Katrina Suherman

667-900-5834 | katyasuherman@gmail.com | linkedin.com/in/katrinasuherman | Portfolio

EDUCATION

University Of California San Diego San Diego, CA
Bachelor of Science - Data Science - 3.94 GPA *Sep 2023 – Mar 2026*

Edmonds College Lynnwood, WA
Associate's in Applied Science - Computer Science - 4.0 GPA *Sep 2021 – Jun 2023*

EXPERIENCE

Data Science Intern Sept 2025 – Present
Optimalex *Remote*
- Developed automated, scalable Python-based data pipelines to extract, transform, and standardize over 6K hospital price transparency files, enriching provider-level data and reliably uploading results to AWS S3.
- Built a prototype *Medical Expenses Calculator* with a Flask backend and dynamic frontend (HTML, CSS, JavaScript, Angular-style components) to help insurance carriers benchmark procedure prices and prevent overpayment.
- Performed exploratory data analysis and statistical validation on regional pricing patterns to identify anomalies, refine pricing models, and support fair-cost provider recommendations.

Data Analyst Intern Jun 2025 – Present
Eluve Health *Remote*
- Prompt-engineered and evaluated multiple AI models to generate tailored clinical recommendation notes for doctors, performing fact verification to enhance relevance and consistency across specialties.
- Designed and analyzed a Metabase Edit Distance Evaluation Dashboard to measure accuracy gaps between AI- and physician-edited notes, informing prompt refinements and workflow improvements.
- Defined key performance indicators (KPIs) to evaluate model precision and physician adoption rates over time.
- Collaborated with product and data science teams to iterate on model prompts, improving factual accuracy and physician usability based on quantitative feedback.

Data Science Intern Jun 2025 – Sept 2025
Nobu Bank *Remote*
- Supported development of an Agentic AI workflow system integrating Microsoft 365 Copilot and Google AgentSpace to automate decision-making across operations and logistics.
- Implemented Robotic Process Automation (RPA) to streamline repetitive banking tasks and reduce manual processing time.
- Designed and maintained interactive Power BI dashboards visualizing loan portfolio performance, customer engagement, and sales KPIs to support data-driven management decisions.

Cyber Security Consultant Intern Jun 2024 – Sept 2024
KPMG *Jakarta, Indonesia*
- Conducted data-driven cybersecurity maturity assessments, analyzing structured risk datasets to identify critical vulnerabilities and provide actionable, data-informed recommendations to enhance fraud detection and risk management.
- Produced detailed reports and visual dashboards to communicate trends in cybersecurity metrics, translating complex technical findings into actionable strategies for executive stakeholders.

PROJECTS

Inside the OR: Interactive Surgical Outcome Explorer | Live Site Jun 2025
- Designed a D3.js-based scrollytelling dashboard to visualize 6,000+ real-world surgeries from the VitalDB dataset.
- Implemented dynamic filters, tooltips, and time-series visualizations to compare surgical approaches and patient outcomes.
- Collaborated on data wrangling, chart rendering, and user interaction logic using JavaScript, HTML/CSS, and JSON.

Predictive Model for Damage Caused by Wildfires | Live Demo Mar 2025
- Conducted EDA on 130k+ wildfire damage records and engineered key structural and property features.
- Utilized data preprocessing pipelines for consistent scaling, encoding, and imputation.
- Developed and tuned ML models using GridSearchCV, with XGBoost emerging as the top performer

Healthy Recipe Analysis | Project Link Dec 2024
- Engineered features to enhance the dataset's predictive power, such as nutrient ratios and ingredient diversity scores.
- Built and optimized machine learning classification models using scikit-learn, achieving an 85% accuracy rate through iterative hyperparameter tuning.

TECHNICAL SKILLS

Languages: Python, SQL, Java, JavaScript, HTML/CSS, TypeScript
Data Analysis: Exploratory Data Analysis, Hypothesis Testing, Feature Engineering, Predictive Modeling, Regression Analysis, Statistical Validation, KPI Tracking
Frameworks & Tools: pandas, NumPy, scikit-learn, matplotlib, seaborn, Flask, Angular, Dask, AWS S3, AWS Athena, Tableau, Power BI, Excel, Beautiful Soup, regex, Jupyter Notebook, Git/GitHub, PostgreSQL, D3.js, Mapbox, JSON
Design & Visualization: Responsive Web Design, Data Storytelling, User Interaction Design, SVG Graphics