

Jacob Merrell

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

Data Scientist II, *Amazon* (2022 – Present)

- **Boosted revenue by \$40MM by enhancing recommendation system recall** from 45% to 67% using neural networks, embeddings, and unsupervised learning. Orchestrated cross-functional collaboration with over 30 multinational stakeholders for system deployment. Productionalized models in AWS
- **Created Gen-AI insights, saving hundreds of labor hours** by implementing AI agents, RAG, vector databases, and LLMs for weekly analytical reports
- **Increased Efficiency by \$7.2MM** through predictive anomaly detection with AWS SageMaker, deploying scalable solutions via Lambda, S3, and Redshift
- Leveraged LLMs, Amazon Bedrock, prompt engineering, and few-shot learning to mine data from product descriptions. Methodology used in product recommendation system

Senior Data Scientist, *Nav Technologies* (2021 – 2022)

- **Doubled credit card revenue** with an XGBoost model in Python for credit approval probability prediction. Used airflow for orchestration, and AWS to productionalize
- **Managed team** to report organization-level critical financial health to CEO and all management
- Data engineered an ETL process utilizing APIs, web scraping (Selenium), SQL, AWS, and Redshift

Senior Data Analyst-Data Scientist, *Mercer* (2019 – 2021)

- **Affected 12 million individuals** by leading risk adjustment project for California's Medicaid program
- **Imputed data on a 2TB Medicaid dataset** with 89% accuracy. Used clustering methods and geospatial data
- Developed linear regression models for pharmacy dispensing fee predictions, managing data preprocessing in Python

Data Scientist, *Brigham Young University* (2017 – 2018)

- Linked individuals with 90% accuracy from a dataset of 60,000 European immigrants to a pioneer dataset using random forests and logistic regression in Python
- Created clustering model in R to identify malignancy of tumors, managed two statistics analysts

Actuarial Analyst, *Milliman* (2015 - 2017)

- Used R to simulate large claims experience and based accrual recommendations on results

EDUCATION, PASSION PROJECTS AND SKILLS

- Python, SQL, machine learning, LLMs, AI, NLP, deep learning, Sagemaker, S3, Redshift, PostgreSQL, AWS, R, SAS, pandas, scikit learn, regression, classification
- Predicted success of YouTube videos and built neural network for facial recognition
- Explained 80% of the variation in box office revenue by training a random forest model
- I love playing tennis, ping pong, chess, I'm fluent in Spanish and am learning Portuguese

Bachelor of Science, Actuarial Science
Brigham Young University, Provo, UT

- Cumulative GPA 3.91/4.0

Master of Science, Data Analytics
Western Governors University, Salt Lake City, UT