

Stephanie Patricia Anshell

sanshell@ucsd.edu | (925) 348-7120 | [linkedin.com/in/stephaniepatriciaanshell](https://www.linkedin.com/in/stephaniepatriciaanshell) | github.com/stephaniepatriciaans | stephaniepatriciaans.github.io/portofolio

Data Science student at UC San Diego with Finance minor, specializing in machine learning, forecasting, and data-driven decision support across real-estate, risk, and operations.

EDUCATION

UNIVERSITY OF CALIFORNIA SAN DIEGO

Bachelor of Science in Data Science and Minor in Finance

La Jolla, California, USA

Expected 2026

UNIVERSITY OF CALIFORNIA BERKELEY EXTENSION

Certificate Data Analytics, Data Analytics Bootcamp

Berkeley, California, USA

November 2023

SKILLS

Tools / Skills: Python, SQL, NoSQL, Tableau, Jupyter Notebook, Microsoft Excel, JavaScript, MATLAB, C++, Streamlit

Frameworks / Libraries: Pandas, NumPy, scikit-learn, TensorFlow, SciPy, Tableau, Jupyter, Streamlit, GeoPandas, PySAL

Methods: Machine Learning, Regression, Forecasting, Time-Series Analysis, Statistical Modeling, Spatial Analytics, Data Visualization

Languages: English (fluent), Indonesian (native)

EXPERIENCE

IDX Exchange (*Real Estate Data Science Platform*)

Data Scientist Internship

Remote, USA

September 2025 - December 2025

- Built and evaluated machine learning models to predict U.S. residential property prices using large-scale listing data from Cotality's Trestle API.
- Designed an end-to-end pipeline spanning data extraction, cleaning, feature engineering, model tuning, and evaluation.
- Optimized regression models including Random Forest and XGBoost, achieving $R^2 \approx 0.89$ and $MDAPE \approx 6\%$ on held-out data.
- Produced reproducible Jupyter workflows and model diagnostics to identify major drivers of housing valuation and support iterative model improvement.

GAF (*Authorized distributor of GAF Roofing materials in Indonesia, focused on supply chain operations.*)

Data Scientist Internship

Jakarta, Indonesia

January 2024 - September 2024

- Developed ARIMA and decision-tree forecasting models that improved demand forecast accuracy by **15%** and reduced inventory stock-outs by **20%**.
- Built Tableau KPI dashboards used to monitor supply-chain performance and support faster operational decision-making.
- Analyzed sales and inventory patterns to improve planning accuracy and reduce avoidable supply disruptions.

STAL Cooperation (*Printing supplier for hospitals/clinics*)

Accounting & Data Analyst Intern

Jakarta, Indonesia

Jul 2021 - Aug 2022; Jun 2023 - Aug 2023

- Built ARIMA-based forecasting models that reduced material procurement costs by **15%**, improved planning accuracy by **10%**.
- Digitized expense and procurement tracking workflows, improving financial visibility and supporting more accurate budgeting.
- Combined historical sales and external market indicators to strengthen demand planning and purchasing decisions.

PROJECTS (Full portfolio: stephaniepatriciaans.github.io/portofolio/projects/)

Home Price Predictions - Real Estate ML Pipeline | [GitHub: github.com/stephaniepatriciaans/Home_Price_Predictions](https://github.com/stephaniepatriciaans/Home_Price_Predictions)

Python, Pandas, Scikit-learn, Matplotlib, REST APIs, Git, Tableau

- Built a machine learning pipeline to predict residential property values using **80k+ housing records**.
- Cleaned and transformed structured real-estate data, engineered interaction features, and compared multiple regression models.
- Achieved $R^2 \approx 0.89$ and $MDAPE \approx 6\%$, with diagnostics used to explain the strongest pricing drivers.

Loan Charge Offs - Credit Risk Modeling | [GitHub: github.com/stephaniepatriciaans/Project-Loan-Charge-Offs](https://github.com/stephaniepatriciaans/Project-Loan-Charge-Offs)

Python, scikit-learn, TensorFlow

- Developed Random Forest and neural-network models to classify credit-card charge-off risk from borrower-level data.
- Achieved **94.9% accuracy**, outperforming a baseline logistic model by roughly **6 percentage points**.
- Built an interactive analysis workflow to inspect threshold trade-offs and model misclassifications.

Urban Transit and Crime Analysis - Data Integration + Spatial Analytics | [GitHub :](https://github.com/stephaniepatriciaans/Urban-Transit-and-Crime)

github.com/stephaniepatriciaans/Urban-Transit-and-Crime

Python, seaborn, GeoPandas, PySAL, data analysis | Team project

- Analyzed **1M+ NYC and Chicago records** to study relationships among transit access, crime, and socioeconomic conditions.
- Used **geographically weighted regression** and spatial visualization to identify localized hotspots and public-safety inequality patterns.

U.S. Bank Failures Across Recessions - Financial Data Dashboard | github.com/stephaniepatriciaans/Bank-Failures

SQL, JavaScript, regression analysis

- Built an interactive dashboard analyzing **570+ FDIC bank failures** alongside regional economic indicators.
- Used regression-based analysis to examine how bank failures related to regional economic vulnerability and recovery patterns.