**Jacob Merrell**

 303-960-5156 • merrell.jacob@gmail.com • [data science website](https://jmmerrell.github.io/)

**Experience**

**Data Analyst/Data Scientist** *Mercer*

February 2019 – Present

* Used discriminant classification model in SAS to impute missing data (with 89% accuracy on test data) on a 2 TB Medicaid dataset
* Ran multiple regression models in SAS to create cost weights for the state of California’s risk adjustment program. The cost weights are now used in calculating statewide risk adjustment amounts.

**Data Science Analyst** *Brigham Young University*

September 2017 – November 2018

* Linked individuals (using ensemble learners in R) with 90% accuracy from a dataset of 60,000 European immigrants to a pioneer dataset recorded in the Utah Valley.
* Created quadratic discriminant analysis model to identify malignancy of cancerous tumors
* Involved in denumeration project which uses linear models to de-identify health data

**Actuarial Analyst** *Milliman*

December 2015 - August 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 which showed high and low estimates for the gain/loss

­­

**Statistics/Property Valuation Intern** *The Church of Jesus Christ of Latter-day Saints*

May 2014 - Dec 2014

* ­Saved $72 million by building a VBA model to predict membership utilization of meeting houses
* Created a multiple regression model in R to predict land values

**Passion Projects and Skills**

* Languages: Python, R, SAS, SQL,VBA, HTML, JavaScript
* I have two YouTube channels ([JADoinStuff](https://jadoinstuff.com/) and [El Güerito](https://www.youtube.com/channel/UC_3VPXVC1-ym83R3fOIMpVQ)) I do for fun! Using YouTube’s API to analyze competitor channel data. Employing Google Vision and OpenCV for facial and text recognition
* Scraped box office data using Beautiful Soup for nearly 13,000 movies. Trained random forest model which explained 80% of the variation in box office revenue given all the data scraped. ([See More](https://jmmerrell.github.io/movies/))
* Used spatial regression to predict air pollution for most of the USA. The adjusted R-squared was 0.6129, with a normalized RSME of 0.174. ([See More](https://jmmerrell.github.io/spatial/))
* Estimated the expected savings ($86.69) for using solar energy using a time series AR(1) model. The model had an adjusted R-squared of 0.9376, and used the findings to predict how many years it would take the individual who provided the data to recoup initial cost of the solar panels. ([See More](https://jmmerrell.github.io/time/))
* Public speaking; addressed an audience comprised of members of the Federal Reserve Bank, and 500 other students
* Fluent in Spanish (speaking, writing, and reading)

**Education**

**Brigham Young University**, Provo, UT

*Bachelor of Science, Actuarial Science*

* ­Cumulative GPA 3.91/4.0
* Certificates: SAS Programming, and Advanced Statistics
* Actuarial Exams: Financial Mathematics (FM), Probability (P), and Models for Financial Economics (MFE)