

# Saumya ranjan Sahu

Hyderabad, Telangana, India



sahu.saumya26@gmail.com



7008229203



[linkedin.com/in/saumyaranjansahu](https://www.linkedin.com/in/saumyaranjansahu)

## Summary

Saumya is a data engineer with 6 years of total experience having multiple technical skills and passionate about providing simple solution out of complex problems. He is currently delivering value to customers with his Strategy, vision, and implementation and provided sustainable and effective business solution. He is a thinker, a continuous learner, a technologist.

Skills:

Hadoop | NiFi | Airflow | Hive | Kafka | Apache Spark

Python | SQL | BigQuery

Informatica | SSIS

Power BI

Azure | GCP | SNOWFLAKE

## Experience



### Data Engineer

Micron Technology

Feb 2021 - Present (9 months +)

- Designed and Build automated monitoring and notification systems for Data Pipelines
- Explore ways to optimize cost and time for BigQuery
- Reduce the cost of BigQuery SQL from \$7 per hour to \$1 per hour
- Created an automated Framework to migrate data from BigQuery to MySQL
- Create automated python scripts to create,modify,delete tables, views, Sps etc in BigQuery
- Migrate Data from SQL Server to Big Query using Apache Nifi for Load type Full,CDC,Incremental
- Load Data from Filer in HDFS to Big Query
- Load around 45gb of data from SQL server to Big Query using DataProc
- Load data from Big Query to Cloud SQL using PUB/SUB and Cloud Function
- Create SP in Snowflake to transform data
- Load Data from Kafka topics to Big Query using DataProc
- Load data from SQL Server to Big Query using DataFlow
- Deployment using Azure Devops
- Building APIs to access the data



### Data Engineer

EY

Aug 2019 - Feb 2021 (1 year 7 months)

- Architect, Design, build, Deploy and own all the components of a high volume data warehouse end to end adhering to industry standards and best practices
- Build efficient data models and metadata for ad hoc and pre-built reporting
- Interface with business customers, gathering requirements and delivering complete data & reporting

solutions owning the design, development, and maintenance of ongoing metrics, reports, analyses, dashboards, etc. to drive key business decisions

- Continually improve ongoing reporting and analysis processes, automating or simplifying self-service support for customers

- Own the functional and nonfunctional scaling of software systems in your ownership area.

- Provides input and recommendations on technical issues to BI Engineers, Business & Data Analysts and Data

Scientists.

- Support the development and validation required through the lifecycle of the DWH and Business Intelligence

systems, maintain user connectivity, and provide adequate security for data warehouse.

- Experienced in using Open source technologies (Apache Spark, Hadoop, Hbase, Hive, Pig, Kafka, etc.);

- Designed and building a data warehouse, streaming architecture, and machine learning pipelines;

- Ensure on time and on budget delivery which satisfies project requirements, while adhering to enterprise

architecture standards

- Implemented Big Data technologies including Hadoop, HDInsight, Data Bricks, ADF on Azure

- Have coded some advanced SQL

- Created some Power BI reports ,Dashboard and Apps

- Worked on automating SSIS Package generation using BIML and Python

- Worked in Document migration of around 1Tb of size

- Worked in a Scaled Agile Framework

- Understanding CI\CD concepts and hands on implementations experiences using GIT

- Hands-on Experience in Azure Databricks, Azure Data Lake, Azure Data Factory, Azure Event Hub, SQL Server, Cosmos DB

## **Technology Analyst**

Infosys

Apr 2019 - Jul 2019 (4 months)

- Analyze data, develop transformation scripts, remediate quality issues, and work collaboratively with SMEs to

implement and test new data products

- Coordinate with technology teams to design, construct, install, test and maintain data management systems

- Create reusable components and data tools for broader Research department use in accessing and analyzing

Research datasets

- Help create and evolve data standards and product documentation

- Supervise design throughout implementation process.

- Interface with other technology teams to extract, transform, and load (ETL) data from a wide variety of data

sources

- Reduced cycle time of routine data migration activities by 50%.

- Ensured end to end process of data extraction until reporting.

- Worked With NoSql Technologies

## **Senior Software Engineer**

## Infosys

Apr 2017 - Apr 2019 (2 years 1 month)

- Database management, data quality and integrity; maintained a 99% data accuracy standard
- Automated the manual data reconciliation process between the different sources, increasing the data accuracy and efficiency
- Monitor the DWH and BI systems performance and integrity provide corrective and preventative maintenance as required.
- Responsible for Analysis the data using data model using Microsoft SSAS and changes to Power BI online reports as per User requirements.
- Extensively worked on SCD type 2 using Look up transformation
- Worked in Upgrading Informatica 9.6 to Informatica 10.1
- Created SSIS packages to kill jobs running for a long time to protect database performance



## System Engineer

Infosys

Oct 2015 - Apr 2017 (1 year 7 months)



## Intern

Tata steels jamesdpur

May 2013 - Jun 2013 (2 months)

## Education



### Institute of technical education and research

Bachelor of Technology (BTech), Mechanical Engineering

2011 - 2015



### Dav public school

12, Physics chemistry mathematics

1999 - 2011



### Dav public school

10th, Science

2008 - 2009

## Licenses & Certifications



EY Analytics - Data integration - Bronze - EY



Microsoft Certified: Azure Fundamentals - Microsoft



## **Microsoft Certified: Azure Data Engineer Associate - Microsoft**

Issued Sep 2020 - Expires Sep 2022



## **Microsoft Certified: Data Analyst Associate - Microsoft**

Issued Oct 2020 - Expires Oct 2022



## **HackerRank Python certified - HackerRank**

## **Skills**

Apache Spark • Python (Programming Language) • SQL • Extract, Transform, Load (ETL) • Data Warehousing • Database Design • Apache NiFi • Big Data • Google Cloud Platform (GCP) • Microsoft Azure

## **Honors & Awards**



### **Bravo Points - Micron Technology**

Jul 2021

Have got Bravo points several time for good work



### **Extra Mile Award - EY**

Mar 2020

For Exceptional Client Delivery



### **EY Appreciate - EY**

Dec 2020

Saumya has done an excellent job leading 2 streams in TAL - Claims and Lifebook and has received countless appreciations from onsite and clients. His commitment and technical skills are commendable. The most complex mappings in the migration was created by him and he was the go to person from onsite BAs for any logic implementation and analysis. He constantly helped with the deployments to various environment and led the production deployments as well. He has put in lot of effort on personal dev