives prior to attending the sessions increased from *not at all confident* and *somewhat confident* but **Confident** after attending the sessions.

- **Session Structure**

- Overall, participants felt the use of examples and sharing of relevant code were effective in increasing their understanding of session content, and Q&A and collaboration with other participants contributed to their learning.

training.tamids.tamu.edu/biomedds



TEXAS A&M
Institute of
Data Science

BIOMEDICAL DATA SCIENCE ONLINE TRAINING PROGRAM

Challenges and Future Work

- **Non-attending registrants**
 - 338 session registrants -> 97 participants
- **Diagnostic included in post-session surveys**
 - Non-participants reported scheduling conflicts and session length as impacting factors
- **Making curriculum available more broadly**
 - Reimplementation as fully asynchronous modular offering
 - Will enable self-paced study over a set of shorter modular units
- **Current course**
 - Recordings and materials remain online at [TAMIDS](https://tamids.tamu.edu) and [YouTube](https://www.youtube.com) (407 total views)
- **Publication**
 - Evaluation report proposed to Conference of Southwest Educational Research Association