Shubham More

Phone number: (+91) 8888591881 (Mobile) | **Email address:** shubhammofficial12@gmail.com | **LinkedIn:** https://www.linkedin.com/in/shubham - more - 55a704285 | **Address:** India (Home)

ABOUT ME

- **Results-driven Experienced Data Engineer** with **4 years** of experience in pipeline designing, developing, and optimizing efficient ETL processes to streamline data integration and transformation.
- **Proficient in Azure Cloud, Python, SQL, PySpark** with a strong ability to leverage these technologies for scalable data processing and analysis.
- A **strong team player**, skilled at integrating with teams to deliver quality solutions.
- Experienced in designing, developing, and optimizing **ETL pipelines**, leading to **40% faster data processing**, **20% improved data accuracy**, and **50% reduction in system downtime**.

DIGITAL SKILLS

Azure Cloud

Azure Data Lake | Azure Data Factory | Azure Databricks | Azure Synapse | ETL Processing | Azure blob storage

Language

Python | SQL | PySpark | spark SQL

Other

GIT (Git Hub Git Bash) | CI/CD (Gitlab CI)

Big Data Technologies

Apache Spark, Apache Hive

Database

Snowflake (Cloud-based Data Storage and Analytics Service) | Microsoft SQL SERVER

WORK EXPERIENCE

05/2021 – CURRENT Bangalore, India

DATA ENGINEER ADUNHILL TECHNOLOGY

PROJECTS

2022 – CURRENT

Notification System

- **Client**: American multinational strategy and management consulting firm
- **Description**: Migrated Alteryx pipelines and developed automated ETL workflows using Azure and Snowflake identify and alert potential risks in operational workflows. Built efficient notification systems to deliver timely alerts to stakeholders, enhancing the ability to address risks proactively.
- **Technology Stack**: Azure Datalake, Azure Data Factory, Databricks, Snowflake, JIRA, GIT, Python, GitHub, Pyspark, Postman.
- **Responsibilities**:
  - Engineered **high-performance data pipelines** using **PySpark scripts** to extract, transform, and load **terabytes of data** from **Snowflake tables** within **Azure Data Factory**, reducing processing time by **35%**.
  - Developed logging functions to store job logs and manage metadata effectively.
  - Designed **dynamic email templates** with configurable settings, enhancing stakeholder communication efficiency by **50%**.
  - Utilized API calls to retrieve data and created views to optimize data retrieval across multiple tables.
  - **Implemented scalability and fault tolerance** along with various transformations in the application.

02/2022 – 11/2022

Data Management Platform

- **Description: Designed and deployed automated ETL pipelines** for the multiple data sources **(Oracle, ADLS Gen2, APIs)** to push the data from source to target **(OpenSearch Indexing Tool)** after performing transformations.

- **Technology Stack:** Azure Data Factory, ADLS Gen2, Azure Functions, Azure Logic Apps, Azure Event Grid, JIRA, GIT.

- **Responsibilities:**
  - Developed scalable PySpark pipeline for seamless data migration from on premise to azure cloud thereby **enhancing availability and significantly reducing operational costs.**
  - Implemented **incremental data loading**, reducing redundancy and improving efficiency by **45%**. and then push the data in batches to OpenSearch indexing tool.
  - **Led a cost-saving initiative** by identifying long running jobs and make them in optimized for data manipulation
  - Developed Azure functions to manage the metadata information of the jobs.
  - Implemented event-triggered alert systems for identifying job failures and errors, reducing downtime by 15% and improving customer satisfaction.

07/2021 – 12/2022

Data Management Platform

- **Description**: Project involved migration of legacy system on to the modern-day cloud-based architecture.

- **Responsibility**:
  - Designed and developed py-spark script to ingest the data from ADLS Gen 2 into the redshift table
  - Enhancements done into the existing raw to staging framework.
  - Created the hive DDL for external hive tables.

EDUCATION AND TRAINING

Aurangabad, India

BACHELOR OF ENGINEERING JAWAHARLAL NEHRU ENGG. COLLEGE

Field of study Electronics And Telecommunication