C. Morgan Davis

Active Secret Clearance | Data Engineer | Python, SQL, ETL | Engineering Systems Background

Phone: 1-415-729-5387
Email: cmdavis25@proton.me
GitHub: github.com/cmdavis25
Website: cmdavis25.github.io

Summary

Data professional with 16+ years building quantitative models, pipelines, and analytical capabilities across aerospace/defense, medical devices, automotive, and consumer products sectors. Engineering background provides domain expertise in simulation data, manufacturing cost modeling, and industrial process analysis. Focus on Monte Carlo simulation, statistical modeling, and machine learning applications. Pursuing a Master of Data Analytics degree at Penn State.

Top clients include Air Force Research Laboratory (AFRL), NASA, Northrop Grumman, Delta Faucet, Rivian, & Mercedes-Benz.

Skills

Data Engineering: SQL, Python (Pandas/NumPy), R, Minitab, Excel, VBA, MATLAB, Bash, Slurm, PBS, ETL Pipelines
Data Science / Analytics: Regression Analysis, Hypothesis Testing, EDA, NLP, PCA, scikit-learn, NetworkX, Monte Carlo Simulation
Visualization / Dashboards: Power BI (DAX), Tableau, Matplotlib, Seaborn, Plotly, ggplot2
DevOps / Version Control: Git, GitHub, GitLab

Professional Experience

Engineer - Computational Physical Sciences

Applied Research Associates, Inc.
Dayton, OH
10/2023 to 10/2025

- Designed and automated Python- and Bash-based data pipelines to productionize electromagnetic simulation workflows across distributed HPC environments, orchestrating mesh generation (CUBIT) and RF simulation runs across up to 32 nodes on an HPE Cray EX supercomputing cluster - with peak memory utilization of 490 GB per node.
- Parsed and analyzed large simulation output datasets to evaluate RF performance characteristics, visualized scalability metrics using Matplotlib, and reported findings in periodic PowerPoint presentations to inform defense research decisions.
- Supported model validation and comparative analysis across simulation platforms (CREATE-RF SENTRi, Xpatch, FEKO).

PDM System Administrator / Engineer

Aerobiotix, LLC
Miamisburg, OH
06/2021 to 08/2023

- Designed and administered a SQL-backed data governance system (PDM vault), establishing data standards, metadata protocols, and 6 automated workflows across 3 departments: Engineering, Quality Assurance, and Operations.
- Developed custom SQL queries and reporting tools to extract and analyze product lifecycle data.
- Automated engineering processes through scripted macros and workflow logic, reducing repetitive manual tasks and improving data integrity.
- Conducted structured root cause analyses using statistical methods and modeling to guide product improvements.
- Collaborated cross-functionally (engineering, quality, manufacturing) to translate operational data into actionable product decisions.

Senior Engineer - Sales and Design

Rack Processing Company

Moraine, OH

02/2019 to 06/2021

- Built Python- and Excel-based cost estimation models using multivariate regression to map product feature inputs to pricing outputs, reducing quote turnaround time from ~48 hours to 10 minutes (~99% reduction).
- Partnered with Sales and Operations to analyze historical product and sales data to improve pricing accuracy and margins.
- Managed $1.5M+ in annual revenue accounts, leveraging data-driven pricing strategies to optimize profitability.

Projects

InstaCart Market Basket Analysis (Penn State - Data Mining - Spring 2026)

Tech: Python, SQLite3, pandas, numpy, matplotlib, mlxtend

Leading a team analysis of 33M+ grocery purchase records across 200K+ users to predict customer reorder behavior. Defined scope, assigned tasks, and coordinated execution across data ingestion, EDA, and association rule mining (Apriori, FP-Growth) phases. Developed a frequency-based one-hot encoding strategy to reduce memory requirements from 480+ GB to under 32 GB across 49K+ unique product IDs, identifying association rules with confidence up to 0.33 and lift up to 3.05.
Ongoing: clustering, PCA, and supervised learning for next-basket prediction.

Ganttoro - Probabilistic Project Scheduling Tool (In Development)

Tech: Python, PySide6, NumPy, SciPy, NetworkX, Matplotlib, PyQtGraph

Desktop application that replaces deterministic Gantt charts with Monte Carlo simulation to model real-world schedule uncertainty and optimize project planning decisions. Runs 10K+ vectorized simulations in seconds from user-defined duration distributions (triangular, PERT, normal) and dependency chains, producing P10/P50/P90 completion estimates and box plot timelines.

GitHub: github.com/cmdavis25/Ganttoro

OneTaskAtATime — Intelligent Task Manager (Beta Release)

Tech: Python, SQL, PyQt5, Pytest

Desktop productivity app using an Elo ranking algorithm to resolve priority conflicts and surface the single highest-importance actionable task. Features dependency tracking, deferral trend analysis, and configurable Windows notifications. Codebase comprises 31K+ lines of application code and 28K+ lines of test code, reflecting production-level software engineering standards.

GitHub: github.com/cmdavis25/OneTaskAtATime

Education

Pennsylvania State University

Master of Data Analytics (Big Data Systems Option)

01/2026 to 05/2028 (Enrolled)

University of Maryland, Baltimore County

Bachelor of Science, Mechanical Engineering

01/2005 to 05/2009 (Graduated)

Honors: Tau Beta Pi Engineering Honor Society

Certifications

- IBM - Data Science Professional Specialization
- University of Colorado Boulder – Excel/VBA for Creative Problem Solving Specialization
- Microsoft - Data Analysis and Visualization with Power BI | Intermediate DAX & Time Intelligence
- Tableau Learning Partner - Advanced Data Visualization with Tableau