

Ian James Douglas

Austin, TX | (737) 336-9698 | jjdouglas6492@gmail.com

[LinkedIn](#) | [Website](#) | [Github](#)

SKILLS

Languages: python, R, SQL, Bash/Linux, Swift, HTML, Git, Docker, Singularity

Packages: pytorch, tensorflow, sklearn, open-cv, pandas, matplotlib, ggplot2, shiny, tidyverse

Statistical Analysis: Machine Learning & AI (CNN/RNN, XGBoost, SVM, Random Forest...)

Data Visualization, Dimension Reduction (PCA, Network Analysis, t-SNE...), Clustering

(KNN, MDS...), Regression/ANOVA, Model Inspection, Feature Selection, A/B Testing

Other: Communication, Model Insight Translation, Presentation, Writing, Teaching, Leadership

PROFESSIONAL EXPERIENCE & ACADEMIC RESEARCH

Data Scientist [contractee], *Data Science Insights, LLC* Jan 2023 - Present

- Clients: [Harvard Business School](#), [Illuminas](#), [ESADE Business School](#)
- I published an [online tool](#) to conduct A/B testing & power analysis for market research
- I used anomaly detection to identify hospitals posing the greatest risk to insurance providers, creating user-friendly results that my client could discuss and present to stakeholders
- [I used latent Dirichlet allocation](#) to identify salient topics to consumers who were interviewed about their purchases to drive a targeted marketing campaign

Data Analyst, Provider *Finance Dept., TX Health and Human Services* Mar 2023 - Present

- [I developed a novel gradient boosting algorithm](#) to minimize financial burden on health insurance/maintenance organizations while maximizing payments to physicians for paid medical claims, saving my department millions in lawsuits and litigation
- [I applied network analysis to geospatial data](#) to predict where policy changes are and will be most impacting insurers and providers, which dictated their implementation across TX
- As the Data/Research Team Lead, I review code, formulate analytic questions, supervise implementation, automate ETL pipelines, and conduct [statistical analyses to present findings](#) to diverse departmental and public audiences on quarterly fiscal metrics

Data Analyst [contractee], *Trinity Healthcare Resources* (Austin, TX) Jan 2023 - Mar 2023

- I used data build tool (dbt) and SQL to automate ETL pipelines to organize, analyze, and present billions of rows of cost report data from hospitals across the USA; computing metrics embedded in a subscription-based product the company sold to participating hospitals

Ph.D. Researcher, *the University of Texas at Austin* Aug 2020 - Jan 2023

- I applied computer vision and transfer learning to images I scraped using the Reddit API, to identify valuable aesthetic features and predict upvotes by reddit users [GitHub](#)

Data Analyst, *Columbia University in the City of New York* Dec 2019 - Aug 2020

- I applied gradient boosted decision tree classification and custom variable importance analysis to identify brain morphology associated with childhood trauma exposure, leading to discoveries on brain development and academic publication [GitHub](#), [Paper](#)

EMPLOYMENT HISTORY (RECENT)

Data Scientist, *Self-employed*, [remote] Mar 2023 – Present

Data Analyst, *Dept. of Provider Finance, TX HHSC*, [remote & Austin, TX] Mar 2023 – Present

Data Analyst, *Trinity Healthcare Resources LLC*, Austin, TX Jan 2023 – Mar 2023

PhD Researcher/Teaching Assistant, *The Univ. of Texas, Austin*, TX Aug 2020 - Jan 2023

Data Analyst, *Columbia University*, New York City, NY Jan 2019 - Aug 2020

EDUCATION

Ph.D., Psychology Research, The University of Texas, *Austin, TX* 2023

M.S., Applied Statistics, Columbia University, *New York City, NY* 2020

B.S., Psychology, Trinity College, *Hartford, CT* 2015