Hailey Ernest

San Diego, CA | Hailey.ernest@icloud.com | 425-491-0996 | haileyernest.com

SUMMARY

Statistics M.S. candidate with two years of data analytics experience at Ondine Biomedical, building predictive models and validated data pipelines for clinical data. Developed an unsupervised ML pipeline for NASA Ames with 99% clustering stability. Skilled in Python, R, and SQL, and in turning complex results into clear visualizations for non-technical stakeholders.

WORK EXPERIENCE

Ondine Biomedical, Data Analytics Consultant Bothell, WA | Sept 2024 - Sept 2026

- Engineered reproducible, validated data cleaning and transformation workflows across 4 types of clinical and laboratory data, including multi-run observational datasets.
- Designed automated laboratory inventory and batch-record management workflows for sensitive data.
- Delivered statistical findings to non-technical audiences through clear visualizations and summaries.

Ondine Biomedical, Data Analytics Intern Bothell, WA | June 2024 - Sept 2024

- Built predictive models for R&D and clinical trial data supporting pre-surgical infection prevention research.
- Developed analytical pipelines in Python, R, and Excel to process clinical lab data.

Cal Poly Statistics Department, Instructional Student Assistant San Luis Obispo, CA | Sept 2025 - Present

- Assist with instruction for statistics courses in the Bailey College of Science and Mathematics.
- Grade coursework across courses ranging from Business Statistics and Engineering Statistics to Applied Multivariate Analysis and Survival Analysis.

PROJECTS

NASA Ames Research Center, Acoustic Boiling Regime Classification Jan 2026 - June 2026

- Built an unsupervised ML pipeline (UMAP + HDBSCAN) that classified cryogenic boiling regimes from 446 acoustic accelerometer runs, achieving 99% clustering stability across parameter changes.
- Engineered 70+ statistical features and removed runtime-correlated bias with per-feature linear regression detrending and hyperparameter tuning.
- Validated clusters with leave-one-out resampling and designed a frozen-embedding GMM to classify incoming data in real time without retraining.

World Health Organization & Cal Poly Kinesiology and Public Health Dept. Aug 2026 - Present

- Clean, map, and validate physical activity data for a WHO-aligned international consortium (Cal Poly, NCI, Oxford, Monash).
- Evaluate open-source wearable algorithms against physical activity metrics using bias, precision, and equivalence analysis.

Global Measles Trends Dashboard & R Package Apr 2026 - June 2026

- Built mixed-effects, logistic, and multivariate regression models in R with country-level random effects, applying Bonferroni correction and likelihood ratio tests.
- Developed a custom R package and deployed an interactive Quarto and R Shiny dashboard to share results.

EDUCATION

California Polytechnic State University, M.S. Statistics San Luis Obispo, CA | Expected June 2027

California Polytechnic State University, B.S. Statistics & Data Science San Luis Obispo, CA | March 2026

Relevant coursework: Machine Learning, Data Science, Statistical Computing, Database Management Systems (SQL), Generalized Linear Models, Design & Analysis of Experiments, Survival Analysis

Honors: Dean's List (five terms)

LEADERSHIP

AWS Cloud Club, President/Captain Mar 2024 - June 2026

- Led technical workshops on AWS cloud services for Cal Poly students.
- Guided members in building and deploying real-world application prototypes on cloud infrastructure.

SKILLS & TECHNICAL TOOLS

Data Science & Modeling: Predictive Modeling, Machine Learning (XGBoost, Random Forests, SVMs, ensemble methods), Clustering & Dimensionality Reduction (UMAP, HDBSCAN, GMMs), Statistical Modeling (GLMs/GLMMs, mixed-effects models, survival analysis), Time Series Forecasting, Exploratory Data Analysis

Experimentation & Analytics: A/B Testing, Logistic, Multivariate, and Poisson Regression, Hypothesis Testing, Experimental Design, Equivalence Testing, Cross-Validation, Quantitative Analysis

Languages & Databases: Python, R, SQL, SAS, SPSS, Excel/VBA, Data Warehousing

Data Engineering & Cloud: AWS (SageMaker, EC2, S3), Apache Spark, ETL Pipelines, Data Engineering, Version Control (Git/GitHub)

Visualization & Deployment: Tableau, Power BI, R Shiny, Quarto, Streamlit, ggplot2, Dashboard Development