Esteban Buz

Vancouver, BC, Canada | esteban.buz@gmail.com | +1 604-868-8088

www.esteban.bz | github.com/ebuz | linkedin.com/in/esteban-buz

Skills

Principal languages:	Python (incl. Numpy, Pandas, Scikit-learn), R (incl. tidyverse), SQL, JavaScript, Bash
Analytics/ML:	linear & logistic regression, Bayesian statistics, non-parametric statistics (bootstrapping, permutation), classification (random forest, SVM, XGB), time series/forecasting
Tools & Frameworks:	AWS (SageMaker), Jupyter, R Shiny, Flask, Dash, FastAPI, Docker, Git/Github
Domain expertise:	experimental methods (hypothesis, A/B, and multi-armed bandit), cognition, decision making, speech perception/production, linguistics

Experience

SonderMind Senior data/applied scientist   Remote/Denver, CO | Aug 2021—Present
- Developed tooling to enable continuous, reproducible, and automatic deployment of ML services via AWS
- Developed and deployed a client-provider recommendation system that increased rate of initial therapy sessions
- Developed an A/B and Multi-armed bandit testing framework for ML experimentation of production services
- Conducted statistical analyses of factors influencing client needs in mental health and identified steps to improve how providers can communicate with clients

Qntfy (acquired by SonderMind) Senior data scientist   Remote/Boston, MA | Aug 2020—Aug 2021
- Developed a client-facing tool in Python and Google Sheets that integrated ML analytics in a financial portfolio management system
- Developed individual and group level mental health measures based on social media data to understand mental health needs at the individual and group level for non-profit organizations who can meet those needs
- Developed analytic and reporting pipelines of large scale social media data (10k> users and millions> of messages) for internal and external stakeholders to understand emergent trends and clusters

Insight Data Science Data science fellow   Boston, MA | Jan 2020—Apr 2020
- Built a tool to predict availability of plants in Massachusetts' public parks with machine learning techniques (e.g. SVM) using millions of plant ID records, climate data, and geo-biome data from the US geological survey

Princeton University Postdoctoral fellow   Princeton, NJ | Sep 2016—Aug 2020
- Developed bootstrapping, permutation, and Bayesian analytic tools in R to quantify asymmetric changes at the lower and upper range of speech outcomes, finding that humans modify their speech at the scale of milliseconds
- Designed experimental protocols to investigate human adaptations to dialects, developed a React, Node.js and PostgreSQL experimental tool to administer these protocols, and an R pipeline to process collected audio data
- Designed and taught an undergraduate seminar on psychological and linguistic approaches to understand conversation, including developing an automatic analysis pipeline in R for students to use in class projects

Vanderbilt University Postdoctoral fellow   Nashville, TN | Sep 2018—Aug 2019
- Modeled speech tone importance for language comprehension using classification techniques (e.g., Bayesian logistic regression) to understand comprehension deficits in individuals with cochlear implants

University of Rochester Graduate researcher and instructor   Rochester, NY | Sep 2010—Aug 2016
- Designed experimental protocols to assess psychological theories of speech change, developed a Flask and PostgreSQL audio/visual web-based recording tool deployed to AWS to administer these protocols, and developed an R analytics pipeline to assess the hypotheses using thousands of speech records
- Contributed to a multi-person effort to design, develop, and deploy a Django+Shiny web-based tool for hundreds of undergraduates to self-design, conduct, and auto-analyze human behavioral studies

Education

PhD in Brain and Cognitive Sciences, University of Rochester	2016
BA in Cognitive Sciences w/minor in Computer Science, Johns Hopkins University	2008