Data Engineer Intern

Michael Johnson

📞 +1 (555) 234-5678

✉️ michael.johnson@email.com

📍 New York, NY

Professional summary

Detail-oriented and analytical Data Engineer Intern seeking to contribute technical expertise in data processing, pipeline design, and big data technologies. Ready to apply academic knowledge of data engineering in a professional environment and further develop skills at a leading data-driven company.

Links

🌐 LinkedIn: /in/michaeljohnson

Experience

Data Engineer Intern

June 2025 - August 2025

NextGen Analytics / New York, NY

- Assisted in building and deploying ETL workflows for data processing pipelines.
- Worked with data scientists to prepare datasets for machine learning model training.
- Performed wrangling and preprocessing tasks to clean and normalize raw data.
- Supported the integration of data sources from various platforms, including AWS and Google Cloud.

Software Engineering Intern

June 2023 - August 2023

TechSolve Innovations / New York, NY

- Contributed to the development of backend services using Python for data processing tasks.
- Assisted in integrating third-party APIs for real-time data collection.
- Participated in the creation of user-facing dashboards to visualize key metrics using Python libraries.

Awards and Achievements

NYU Dean's List

New York University – New York, NY

Fall 2024, Spring 2025

- Recognized for academic excellence in the Computer Science program with a GPA above 3.8.

Projects

Real-time Data Pipeline

Developed a real-time pipeline using Apache Kafka and AWS Lambda for processing streaming data from various sensors in an IoT environment. The pipeline included data ingestion, processing, and storage in a NoSQL database.

Education

Bachelor of Science in Computer Science

New York University, NY

- Expected Graduation: May 2026
- Relevant Coursework: Data Engineering, Data Structures, Distributed Systems, Cloud Computing

Skills

Skill	Level
Python (Pandas, NumPy)	◆◆◆◆◆
SQL (PostgreSQL-, SQLite)	◆◆◆◆◆
Data Visualization (Matplotlib, Seaborn)	◆◆◆◆◆
Apache Hadoop	◆◆◆◆◆
Cloud Computing (AWS, Google)	◆◆◆◆◆
Data Cleaning and Transformation	◆◆◆◆◆
Git	◆◆◆◆◆
Jupyter Notebooks	◆◆◆◆◆
Data Warehousing (Snowflake, Google BigQuery)	◆◆◆◆◆
RESTful APIs	◆◆◆◆◆