

Lance Penaflor

850-686-2389 | penaflor.lance11@gmail.com | LinkedIn | Portfolio | GitHub

EDUCATION

University of West Florida

Bachelor of Science in Computer Science; Minor in Statistics; Certificate in Data Science

Pensacola, FL

May 2027

EXPERIENCE

Data Systems Engineer Intern Andrews Institute

Sep. 2025 – Jun. 2026

- Engineered **R Shiny applications** for athlete performance and injury data analysis, providing interactive dashboards and data-entry modules used by coaches and medical staff.
- Built **dashboards** summarizing athlete statistics, test results, and recovery trends from **SQL-based databases**.
- Designed optimized **SQL schema** in **Azure SQL** to ensure scalability and accessibility.
- Streamlined **data pipelines** for automatic collection, and visualization of athlete data, reducing manual entry time.
- Implemented **role-based data access controls**, protecting sensitive athlete information in compliance with HIPAA guidelines.

Undergraduate Teaching Assistant University of West Florida

Jan. 2025 – Present

- Supported **120+** students across C++ Algorithms and Statistics courses, boosting performance through targeted guidance.
- Replaced **calculator workflows** with **Python labs**, improving reproducibility and aligning with industry practices.
- Co-Produced accessible course materials meeting **100% compliance** with disability mandates.

Machine Learning & Statistics Research Assistant UWF CSDA Lab

Oct. 2024 – Present

- Implemented a data pipeline on mass spectrometry data (**75% accuracy, 80% specificity**) using Python libraries; **Pandas, scikit-learn, NumPy, and Matplotlib**.
- Reduced data dimensionality by **99.8%** with wavelet transforms and PCA, enhancing clarity and processing efficiency.
- Delivered model evaluation metrics (accuracy, precision, recall, F1) to guide algorithm selection for research reporting by visualizing results.

Data Science Educator University of West Florida

Aug. 2024 – Present

- Spearheaded a **35+ member Python/data science community**, lowering barriers to Data Science.
- Created **10+ interactive Jupyter notebooks** for Python, statistics, and ML instruction.
- Led weekly coding sessions (10+ attendees) to explain technical concepts to beginner students.

PROJECTS

Music & Fashion Analysis (In Progress) | Python, Pandas, Scikit-learn, ETL

- Built ETL pipeline on **3 Kaggle datasets**, cleaning duplicates, removing nulls, and filtering using **Pandas**.
- Feature engineering by integrated **LastFM API** to enrich Spotify data with genre tags for all unique songs.
- Performed Exploratory Data Analysis on data using **matplotlib** to spot trends and gain insight on data.

Data Science Learning Website | React, Node.js, TypeScript, Express

- Deployed browser-based coding + AI tutor platform with **50+ users** using **React** and **Express**.
- Directed a **3-person full-stack team** to deploy and maintain a browser-based platform using **React/Express**, implementing CI/CD via Vercel/GitHub for stability.

Embedding Data Preprocessing for LLM Interpretability (In Progress) | Python, PyTorch, NumPy

- Researching wavelet-based embedding perturbations preserving semantics in GPT-2 outputs, advancing robustness.
- Compared logits and distributions to quantify differences, contributing to **LLM interpretability and safety**.

TECHNICAL SKILLS

Programming: Python, SQL, R, C++, JavaScript, TypeScript, Java

Data Tools: NumPy, Pandas (ETL, EDA), Scikit-learn, PyTorch, TensorFlow, Matplotlib, R Shiny

Data Concepts: Data Analysis, Data Visualization, Machine Learning, Statistics, Communication

Web: React, Node.js, Express, Docker, Git, Agile Development, REST API, Azure Storage, MySQL