

Michelle D. Davies

Greater Boston, MA
[Portfolio](#)
[LinkedIn](#)
[Credly](#)

Education

University of California, Berkeley – Master of Information & Data Science, 2025

Cornell University, College of Engineering – Bachelor of Science in Electrical and Computer Engineering, 2022

Experience

[IBM, Massachusetts – May 2021 - Present](#)

Senior Data Engineer; Cambridge, MA – April 2024 - Present

- Designed, developed, and deployed business process management applications using Appian.
- Designed, developed, and deployed ETL pipeline applications using IBM DataStage.
- Created intuitive and user-friendly interfaces for BPM applications to facilitate easy navigation and use by non-technical users.
- Designed and implemented efficient data ingestion pipelines to collect data from various sources, ensuring the scalability and reliability of data storage solutions.
- Implemented security measures to protect data integrity and privacy, ensuring compliance with data protection regulations.
- Worked closely with business analysts, data scientists, and IT teams to define data requirements and deliver actionable insights.
- Continuously monitored data pipelines and performed optimizations to enhance performance and reduce costs.

Data Engineer & BPM Application Developer; Cambridge, MA – October 2023 - April 2024

- Designed, developed, and deployed business process management applications using Appian.
- Designed, developed, and deployed ETL pipeline applications using IBM DataStage.
- Created intuitive and user-friendly interfaces for BPM applications to facilitate easy navigation and use by non-technical users.
- Developed and optimized data processing pipelines using big data technologies to cleanse, structure, and enrich raw data into usable formats.
- Customized BPM software according to business needs, including configuring process models, rules, and notifications.
- Integrate new technologies and data management tools into the existing infrastructure to improve data processing and analytics capabilities.

Associate Data Scientist; Cambridge, MA – October 2022 - October 2023

- Wrote programs to cleanse and integrate data in an efficient and reusable manner using leading-edge and open-source tools such as Python, R, and TensorFlow, combined with IBM tools and our AI application suites.
- Worked in an Agile, collaborative environment, partnering with other scientists, engineers, consultants, and database administrators of all backgrounds and disciplines to bring analytical rigor and statistical methods to the challenges of predicting behaviors.
- Communicated with internal and external clients to understand and define business needs and appropriate modeling techniques to provide analytical solutions.
- Assisted in leading workshops to gather insight from users about how to improve applications.
- Evaluated modeling results and communicated the results to public sector audiences.

Client-Facing Technical Support CO-OP; Littleton & Lowell, MA – May 2021 - December 2021

- Resolved support requests and debugged issues with the IBM InfoSphere Information Server (IIS) product.

Specialized Skills

- Machine Learning
- Business Process Management (BPM)
- Government, Human Services & Healthcare Industry Experience
- Research
- Product Engineering

Projects

RealPage Rent Impact, University of California-Berkeley Capstone

Skills: SQL · Python · Pandas · Data Engineering · Business Intelligence

Analyzed 100M+ rental listings with SQL and Python to uncover pricing anomalies linked to RealPage's rent-optimization algorithm.

Context-Enriched NER for Identifying Emerging Trends in Video Comments, University of California-Berkeley

Skills: NLP · Deep Learning · Transformers (BERT) · Python · PyTorch

Fine-tuned BERT models to extract emerging trends and locations from YouTube comments using NLP and PyTorch.

Reidentification Project, University of California-Berkeley

Skills: Privacy Engineering · Python · Data Privacy · Feature Engineering · Machine Learning

Explored privacy vulnerabilities in de-identified data using ML and clustering to model real-world re-identification risks.

Machine Learning Systems Engineering: Deploying an API, University of California-Berkeley

Skills: ML Ops · Docker · FastAPI · Model Deployment · Cloud (AWS)

Built and deployed a containerized ML prediction API with FastAPI, Docker, and cloud infrastructure.

Non-Linear Regression for Housing Prices, University of California-Berkeley

Skills: Regression Analysis · scikit-learn · XGBoost · Feature Engineering · Model Evaluation

Compared linear vs. non-linear models (e.g. XGBoost) to boost R² on real estate datasets.

Licenses and Certifications

- IBM Data Engineering Specialty - Experienced
- Appian Certified Senior Developer
- Neo4j Graph Data Science Certification
- Azure AI Fundamentals Certification
- Azure AI Engineer Associate Certification
- AWS Partner Accreditation (Technical)
- Databricks Generative AI Fundamentals Accreditation
- IBM Information Server - Datastage ETL Developer
- IBM Fundamentals of Quantum Algorithms
- IBM Quantum Machine Learning