Rebecca Bachtra

(916)300-3021 | beccaashley@berkeley.edu | linkedin.com/in/beccaashley/ | Portfolio

EDUCATION

University of California, Berkeley   Expected Graduation Date: May 2026
B.A. Statistics, Data Science | GPA : 3.8/4.0

Relevant Coursework: Computing with Data, Statistical Inference, Computational Structures in Data Science, Reproducible Data Science, Probability Theory, Concepts of Statistics, Financial Accounting, Principles of Business, Macroeconomic Analysis

SKILLS

Programming & Data Technologies: SQL, Python, R/Rstudio, HTML/CSS, Apache Spark (PySpark & Spark SQL), Pandas, NumPy, Scikit-Learn, PyTorch, NLP
Data Visualization & Tools: Tableau, Matplotlib, PowerBI, Excel, SAS, Jupyter Notebook, Git, Google Workspace, Knime
Languages: English (Native), Indonesian (Native), Chinese (Conversational)

RELEVANT EXPERIENCE

Bureau of Ocean Energy Management (BOEM)   August 2025 - Present
Data Science Intern
- Integrated high-resolution X-ray diffraction (XRD) mineralogical data with large-scale petrophysical and wireline sensor measurements (~100k observations) to support offshore reservoir characterization
- Performed statistical analysis and correlation studies to quantify relationships between sensor features and mineral composition, including clay content and lithologic properties
- Designed a standardized preprocessing pipeline to handle missing, sparse, and heterogeneous sensor data across multiple wells and measurement types

Stanford Medicine   June 2025 - Aug 2025
Process Design and Strategy Analyst
- Conducted exploratory data analysis on healthcare operational processes to assess performance, variability, and bottlenecks across workflows
- Developed data-backed process metrics and analytical frameworks to support strategic decision-making and system-level improvements
- Synthesized qualitative and quantitative data into actionable insights for healthcare leadership

Sinarmas Land   May 2024 - July 2024
Data Science Intern
- Built regression-based pricing models for 10+ commodity products to uncover predictive relationships between input factors and market performance
- Automated preprocessing of large-scale datasets in KNIME and Python, improving data throughput by 30%
- Analyzed variance and correlation structures to detect volatility trends across oil categories, informing pricing and risk strategies

Math Tutor   August 2022 - May 2024
- Tutored 50+ students weekly in probability, statistics, and regression; reinforced concepts with applied examples
- Cultivated a supportive learning atmosphere for students to overcome challenges and achieve their academic goals

RELEVANT PROJECTS

Advanced SQL Product Analytics & Reporting
- Built a **product-level analytical reporting view** by joining transactional fact data with product dimensions and aggregating core metrics including total orders, total sales, total quantity sold, unique customers, product lifespan, recency (months since last sale), and average selling price
- Implemented **rule-based product segmentation and KPI engineering in SQL**, classifying products into High-, Mid-, and Low-Performers based on revenue thresholds, and computing average order revenue (AOR) and average monthly revenue using defensive case logic

News-Driven Stock Predictor
- Designed a machine learning pipeline to predict stock price direction using sentiment signals extracted from global news data and aligned with high-frequency market data
- Engineered sparse, high-dimensional feature sets using NLP sentiment vectors, time displacement, and lag features; applied Lasso regularization for feature selection
- Evaluated multiple modeling approaches (linear regression, logistic regression, neural networks), analyzing bias–variance tradeoffs and generalization performance