



Michael Mullarkey

Principal Data Scientist

April 2021

📍 Lab for Scalable Mental Health

🏠 mcmullarkey.github.io

✉️ mullarkey.mike@gmail.com

🐦 [mcmullarkey](https://twitter.com/mcmullarkey)

🎧 mcmullarkey

in mcmullarkey

About me

Michael is a clinical psychology PhD with expertise in data science and scaling mental health supports.

Education

2014-2020 PhD - Clinical Psychology University of Texas at Austin

2011-2013 MA - Clinical Psychology American University

2008-2011 BA - Psychology American University

Work Experience

2020- Principal Data Scientist Lab for Scalable Mental Health

2019- Consulting Data Scientist Institute for Mental Health Research

2019-2020 Data Scientist Lab for Scalable Mental Health

Selected Projects

- Created an automated screener for suicidality in text data using tree-based and neural network models that saved 20 hours of “by hand” work per week: [Link to Code on Github](#)
- Developed a scalable time series forecasting pipeline using tsfresh feature engineering and a random forest algorithm versatile enough to be used across multiple organizations: [Link to Code on Github](#)
- Built a Shiny dashboard to track available data for mission-critical projects according to stakeholder identified criteria, used at least once weekly within the organization: [Link to Code on Github](#)
- Gave invited presentations at national conferences on how to create reproducible workflows for clinical scientists that resulted in a >90% sign-up rate for reproducible workflow tools: [Link to Presentation Slides](#)

Skills

- Expert in R and RStudio (tidyverse, tidymodels, torch, Shiny, R Markdown)
- Familiar with Python (NumPy, Pandas, Scikit-learn, Keras, TensorFlow)
- Familiar with SQL
- Git and Github
- Time Series Forecasting
- Multivariate Data Analysis

Other Information

- I've published >20 technical papers, cited nearly 1,000 times
- I stream live data science on Twitch
- My cat Boo helps me code because her code review respects piping