# PRAVEEN KUMAR SIDDELA

480-919-5206 | praveenkumarsiddelaasu@gmail.com | https://www.linkedin.com/in/praveenkumarsiddela/ | [Portfolio](Portfolio)

## SUMMARY

Data Engineer with 2+ years of experience in building ETL pipelines, optimizing data workflows, and migrating on-prem systems to AWS/Azure. Proficient in SQL, Python, Apache Spark, and cloud platforms.

## TECHNICAL SKILLS

Languages: Python, SQL, Java, Bash
Big Data Tools: Apache Spark, Hadoop, Kafka, Airflow
Cloud Platforms: AWS (Glue, S3, Redshift), Azure (Data Factory), GCP
Databases: MySQL, PostgreSQL, MongoDB, SQL Server
ETL/BI Tools: AWS Glue, SSIS, SSRS, Tableau, Power BI
Web Scraping: Selenium, BeautifulSoup

## WORK EXPERIENCE

**Data Engineer Intern,** Xtrium, Austin, Texas.                                  January 2025 - Present
- Increased data retrieval speed by 30% by building a Python/Selenium web scraper, extracting 10K+ daily records for AI model training.
- Designed SQL-based relational databases, enhancing structured data storage for 10+ modules AI models.
- Developed Tableau dashboards to track and analyze data validation results, reducing reporting errors by 30%.
- Proposed and implemented an automated data validation with SQL and Python, improving accuracy by 30% and reducing manual effort.

**Data Engineer,** Accenture, Hyderabad, India.                            December 2020 - December 2022
- Engineered complex SQL queries, stored procedures, triggers, and functions, improving business processes for 20+ departments.
- Ensured 95% accuracy and quick health program eligibility determination by designing and optimizing SQL procedures with Apache Spark.
- Supervised and maintained 50+ SSIS packages for automated data integration tasks and enhanced reporting capabilities by deploying 25+ comprehensive reports operating SSRS.
- Led API efficiency during seamless data migration processes to Azure Cloud, resulting in a 30% reduction in data transfer times and a 20% increase in overall system performance.

**Data Engineer Intern,** Netcon Technologies, Hyderabad, India.              January 2020 - June 2020
- Collaborated with cross-functional teams to analyze data using SQL and Python, supporting strategy development for 10+ business units.
- Facilitated regular meetings with senior management to present insights from 15+ Tableau dashboards, offering real-time insights and fostering informed decision-making.

## PROJECTS

**Spotify Data Pipeline Model**                                              January 2024 - May 2024
- Built an AWS serverless data pipeline (Lambda, Glue, S3) for Spotify, cutting down data ingestion latency by 40% and enabling real-time analytics.
- Orchestrated a computerized ETL pipeline with AWS Glue. Enabled efficient data extraction, cleansing, and storage in an S3 data lake. Integrated AWS Glue Crawler and Athena for dynamic schema generation and querying.

**Insurance Policy Data Management System**                                January 2023 - April 2023
- Developed a comprehensive relational database management system with MySQL to manage insurance policy and claim data efficiently.
- Created an interactive policy sales dashboard in Tableau for real-time sales monitoring and analysis.

## EDUCATION

**Master of Science in Information Technology**                                       December 2024
Arizona State University, Tempe, USA.                                                  GPA: 4.0/4.0

**Bachelor of Engineering in Computer Science and Engineering**                        September 2020
Vasavi College of Engineering, Hyderabad, India.                                       GPA: 7.68/10.0

## CERTIFICATIONS

AWS Academy Cloud Foundations (2024), AWS Academy Cloud Security Foundations (2024), Microsoft Certified: Fundamentals of Generative AI (2024), Oracle Cloud Infrastructure Foundations (2023), Microsoft Certified: Power Platforms Fundamentals (2022)