AJAY A

palakkad,kerala ✉ ajaya82717@gmail.com, anayajay58@gmail.com ☐ +917907379758 🔗 in/ajay7777

PROFESSIONAL SUMMARY

DATA ENGINEER
INFOSYS   December 2021 - Present

Results-oriented and performance-driven Data Engineer with over 4 years of experience in designing, building, and optimizing end-to-end data solutions across cloud-based and on-premise environments. Specialized in data development, migration, analysis, and pipeline engineering, with hands-on expertise in Google Cloud Platform (GCP), SQL, BigQuery, DBT, Python, MemSQL, and Apache Airflow. Proficient in data analysis, code optimization, unit testing, and the design and optimization of data workflows. Skilled in building robust ETL processes, enhancing data quality and reliability, and ensuring scalable and maintainable data infrastructure. Known for solving complex technical challenges, collaborating cross-functionally, and driving operational excellence through innovative data engineering practices.

SKILLS

Technical Skills

- **Programming & Scripting**: SQL, Python, Unix
- **Cloud Platforms & Services**: Google Cloud Platform (GCP), BigQuery, Pub/Sub, Cloud Storage (GCS), Cloud Functions, Cloud Composer
- **Databases & Data Warehousing**: MySQL, MemSQL, Teradata, SQL Server
- **ETL & Data Integration**: DBT, Talend ETL, Tibco Jaspersoft ETL, Tibco Streaming, Real-time & Batch Data Processing Pipeline Design
- **Workflow Orchestration & Automation**: Apache Airflow, Docker, Kubernetes, Cron
- **Data Storage & File Systems**: Google Cloud Storage (GCS), Network Attached Storage (NAS), MinIO, SingleStore Database
- **Data Governance & Visualization**: Qlik Sense, Grafana, Tibco EBX, Tibco Data Virtualization (DV)
- **Version Control & CI/CD**: GitHub, GitLab, GitHub Actions, Nexus Repository Manager
- **Development & Collaboration Tools**: Jira, Zoho, Excel, WinSCP, VS Code, Beyond Compare, DBeaver, Infosys IDTW
- **Operating Systems**: Windows, Linux (System Maintenance & Administration)

EDUCATION

Bachelor of Computer Science
University of Calicut • 2021

Master of Computer Applications (MCA)
Indira Gandhi National Open University (IGNOU) • 2025(expected)

PROJECT EXPERIENCE

1. **Data migration project on Google Cloud Platform (GCP)**

Project goal is to migrate the client data from Teradata database to GCP-BigQuery along with the implementation of components for access management, scheduling/triggering tasks, data backup/archival and API integration for user specific requirements

- **GCP Services**: Worked with BigQuery, Cloud Functions, Cloud Storage, Cloud Composer, Dataflow, Cloud Pub/Sub.
- **Teradata to DBT Transformation**: Converted Teradata BTEQ scripts to DBT models, integrated with Control-M workflows.
- **Testing**: Validated code in DEV, QA, and production environments.
- **Proficient in handling large, complex datasets**: Optimized data flow and ensured data integrity across various systems.

- **Airflow & DBT**: Created Airflow DAGs for scheduling and triggering DBT models based on client needs.
- **Optimized SQL Queries**: Improved data retrieval and processing with optimized SQL queries.
- **Agile Methodology**: Delivered incremental improvements, adapting to changing requirements and tight timelines.
- **Data Analysis & Optimization**: Performed data analysis, code optimization, unit testing, and debugging.
- **Documentation**: Created detailed documentation to enhance client understanding and engagement.

2. Big Data Management and Integration Project

The goal of this project is to design and implement a robust big data management and integration platform that efficiently handles both real-time and batch data flows using SingleStore DB/ MemSQL. The system ensures seamless data processing, high performance, and scalability to support business decision-making and analytics.

- **SingleStore Database**: Utilized SingleStore (formerly MemSQL) with HDFS architecture to efficiently manage real-time and batch data, optimizing query performance.
- **Hybrid Data Pipelines**: Built and managed hybrid data processing pipelines using Tibco Streaming and Jaspersoft ETL for real-time and batch data flows.
- **Real-Time Data Streaming**: Implemented continuous 24/7 real-time data flow with Tibco Streaming, monitoring and maintaining jobs from the TAC server. Developed batch data loading jobs for unavailability periods.
- **Data Governance**: Leveraged EBX MDM platform for effective data governance, ensuring consistent and compliant master data.
- **Data Visualization & Monitoring**: Created interactive data visualizations in Qlik and used Grafana for system health and performance monitoring.
- **Data Backup Automation**: Automated data backup from SingleStore to NAS using Python scripts for secure, scheduled transfers, preventing data loss.
- **Cross-Functional Collaboration**: Worked with cross-functional teams to integrate tools and optimize data workflows, enhancing system reliability and failover mechanisms.
- **Client Relationships**: Maintained strong client relationships, ensuring effective communication and delivering tailored data solutions.

CERTIFICATIONS

- Infosys Certified Cloud Developer Professional
- Infosys Certified Full Stack GCP Associate Engineer
- Infosys Certified MySQL Associate
- Infosys Certified Python Associate
- Infosys Global Agile Developer Certification
- Infosys Certified L1 Business Intelligence Programmer

AWARDS AND HONORS

- **Data and Analytics Unit Rise Award (FY25 Q1) – Infosys**

Recognized by Infosys for exceptional performance and outstanding contributions to the company's culture of excellence and innovation.

- **Data and Analytics Unit INSTA Award – Infosys**

Honored by Infosys for exceptional performance and timely delivery of project deliverables, demonstrating a commitment to high-quality execution and project success.

- **Commendable performance Rating (Consecutive Quarters) – Infosys**

Acknowledged by Infosys for consistently exceeding performance expectations across multiple quarters, ensuring excellence in project delivery and maintaining high client satisfaction.

- **Busy Bee Award – Infosys**

Awarded by Infosys for relentless commitment and consistent hard work in managing multiple tasks efficiently, ensuring smooth project execution and timely completion of key deliverables.