# Ian James Douglas

Austin, TX | (737) 336-9698 | ijdouglas6492@gmail.com

LinkedIn | Website | Github


## SKILLS

**Languages**: SQL, python, R, Linux, Matlab, HTML, Javascript, Git, Docker, Singularity

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

**Statistical Analysis**: Predictive Machine Learning (e.g., XGBoost, SVM, Random Forest, etc) Data Visualization , Computer Vision, Dimension Reduction (PCA, Networks, t-SNE, etc), Clustering (KNN, MDS, etc), Decision-Trees, Regularization, Regression/ANOVA, Model Inspection / Feature Selection, Bootstrapping, Model Diagnostics, AB Testing

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

## PROFESSIONAL EXPERIENCE & ACADEMIC RESEARCH

**Data Analyst** [contract], *Trinity Healthcare Resources*, Austin, TX Jan 2023 - Mar 2023

- I built proprietary SQL models to extract raw healthcare claims data, transform it efficiently, and feed pipelines for analysis. I wrote analyses to serve a dashboard to track business objectives, which improved the efficiency of team members to divert efforts where needed.
- I wrote statistical analysis pipelines to track the influx of different patient populations at hundreds of hospitals across the USA, culminating in reports with visualizations of trends, which contributed to a subscription-based product that hospitals subsequently purchased.

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

- Using authorized APIs I scraped posts from Reddit containing images that I used to build a computer vision model of user preferences, showing that common aesthetic principles guided the perception of images from different sub-Reddits GitHub
- Using network analysis, I mined surveys about behavioral and emotional responses to the COVID-19 pandemic to determine common patterns of experiences and participant clusters, culminating in a presentation of the findings at a relevant conference GitHub

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

- Working in a neuroscience lab I built an interpretable machine learning classification pipeline to identify brain morphology associated with childhood trauma exposure, which led to new and undiscovered insights about brain morphology and development GitHub
- Working in a neuroscience lab I programmed an analysis pipeline to automate machine learning predictive model comparison and feature selection, resulting in peer-reviewed publication reporting the relation between gastrointestinal symptoms and anxiety Paper

## EMPLOYMENT HISTORY

**Data Analyst, *Trinity Healthcare Resources LLC*,** Austin, TX Jan 2023 – Mar 2023
**Statistics Consultant, *The Univ. of Texas*,** *Austin*, TX Aug 2022 - Jan 2023
**Graduate Research Analyst, *The Univ. of Texas*,** *Austin*, TX Aug 2021 - May 2022
**PhD Research/Teaching Assistant, *The Univ. of Texas*,** *Austin*, TX Aug 2020 - Jan 2023
**Data Analyst, *Columbia University*,** *New York City*, NY Jan 2019 - Aug 2020
**Clinical Research Site Manager, *Vanguard Research Group*,** *Glen Oaks*, NY Aug 2016 - Aug 2018
**Research Coordinator, *Vanguard Research Group*,** *Glen Oaks*, NY Aug 2015 - Aug 2016

## 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