

# SAI MAHESWAR REDDY | Data Engineer

+1 (716) 398 6783 | [maheswareddysai@gmail.com](mailto:maheswareddysai@gmail.com) <http://linkedin.com/in/saimaheswar>

## SUMMARY

- Around 4 years of professional Data Engineer experience in Big Data, Hadoop Ecosystem, Cloud Engineering, and Data Warehousing.
- Generated, maintained & analyzed Azure monitoring dashboards, reports, trends, minimizing customer pain points by 20%.
- Implemented ETL process to streamline the import of data from different sources into the BigQuery warehouse.
- Designed and deployed data pipelines using Talend and Informatica, resulting in a 20% improvement in data accuracy.
- Manage data, including developing SQL Server Report Services to produce reusable code and an automated reporting system, and designing user acceptance tests to give clients a chance to offer helpful criticism.

## EDUCATION

University at Buffalo, The State University of New York, New York, USA | Master of Science - Data Science | Dec 2023  
SRM Institute of Science and Technology, India | B. Tech in Electronics & Communication Engineering | Jun 2022

## TECHNICAL SKILLS

**Programming Languages:** Python, SQL, Scala

**Database Systems:** SQL (MySQL, PostgreSQL), Redshift, MongoDB

**ETL Tools:** Apache Spark, Apache Kafka, Talend, Informatica, Apache Airflow, Databricks, Pipelines

**Data Warehousing:** Amazon Redshift, Snowflake, Google BigQuery, Azure SQL Data Warehouse

**Packages:** NumPy, Pandas, Matplotlib, SciPy, ggplot2

**Big Data Technologies:** Hadoop, Apache Spark, Apache HBase, Apache Hive, MapReduce

**Cloud Platforms:** AWS (S3, EC2, Redshift), Azure (SQL Database, Data Lake)

**Machine Learning:** TensorFlow, scikit-learn, MLlib (Spark), Seaborn

**Version Control & Collab. Tools:** Git, SVN, Jira, Confluence, Trello, GitHub, Docker, Kubernetes

**Other Skills:** Documentation, Team Collaboration, Data Cleaning, Data Wrangling, Troubleshooting, Data Warehousing, Data Cleaning, Data Analysis, Data Migration, Data Visualization

## WORK EXPERIENCE

### MetLife, USA | Data Engineer | Jan 2024 – Present

- Spearheaded the implementation of a robust ETL process in collaboration with the operations team, resulting in a 25% increase in data accuracy and enhancing overall reporting efficiency by 40 hours per month.
- Engineered tailored data models through Power BI and Excel for 10 clients, enabling real-time monitoring of performance metrics; improved client satisfaction scores by 25% and facilitated quicker decision-making processes.
- Performed data cleaning including transforming variables and dealing with missing values and ensured data quality, consistency, and integrity using Pandas and NumPy.
- Optimized data workflows using Azure Synapse Analytics, and Azure Data Factory, enhancing processing speed by 50%.
- Established a Hadoop architecture to monitor user interactions in real-time; this setup facilitated the collection of 1M+ data points, enabling the UI team to optimize user experience based on concrete behavioral data.

### Ivive Labs Pvt Ltd, India | Jr. Data Engineer | Aug 2020 – Jul 2022

- Optimized data integration workflows, enhancing quality and consistency with a 40% reduction in load times.
- Implemented Python-based predictive analytics for a 30% improvement in clinical trial prediction accuracy and 15% optimization in vaccine supply chain planning.
- Secured sensitive data and managed cryptographic keys using Azure Key Vault, ensuring data protection and compliance with security standards.
- Developed an automated AWS pipeline solution that processed over 10TB of data per month while reducing operating costs by 30%.
- Pioneered waterfall methodology for building an internal application and SDLC and defect lifecycle management.
- Executed Hadoop ecosystem components (HDFS, YARN, MapReduce) for distributed data processing.
- Crafted and implemented a robust cloud data architecture that increased data retrieval speeds by 40% while simultaneously reducing operational costs by 30%, enabling more efficient data analysis for strategic decision-making.

### Data Engineer Intern | Oct 2019 - Jul 2020

- Created over 10 interactive Power BI dashboards that provided actionable insights into marketing performance; leveraged trend analysis to identify major cost-saving opportunities, resulting in a 20% reduction in marketing expenses.
- Processed Data Analysis, Data Validation, Data Cleansing, Data Verification, identifying data mismatches, Data Import, and Data Export using Microsoft SQL Server Integration Services (SSIS).
- Prompted analysis and interpretation of comprehensive and intricate datasets sourced from Excel, FLAT files, JSON, and SQL Server, leveraging Python programming.
- Streamlined Tableau dashboards by optimizing SQL procedures, cutting query execution time by 20%.
- Conceptualized Agile methodology in developing the application, which included iterative application development, weekly status reports, and stand-up meetings.

## CERTIFICATIONS

AWS cloud practitioner | Google Cloud (Professional Data Engineer) | Microsoft Certified: Azure Data Engineer Professional  
| Databricks Certified Data Engineer Associate