

Yashashvi Dethe

Mumbai, India

+91-7057806052 | yashashvuidethe3108@gmail.com | www.linkedin.com/in/yashashvi_dethe

SUMMARY

Data Engineer with 2.7 years of experience in building scalable data pipelines, cloud migrations, and optimized ETL workflows. Proficient in BigQuery, Dataflow, Airflow, and Cloud Run with strong expertise in Python, SQL, and GCP-based deployments. Experienced in data modeling, automation, and implementing CI/CD practices for production-ready solutions.

EDUCATION

Batchelors of Engineering | SPPU Session: 2017 – 2021 |

Score: 8.67 SGPA

SKILLS

Technical: SQL, Python, PySpark, ETL, Data Modelling

Cloud: Google Cloud Platform – BigQuery, Cloud Storage, Cloud Composer, Cloud Function, Airflow, Dataflow, Pub/Sub

Databases: SQL Server, PostgreSQL, Teradata, BigQuery, SAP HANA

Tools & Technologies: Jira, Git, GitHub, CI/CD, Tekton, Data Lake, Tableau, Qlik, Power BI, Data Visualization

CERTIFICATIONS

- Associate Cloud Engineer
- Professional Data Engineer
- AWS Cloud Practitioner
- Professional Cloud Database Engineer

EXPERIENCE

Quantiphi Analytics

Mumbai, India

Data Engineer

Nov 2022 – Present

- Designed and implemented scalable data pipelines for batch and real-time processing using Dataflow, Apache Beam, and Cloud Run.
- Migrated complex data workflows from legacy systems like ADLA and PostgreSQL to Google BigQuery, optimizing performance and cost.
- Converted PySpark-based Databricks notebooks into efficient SQL stored procedures in BigQuery, improving maintainability.
- Automated end-to-end ETL workflows using Apache Airflow, enabling reliable scheduling and monitoring of pipelines.
- Engineered API integration pipelines with Cloud Run and Python to fetch, process, and load external data into BigQuery.
- Skilled in data modeling, implementing SCD Type 1 & 2 logic, and optimizing SQL queries for large-scale analytical workloads. Experienced in performance optimization, data modelling and building scalable, reliable data pipelines to support analytics and business intelligence initiatives.

PROJECTS

ADLS to BigQuery Migration (6 Months)

- Migrated enterprise pipelines from Azure Data Lake Analytics (ADLA) to BigQuery, improving performance and scalability.
- Implemented SCD Type 1 and Type 2 logic and used Airflow for orchestration of Dataflow jobs and Cloud Run APIs.

PostgreSQL to BigQuery Pipeline (5 Months)

- Built Dataflow pipelines to extract data from PostgreSQL, process .gz files, and load into BigQuery.
- Handled schema mapping, data quality checks, and batch-load optimizations.

Databricks Notebook to SQL Conversion (4 Months)

- Converted Databricks PySpark notebooks into BigQuery SQL stored procedures.
- Scheduled and monitored workflows using Apache Airflow, enhancing performance and maintainability.

API to BigQuery Ingestion with Cloud Run (5 Months)

- Developed a Python Cloud Run service to ingest data from external APIs into BigQuery.
- Implemented a YAML-driven config system with robust logging, retries, and error handling.