Caroline Smith

Data Engineer

—

Contact

carolinesmith527@gmail.com

—

Technology

Python, Databricks, Scikit-learn, Pytorch, Pyspark, AWS, Azure, Pandas, Geopandas, Numpy, LLMs, ArcGIS, GIS, R, SQL, SAP, Oracle, Stata, Tableau, PowerBI, AutoCAD, Sketch Up

—

Skills

Machine Learning, Data Science, Data Engineering, Generative AI, Data Modeling, Mathematical Programming & Modeling, Entity Resolution, Application Design and Build, Data Analysis, Spatial Analysis, Drafting, Computer Modeling

Employment History

Shields Health Solutions

Software Engineer JULY 2025 - CURRENT
- Built a comprehensive, automated regression test suite using Python, Pytest, and Playwright to ensure application stability and quality.
- Extended data processing ETLs to integrate new healthcare partners.

Senior Analytics Engineer JULY 2024 - JULY 2025
- Architected and developed end-to-end data pipelines including the DDLs, PL/SQL, bash scripting, and cron scheduling.
- Optimized SQL performance, resolved complex bugs, and added new features to Python and SQL based data processing pipelines.
- Led expansion of Patient Risk Management data science project to new patient population; refactored training datasets to Python and advised on ML best practices.

Accenture

Data Science and Engineering Specialist NOV 2022 - JUNE 2024, BOSTON
- Designed an AWS-based ML solution to automate Master Data Management (MDM) for a major biopharma client.
- Built Python/SQL tools to automate data quality checks and resolve backlogs.
- Developed SAP calculation views to merge cross-system data and apply transformation rules.
- Developed Python scripts for data enrichment via webscraping and APIs.

Data Science and Engineering Senior Analyst MAY 2021 - OCT 2022, BOSTON
- Developed text processing algorithms and toolkits to perform entity resolution and clustering across multiple systems using Python and SQL.
- Created Python and SQL data quality analyses to resolve compliance and integration issues, ensuring continued data quality at healthcare and life sciences clients.
- Executed ad-hoc statistical analyses to drive research & data-driven decision-making.

Data Analyst OCT 2019 - APRIL 2021, LOS ANGELES
- Performed QA testing and analysis for a large-scale SAP implementation focused on renewable energy "Net Metering" data.
- Authored technical documentation to facilitate client education.

Otis College of Art & Design Office of Institutional Research / Data Analyst
OCT 2018 - JUNE 2019, LOS ANGELES
- Performed data analysis in SQL to support strategic planning & institutional research.

University of Kentucky Center for Health Services Research / Data Analyst Intern
SEP 2017 - JUNE 2018, LEXINGTON KY
- Performed data analysis and visualization in SQL and Tableau to support healthcare research.

Education

University of Kentucky / Mathematical Economics B.S. / 2014 - 2018, LEXINGTON KY
Minor in Dance
Summa Cum Laude / Departmental Honors / 3.975 GPA

Udemy / ZTM Complete A.I. & Machine Learning, Data Science Bootcamp / 2024