

Akash Singh

☎ +91-7055603344 🔗 [linkedin.com/in/akash-singh-1b9586178](https://www.linkedin.com/in/akash-singh-1b9586178) ✉ singh.akash1dec@hotmail.com 🌐 portfolio.s7labs.dev

SUMMARY

Data Engineer with **5+** years building cloud-native data platforms on **GCP** and **Azure**, specializing in **ETL/ELT pipelines**, **data warehousing (BigQuery, Snowflake, Databricks)**, and **GenAI/Agentic AI** solutions.

EXPERIENCE

Best Buy India | *Software Engineer II (Data)*

April 2025 – Present

- Designed and built a **subscription lifecycle framework** that consolidated historical and current data into an **SCD Type-2** model, enabling 20+ business consumers to query the point-in-time state of any subscription.
- Introduced **conversational analytics** that let business users get answers from large datasets by simply asking questions in plain English, empowering 50+ non-technical stakeholders to self-serve and reducing ad-hoc requests to the data team by 45%.
- Designed and implemented a **Snowflake Data Clean Room (DCR)** to establish a secure environment for joint analytics with vendor partners without exposing underlying raw data, enabling privacy-preserving collaboration.
- Optimized a **GCP/BigQuery** digital-subscriptions data pipeline spanning 20+ source tables, cutting query runtime ~40% and eliminating memory-limit failures while preserving exact output across 120+ fields.
- Modernized legacy on-premise datasets by integrating multiple data sources into **Google BigQuery**, enabling the retirement of **SAP BusinessObjects** and **QlikView** reporting platforms and reducing annual licensing costs by approximately \$740K.

Shell India Markets | *Data Engineer*

Oct 2021 – April 2025

- Engineered, optimized, and managed **ETL pipelines** using **Azure Data Factory** and **Databricks**, integrating **Delta Live Tables** and **Spark Streaming** for real-time data processing, resulting in a 25% reduction in data latency and a significant increase in pipeline reliability.
- Built an end-to-end pipeline that leveraged **web scraping** and **API requests** to consolidate public and vendor-specific datasets into a unified source, powering forecasts of power demand and generation across Singapore, the Philippines, and Japan, boosting analytics efficiency by 40%.
- Created a secure data-sharing framework using **Azure LogicApps** that enabled 10+ third-party vendors to access curated data products for analytics with all access permissions and governance controls satisfied, accelerating secure external collaboration.
- Automated pipeline testing using the **Databricks AI Assistant**, reducing development time by 10% and accelerating delivery of production-grade pipelines.
- Built and optimized **CI/CD pipelines** to automate deployments across dev and prod environments, cutting deployment time by 30% and eliminating repetitive manual tasks.

SKILLS

Languages: Python, PySpark, SQL

Cloud, Warehousing & Data Engineering: GCP (BigQuery, Vertex AI, BigQuery ML), Azure (Data Factory, Data Lake, LogicApps, Functions), Snowflake, Teradata, **ETL/ELT Pipelines**, **Airflow**, **Databricks**, **Data Modeling**

AI/GenAI: Agentic AI, LangChain, ReAct Agents, LLM Orchestration, **RAG**

DevOps & Tools: Git/GitHub, GitHub Actions, **CI/CD**, Databricks Asset Bundles, Looker Studio

EDUCATION

VIT University, Vellore

June 2021

B.E. Computer Science

ACHIEVEMENTS

- Awarded the **Q3 2022 VP CIO Award** for the AskBrand Auto Approval project, projected to save Shell \$1.3M annually.
- Databricks Certified Data Engineer Associate**
- GCP Professional Data Engineer**