

Raktim Das

Data Engineer | +91-8638194454 | raktimd359@gmail.com | [Linkedin](#) | [Github](#) | Guwahati, Assam

TECHNICAL SKILLS

- **Programming Languages:** Python, SQL
- **Frameworks & Libraries:** Django, Flask, FastAPI, Pandas, Numpy
- **Databases:** MySQL, PostgreSQL, MongoDB, Cassandra
- **Data Warehouses:** Hive, AWS Redshift, BigQuery, Snowflake
- **Message Queue:** Confluent Kafka
- **Distributed Computation Frameworks:** Apache Hadoop, Apache Spark, DataBricks
- **Dashboarding Tools:** PowerBi, Apache Superset
- **ETL & Ingestion Tools:** DBT (ELT), Airbyte
- **Workflow Management:** Airflow
- **Cloud Platforms:** AWS, GCP
- **AWS Services:** S3, EC2, RDS, Lambda, Cloudwatch, DynamoDB, Redshift, Athena, Glue, SNS, SQS, Kinesis, Step Function
- **CICD:** Docker, Git, Github Actions, Terraform (Inc)

EXPERIENCE

Technical Support - e-Prastuti Project Under GAD, Govt Of Assam.

August -2024

- The role involves Drupal administration for the government of Assam Websites.
- Training Master Trainer, Content Managers of Each Concerned Department in Drupal CMS.
- Using Python and other tools for monitoring the websites, developing scripts for deploying in the server.

PROJECTS

Realtime Modern Streaming Data Pipeline

[Github](#)

Tech Stack: Python, AWS(EC2 , RDS , CloudWatch), Airbyte, Snowflake, DBT, Airflow, Github Actions, Docker, Terraform

- Designed and implemented a data pipeline to generate and ingest synthetic data using Python on AWS EC2 and RDS PostgreSQL, integrated with Airbyte for daily ingestion into Snowflake.
- Automated data transformation workflows with dbt and Apache Airflow, scheduling refreshes for Snowflake tables.
- Enabled real-time analytics and reporting by integrating Looker for visualization, ensuring seamless monitoring with CI/CD pipelines using GitHub Actions, Docker, and Terraform.

Incremental Ingestion Pipeline - Airline Data

[Github](#)

Tech Stack: PySpark, S3, Glue crawler, Glue ETL, EventBridge, Redshift, SNS, Step Functions, SQL

- Build generic & optimized ETL pipeline to process daily incoming data of airline from S3 for daily incremental load in Redshift Table.
- This helped in predicting flight delays, optimizing schedules, enhancing customer satisfaction and profitability

EDUCATION

Bachelor of Computer Applications
Gauhati University, Pragjyotish College
Aug 2020 - Aug 2023
CGPA - 8.5