

# Matthew Coleman

408-882-7585 | coleman.matthew.b@gmail.com | www.linkedin.com/in/matthewbriancoleman | matthewbriancoleman.github.io/

---

## WORK EXPERIENCE

### SAP

Data Scientist (SAP Data Custodian) September 2021 – March 2022

- Developed a sensitive information NER model in Python and improved the model's F1 score from 76% to 93.6%.
- Used Python for EDA, data cleaning, ML model training, and performance evaluation.
- Collaborated with global teams to plan and implement further improvements to the existing NER model and API.

Data Scientist (SAP SuccessFactors) March 2022 – September 2022

- Used association rule mining in PySpark to serve related skills to user's job profiles, improving job opportunities.
- Leveraged an AWS Cluster with parallel computing to increase association rule querying speed by 300%.
- Evaluated team's use of customer data and AI to ensure ethical development practices.

Data Scientist (SAP S/4HANA Cloud Innovation Frontrunner PM Team) September 2022 – Present

- Working with guided implementation of S/4HANA Cloud to decrease customer onboarding time by 500%.
- Collaborating with teams in India, Germany, and the US to meet strict deadlines and deliver results on-time.

### NSF Funded Central Coast Data Science Fellowship

Data Science Fellow September 2020 – June 2021

- Worked with three fellows and a Senior Data Scientist at the biomedical company, Evidation Health.
- Performed data wrangling and feature engineering on raw data to create meaningful variables for teammates use.
- Used multiple change point analysis on smartwatch health data to pre-symptomatically predict COVID-19.

### BigIron Auctions

Data Science Intern (Remote) July 2020 – September 2020

- Developed interactive dashboards to optimize marketing and sales teams' decision-making.
- Created a predictive model in Python to determine and target likely bidders and buyers of equipment.
- Utilized MySQL server to query data necessary for visualization, model building, and prediction.

### UCSB Department of Probability and Statistics

Undergraduate Learning Assistant January 2020 – April 2021

- Enhanced learning experience of 600 students enrolled in *Data Science in R* and *Bayesian Data Analysis* courses.
- Held individual drop-in office hours and supported computer lab sections to assist students with coursework.

Undergraduate Research Assistant January 2020 – April 2020

- Coded formula interface for Wasserstein Regression function in the R package, WRI.
- Authored and coded package vignette which displays the package's functions and usage.

### UCSB Department of Probability and Statistics Administrative Office

Transfer Peer Advisor (Remote) March 2019 – June 2021

- Individually assisted 200+ transfer students with admission steps, course clearance, and schedule planning.

### UCSB Baseball

Analytics Intern May 2019 – December 2019

- Conducted player research with UCSB D1 baseball team during 2019 Big West Conference championship season.
- Performed data analysis, created ML models, wrote research reports in R and Python to guide player development.
  - Bat Point of Contact Effects using LOESS (R); Cluster Analysis of Curveballs in D1 Baseball (Python).

## EDUCATION

### University of California, Santa Barbara (UCSB)

B.S. Statistics and Data Science June 2021

- 3.77 Cumulative GPA | 3.83 Major GPA | Dean's Honors: Winter 2019, Spring 2019, Fall 2019, Spring 2020
- Spanish Biliteracy (speaking, reading, writing).
- Projects: Twitter Hate Speech Classifier (PySpark), Airbnb Price Regression and Median Price Classification (R).