

SAUMYA RANJAN SAHU

Data Engineer

+91 7008229203

@ sahu.saumya26@gmail.com

<https://www.linkedin.com/in/saumyaranjansahu/>

Bangalore



SUMMARY

Driven data engineer with 5 years experience wrangling big datasets designing, deploying and operating highly available, scalable and fault tolerant systems. With a solid understanding of core data engineering concepts like Data Ingestion, Data Transformation, Data Modelling, Data Governance and Performance Optimisation

EXPERIENCE

Data Engineer

EY

08/2019 - Ongoing Bangalore

EY is a global leader in assurance, tax, transaction and advisory services. The insights and quality services delivered help build trust and confidence in the capital markets and in economies the world over

- Managed the integration of systems, data and employees from company acquisitions
- Designed and developed different migration solutions
- Lead a team of 3 in building a generic ETL framework, which is now used as a standard
- What did you achieve in this role?

Data Engineer

Infosys

10/2015 - 07/2019 Bangalore

Infosys is a global leader in next-generation digital services and consulting. It has clients in 45 countries to navigate their digital transformation

- Ensured end to end process of data extraction until reporting.
- Reduced cycle time of routine data migration activities by 50%.
- Database management, data quality and integrity; maintained a 99% data accuracy standard
- Got trained in big data domains for 3 months
- Automated the manual data reconciliation process between the different sources, increasing the data accuracy and efficiency
- What did you achieve in this role?

CERTIFICATIONS



Azure Data Engineer Associate

Why are you proud of this achievement?



Microsoft Certified: Data Analyst Associate

Why are you proud of this achievement?

TECH STACK

Informatica ,SSIS,SQL



Power BI, Kafka
HDFS, Hive, Pig

How did you acquire it? What did it result in?



Python

How did you acquire it? What did it result in?



Azure Cloud

How did you acquire it? What did it result in?



MS SQL Server,DB2,Azure
Blob,Cosmos DB,
DataLake

How did you acquire it? What did it result in?

LANGUAGES

English

Expert



Hindi

Expert



EDUCATION

B.Tech

Institute of Technical Education and Research

📅 2015 - 2019

GPA

7.8 / 10

PROJECTS

EY GDS, Insurance company in Australia

📅 12/2019 - Ongoing 📍 Bangalore

🔗 URL

Summary: The insurance firm acquired a product of another insurance firm, As part of the acquisition my responsibility was to migrate all the data.

- Responsibility:
 - Get the Business requirement form the client and create ETL pipeline according to the business requirement
 - Convert Complex business logic to simple ETL's
 - Performed the Unit Testing
 - Implemented exceptions and exclusion of data's in ETL
 - Build Toggling feature in ETL
 - Deployed and maintained code.
 - Ran the ELT's and monitored the performance and simultaneously improved it.
 - Also helped the Release Management Team.
 - Automated SSIS Package Generation using BIML and Python
 - Automted SQL query Generation using T-sql

Different POC works

📅 2019 - Ongoing 📍 Bangalore

Worked on different POC's involving Big Data technologies

- Create a 4 system cluster using Oracle Virtual Box to implement Kafka
- Created a producer and consumer in Kafka fetch twitter data using Kafka connect
- Also worked on the concepts of HDFS , Map-Reduce , Hive and Pig
- Current working on POC on Apache Spark using DataBricks

PROJECTS

2016–2019, Infosys Ltd, Insurance company in USA

📅 05/2016 - 06/2019 📍 Bhubaneswar
🔗 URL

Build a Data warehouse for Financial Representative and policies they sold and Revenue they generated at different Hierarchy level and Create User readable reports.

- Responsibility:
 - Analysis of requirement specifications provided by the Clients and translating them to technical impacts into systems. Creation of Design Documents, Test Plans and Test Scripts.
 - Responsible for designing ETL mappings through which data will flow from source to targets with some business rule applied with Agile process.
 - Analysis of the detailed data from source and minutely add them to target.
 - Unit Testing all kind of test cases which will give best ETL mapping for data flow.
 - Worked in process of migration of ETL components to Azure.
 - 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
 - Performed data validation in the target tables using complex SQLs to make sure all the modules are integrated properly. and also created mappings for automated data validation.
 - Converted Informatica mappings to SSIS package for azure cloud online databases.
 - Created SSIS packages to kill jobs running for a long time to protect database performance.
 - Worked in Column level encryption and decryption of Table.
 - Created RealTime ETL's
 - Designed Various SQL server database objects (Stored Procedures, Triggers, Views etc.
 - Worked in Creating cubes from views and Created highly usable reports from those Cubes.
 - Created Dimensional and Fact Table from Fact to View which are used in Cube