Dave Austin, Ph.D.

Charleston, SC | austindi1133@gmail.com | LinkedIn | GitHub

Summary

Data Scientist with a Ph.D. in Computational Physics applying scientific methods, statistical reasoning, and large-scale data analysis to real-world systems. Experienced in designing analytical workflows, building reproducible datasets, and translating complex quantitative results into actionable insights. Combines modeling, experimentation, and data engineering to enable reliable decision-making in research and production environments.

Technical Skills

Programming: Python, SQL, Bash
Data Science: Statistical modeling, regression analysis, hypothesis testing, experimental design, feature engineering, uncertainty analysis
Machine Learning: Predictive modeling, classification, clustering, optimization workflows
Data Platforms: Snowflake, PostgreSQL, Parquet, JSON
Visualization: Matplotlib, scientific visualization, analytical reporting
Cloud & Tools: AWS (S3, EC2), Docker, Git, Linux
Scientific Computing: VASP, Quantum ESPRESSO, HPC environments

Experience

Data Engineer / Software Engineer Mar 2025 – Present
Bear Cognition

- Built analytical data pipelines transforming operational and logistics datasets into structured environments supporting reporting and decision-making.
- Developed monitoring systems identifying anomalous pipeline behavior and summarizing operational risks using automated analysis workflows.
- Worked with noisy real-world datasets requiring validation, normalization, and feature extraction prior to downstream analytics.
- Collaborated with stakeholders to translate business questions into measurable data outputs and reproducible analysis workflows.

Postdoctoral Scholar — Computational Modeling & Data Analysis Aug 2024 – Mar 2025
University of Central Florida

- Led independent computational research analyzing electronic and catalytic behavior in complex material systems using large-scale simulation datasets.
- Designed reproducible data analysis pipelines transforming raw simulation outputs into structured datasets suitable for quantitative comparison and interpretation.
- Performed statistical and comparative analysis across multiple configurations to identify trends linking atomic-scale structure to experimentally observable properties.
- Built tooling to extract physical features including energy landscapes, charge distributions, and electronic structure metrics for downstream analysis.
- Collaborated closely with experimental researchers to align computational predictions with measured results, translating analytical findings into actionable scientific insight.
- Managed multi-terabyte datasets generated across national and local HPC facilities, emphasizing reproducibility, traceability, and consistent data organization.

- Mentored graduate researchers on workflow design, data interpretation, and reproducible computational practices.

Graduate Research Assistant — Data Engineering & Scientific Computing Aug 2018 – Aug 2024
University of Central Florida

- Designed and operated automated computational workflows supporting large-scale density functional theory (DFT) studies across national high-performance computing systems (Stampede2/3, Perlmutter, Anvil).
- Built Python-based orchestration tooling to submit, monitor, restart, and validate simulation jobs through SLURM, enabling reliable execution of dozens of concurrent long-running computations.
- Implemented retry and recovery logic for failed or timed-out simulations, reducing manual monitoring and saving days of compute time per project.
- Developed automated data movement pipelines using SSH/SCP to transfer multi-GB simulation outputs for downstream analysis and visualization.
- Parsed complex scientific output files (wavefunctions, charge densities, energies, forces, atomic positions) into structured datasets stored as JSON and CSV for reproducible analysis.
- Created analysis pipelines extracting physical features including density of states, band structures, minimum-energy configurations, and electron localization metrics.
- Built reproducibility tooling capable of regenerating figures and analysis results directly from structured metadata and configuration tracking.
- Managed hierarchical experiment datasets consisting of hundreds of simulation configurations, maintaining traceability across parameter variations and restart trajectories.
- Developed simulation-to-observation tooling that generated STM and STS image predictions from electronic structure data, bridging computational outputs with experimental measurements.
- Collaborated internationally with experimental researchers to validate computational predictions against real measurements, ensuring analytical consistency and data integrity.
- Supported eight peer-reviewed publications through large-scale data generation, analysis, and computational infrastructure development.

Teaching Assistant Aug 2019 – May 2020
University of Central Florida

- Taught undergraduate physics laboratories emphasizing experimental analysis, uncertainty estimation, and quantitative reasoning.
- Guided students in interpreting real measurement data and validating scientific conclusions.

Education

Ph.D. in Physics (Computational Materials Science) 2024
University of Central Florida
Dissertation: First Principles Studies of Nano-Scale Phenomena At Surfaces: From Characteristics of Single Atom Catalysts to Molecular Structure Formation

B.S. in Physics 2018
College of Charleston

Research Impact

8 peer-reviewed publications in computational physics and materials science demonstrating advanced modeling, large-scale data analysis, and interdisciplinary collaboration.