

# Jacob Merrell

303-960-5156 • merrell.jacob@gmail.com • [data science website](#)

---

## EXPERIENCE

### **Senior Data Scientist** *Nav Technologies*

2021 – Present

- Developed XGBoost model in Python to predict probability of approval for credit cards. Increased approval rates by almost 100%. Credit card revenue has also increased by over 100%.
- Created ETL process using APIs, web scraping in Python (Selenium), SQL, AWS, Redshift
- Managed team to report financial health companywide. Complex PostgreSQL queries populate the monthly report.

### **Senior Data Analyst/Data Scientist** *Mercer*

2019 – 2021

- Used discriminant classification/clustering model and geospatial data in SAS to impute missing data (with 89% accuracy on test data) on a 2 TB Medicaid dataset
- Risk adjustment project leader for California's Medicaid program. Oversee running of regression models in R to create cost weights program. Results are implemented state wide
- Predicted dispensing fee costs for pharmacies using Python and linear regression. Preprocessed the data and imputed values where there were errors
- Manage trainings for new employees. Provide mentorship and taught analysts SQL and SAS

### **Data Scientist** *Brigham Young University*

2017 – 2018

- Linked individuals (using random forests and logistic regression in Python) with 90% accuracy from a dataset of 60,000 European immigrants to a pioneer dataset recorded in the Utah Valley
- Created quadratic discriminant clustering model in R to identify malignancy of cancerous tumors
- Led and managed two junior statistics analysts

### **Actuarial Analyst** *Milliman*

2015 - 2017

- Used R to simulate large claims experience and based accrual recommendations on results
- Developed a process to streamline the creation of pro forma scenarios pulling data using SQL

## PASSION PROJECTS AND SKILLS

- Skills and Languages: Python, SQL, PostgreSQL, AWS, GCP, R, SAS, pandas, scikit learn, TensorFlow, regression, and NLP
- Accessed YouTube's API and XGBoost in Python to predict success of YouTube videos with 86% accuracy. Built deep learning model using TensorFlow for facial and text recognition ([See More](#))
- Scraped box office data using Python's BeautifulSoup package for nearly 13,000 movies. Trained random forest model which explained 80% of the variation in box office revenue ([See More](#))
- Other projects include spatial regression and time series AR(1) ([See More](#))
- Fluent in Spanish (speaking, writing, and reading)

## EDUCATION AND CERTIFICATIONS

### **Brigham Young University**, Provo, UT

*Bachelor of Science, Actuarial Science*

- Cumulative GPA 3.91/4.0

### **Western Governors University**, Salt Lake City, UT

*Master of Science, Data Analytics*

- Coursework includes Natural Language Processing(NLP), Neural Networks, Time Series, tree based methods, data mining, SQL, k-nearest neighbors (KNN), and Naïve Bayes