

RAJDEEP REDDY CHARLA

Cincinnati, OH | rajdeepreddy@yahoo.com | 419-378-1526 | [LinkedIn](#)

SUMMARY

Dynamic data engineer with experience developing scalable systems, creating, and maintaining data pipelines, and offering insightful information. Competent in SQL, Tableau, PowerBI, Google Analytics, and Python (Pandas, NumPy, PySpark, SparkSQL). knowledgeable in distributed computing (AWS EMR, Databricks, Spark), database management (DynamoDB, Athena, MS SQL Server, MySQL, Oracle), and data warehousing (Snowflake, AWS Athena, AWS Redshift). Competent with cloud services (AWS, Azure Databricks), software project management (SDLC, Aha, JIRA, A/B testing, Niro), and programming (PowerShell, Python). Shown proficiency in ETL techniques, data visualization, and data integrity, attempting to use these abilities to make creative solutions.

EXPERIENCE

Data Engineer – Nike Inc, Portland, USA

Jan 2020 – Apr 2024

- Developed and supported Nike Global Sales Pipelines, facilitating data consumption for multiple use cases across the organization.
- Designed robust and scalable data integration pipelines using PySpark, optimizing data processing with EMR and EC2 instances, and orchestrated pipelines with Apache Airflow.
- Managed end-to-end data responsibilities including data modeling, ad-hoc analysis, and implementation of algorithms for extracting Sales and Revenue metrics in the AWS environment.
- Created Tableau dashboards for generating sales and revenue reports, enabling strategic decision-making, and orchestrated the process using Airflow.
- Integrated product data feeds from Kafka to Spark processing system and stored order details in Hive databases.
- Improved data integrity and quality assurance through advanced test design and documentation, enabling self-service customer support.
- Implemented Delta tables in Databricks to enhance data reliability, streamline processing workflows, and utilize Delta Lake features for ACID transactions and time travel capabilities.
- Developed materialized views in Databricks to address specific analytical needs and improve query performance for stakeholders.
- Resolved complex data integration challenges through optimized ETL strategies, influencing data-driven opportunities across cross-functional teams.
- Created an email notification service using Pandas for stakeholders requesting data in CSV format.
- Managed code deployment processes through GitHub and Jenkins, establishing robust CI/CD pipelines.

Data Engineer - Carelon, Hyderabad, India

May 2019 – Dec 2019

- Utilized data analysis skills to drive growth initiatives for US-based pharmaceutical companies through insightful data interpretation.
- Developed Spark applications for key operations like data inspection, cleansing, loading, and transformation on extensive datasets.
- Designed and executed ETL pipelines using EMR and Glue, customized for precise business-specific transformations.
- Created data applications and automated Spark pipelines for efficient management of bulk loads and incremental data updates.
- Collaborated with data scientists and stakeholders to refine datasets according to project requirements.
- Implemented robust security models to ensure data quality and compliance with privacy requirements, enhancing governance processes.

LEADERSHIP, AWARDS AND RECOGNITION

- Mentored team members, empowering them with a strong foundation in coding and problem-solving capabilities.
- Recipient of the "Coder of the Month" award for 3 consecutive months at Carelon

SKILLS

- **Data Analysis:** Proficient in Python, Pandas, NumPy, PySpark, and SparkSQL.
- **Data Visualization:** Tableau, PowerBI, and Google.
- **Data Analytics:** Snowflake, AWS Athena, and AWS Redshift, Big Data, Hadoop.
- **Distributed Computing:** Proficient in distributed computing platforms like AWS EMR, Databricks, and Spark.
- **Database Management:** Experienced in managing databases including DynamoDB, Athena, MS SQL Server, MySQL, and Oracle.
- **Programming Languages:** Python, SQL
- **Software Project Management:** SDLC, Aha, JIRA, A/B Testing, Niro.
- **Cloud Services:** Skilled in utilizing cloud services, including AWS (EC2, EMR, S3, RDS, Lambda, Athena, Glue, Aurora) and Azure Databricks.
- **Other Tools:** Proficient in using MS-Excel for data analysis (PowerPivot/DAX, PowerQuery, Optimization), and experienced in version control with GitHub and Bitbucket, as well as using VS Code for development.

EDUCATION

Master of Science in Information System Technologies

Dec 2021

Wilmington University, DE, USA (3.58 GPA)

Relevant Coursework: Data Management, Data Warehouse, Database Management Systems, Project Management

Bachelor of Technology in Electronics & Communications

Apr 2019

Jawaharlal Nehru Technological University, Hyderabad, India (3.12 GPA)

Relevant Coursework: Database Management, Computer Networks, Computer Architecture, Probability and Statistics