

# Zack Mosley

zmosleymis@gmail.com | (209) 487-4826

[Portfolio](#) | [LinkedIn](#)

## Education

### M.S. in Data Science

*University of California, San Diego*

La Jolla, CA

Expected Graduation: March 2026

University GPA: 3.80 /4.00

- Relevant Coursework: Statistics, Linear Algebra, Data Visualization, NLP, Big Data

### B.S. in Management Information Systems

*California State University, Sacramento*

Sacramento, CA

May 2024

University GPA: 3.80 /4.00

- Honors: **Magna Cum Laude**, Dean's Honor List
- Relevant Coursework: Machine Learning, Data Mining, Advanced Analytics, Data Mgmt

*A.S. in Accounting*

May 2021

*A.S. in Business Administration*

## Projects

### Car Value Predictor (Personal Project, [Full Project](#))

End-to-end vehicle price prediction dashboard using scraped data from Cars.com.

- Built **ETL pipeline** to ingest/clean **80K+ vehicle listings** across top 10 U.S. manufacturers
- Engineered features (trim, mileage, year, body style); analyzed price trends via EDA
- Analyzed depreciation trends by make/year to quantify long-term value patterns
- Developed predictive model using LightGBM regressor (**MAE ~\$1.7K,  $R^2 \approx 0.96$** )
- Deployed interactive Dash/Plotly dashboard for real-time pricing & feature insights

### NYC Yellow Taxi Fare Prediction (Academic Project, [My Contributions](#) | [Full Project](#))

Led analysis of **150M+** NYC taxi rides using Spark on SDSC supercomputer, earning **3rd place** in class-wide project competition.

- Conducted EDA on 20GB+ trip data: time trends, fare efficiency, tipping patterns
- Applied filters to outliers/invalid trips; engineered time & efficiency features
- Processed 170M+ trip records on distributed **Spark** cluster, retaining 88% clean data
- Developed fare predictor (**MAE ~\$1.91,  $R^2 \approx 0.94$** ) and tip classifier (**~79% accuracy**)
- Benchmarked against baseline regression, achieving **36% RMSE reduction**
- Evaluated performance with stratified sampling, residual analysis, and baseline benchmarks

## Skills

**Languages & Tools:** Python (Pandas, NumPy, Scikit-learn, Plotly, Matplotlib, Seaborn), SQL (PostgreSQL, MySQL), **D3.js**, Tableau, Power BI, **Git**, JavaScript, Streamlit, Azure, AWS, Excel

**ML & Analytics:** Statistical Analysis, Regression, Classification, NLP, A/B Testing, Feature Engineering, Cross-Validation, Model Evaluation

**Data Engineering:** ETL, APIs, Spark, Dask, Data Pipelines

## Experience

### DBI Beverage

2014 - 2020

#### Lead AM Warehouse Associate

Stockton, CA

- Led team of 7 managing distribution of 50K+ cases/day for Coors, Modelo, and Corona brands
- Spearheaded EasyOps implementation, partnering with vendors to modernize warehouse workflows
- Improved inventory accuracy 12% by analyzing operational data & implementing improvements
- Led warehouse expansion: designed floor layout from planning through implementation