

SRAVAN KUMAR NEETIKONDA

+1-314-439-6850 | neetikondasravan@gmail.com | [LinkedIn](#)
UNITED STATES

PROFESSIONAL SUMMARY

Data Engineer with 5 years of production experience designing and operating reliable Python- and SQL-based data pipelines supporting daily analytics and operational decision-making. Strong background in incremental processing, data quality validation, and delivering analytics-ready datasets to business and technical teams.

TECHNICAL SKILLS

Programming Languages: Python, SQL, PL/SQL, Bash
Data Processing: Apache Spark (PySpark, Spark SQL), Hadoop (HDFS, Hive)
Cloud & Storage: AWS (S3, EMR, Athena)
Databases: Oracle, PostgreSQL, MySQL, Teradata
Data Engineering: Data Modeling (Star & Snowflake), Batch ETL, Incremental & Idempotent loads, Data Quality Validation & Reconciliation
Data Formats: Parquet, CSV, JSON
Data Architectures: Transactional, Analytical, and Time-Series-Oriented Batch Data
Analytics: Power BI, Tableau
Scheduling & Ops: Linux/Unix, cron-based scheduling, job monitoring, Bash automation
Dev & Process: Git, Scrum, Jira, ServiceNow

PROFESSIONAL EXPERIENCE

Data Engineer
eBay Inc.

Feb 2024 – Present

- Designed and maintained Spark-based batch data pipelines using Python (PySpark) and Spark SQL to process large, multi-source transactional and behavioral datasets for daily analytics and operational reporting.
- Translated business and analytics requirements into curated, analytics-ready datasets consumed by **10+ analysts and business users across 5–7 teams**, supporting recurring decision workflows.
- Implemented **incremental and idempotent processing patterns**, reducing reliance on full reloads and improving pipeline stability and data freshness for time-sensitive reporting.
- Introduced automated **data quality checks** (row counts, schema validation, null thresholds), enabling earlier detection of upstream data issues and reducing recurring production incidents.
- Standardized SQL and Python transformation logic across pipelines, improving maintainability, reducing duplicated logic, and simplifying issue diagnosis during production support.
- Operated pipelines across **hybrid data architectures** (HDFS-backed Hadoop and AWS S3/EMR/Athena), gaining experience working with diverse storage and processing environments.
- Supported production workloads using Linux/Unix and Bash, including job monitoring, failure investigation, and lightweight automation.
- Collaborated with analysts and engineers in Agile Scrum teams, managing work through Jira and resolving production issues via ServiceNow.

Data Engineer
Oracle Corporation

Jul 2021 – Dec 2022

- Designed and supported SQL- and PL/SQL-based ETL workflows ingesting data from multiple enterprise source systems into analytical data stores for recurring business and operational reporting.
- Developed and optimized complex SQL transformations, views, and stored procedures supporting **10–15 recurring datasets and reports** used by business and operations teams.
- Improved batch job predictability through query refactoring, indexing, and execution plan analysis, reducing runtime variability for **time-sensitive daily reporting jobs**.

- Built incremental data loading processes using SQL, PL/SQL, and Python, improving data availability within defined reporting windows.
- Implemented data validation and reconciliation checks, reducing reporting discrepancies and shortening investigation cycles for **critical business reports**.
- Delivered curated datasets powering **Power BI and Tableau dashboards**, enabling business users to identify trends and monitor performance metrics.

Data Engineer
Verizon Communications Inc

May 2019 – Jul 2021

- Built and maintained SQL-based transformations supporting recurring operational and business reports across large, high-row-count datasets.
- Supported daily and weekly ETL workflows using SQL and Python to clean, validate, and transform operational data for analytics consumption.
- Delivered Power BI and Tableau dashboards used by **15–25 regular business users** to monitor operational performance and trends.
- Automated routine data checks and refresh tasks using Linux/Unix and Bash, improving reliability of recurring data deliveries.
- Performed data validation and reconciliation to improve confidence in reported metrics used by business stakeholders.

EDUCATION

Saint Louis University
Master of Science in Data Science & Analytics

St. Louis, MO
Jan 2023 – Dec 2024

ACADEMIC PROJECTS

Data-Driven Firewall Traffic Analysis

- Captured and analyzed **packet-level network traffic** using Wireshark; processed raw packet data with **Python** to create structured datasets and traffic summaries.
- Analyzed protocol behavior and traffic patterns to identify anomalies and usage trends.

HCI User Interaction Analytics

- Designed an **event-logging system** to capture user interaction metrics such as task completion time, error rates, and dwell time.
- Analyzed interaction data using **Python and Pandas**, Leading to an estimated **~30% improvement in task efficiency** during testing.

NLP-Powered Text Processing Application

- Built a **Python-based NLP application** supporting text summarization, information extraction, OCR-based text retrieval, and text-to-speech.
- Demonstrated end-to-end processing of unstructured text data.