

# Sanjeevi M

No.14, Kailash Sector, 3rd Street,  
Chinnamettupalayam, Kaladipet,  
Chennai, Tamil Nadu, India – 600019.

**Mob:** +91 8610889030 | **sanjayvm66@gmail.com**

**LinkedIn:** [www.linkedin.com/in/sanjeevi-m](https://www.linkedin.com/in/sanjeevi-m)

**GitHub:** <https://github.com/askmrsanjay>

**Tableau Public:** <https://public.tableau.com/app/profile/sanjeevi.m/vizzes>

---

## Objective

Azure & Databricks-certified Data Engineer with 3.5+ years of experience designing, automating, and optimizing scalable ETL/ELT pipelines using PySpark, Snowflake, Delta Lake, and Azure cloud services. Skilled in building metadata-driven frameworks, improving performance, and delivering enterprise-grade data solutions for Fortune 500 clients.

## Experience

### COGNIZANT TECHNOLOGY SOLUTIONS | ASSOCIATE DATA ENGINEER | JUL 2025 – PRESENT

- Building and optimizing **cloud-native ETL pipelines** using Azure Databricks, PySpark, and Delta Lake for enterprise-scale workloads.
- Designing **metadata-driven workflows** to streamline ingestion, improve reusability, and reduce development effort.
- Implementing **Delta Lake ACID pipelines**, supporting incremental loads, schema evolution, and time-travel.
- Integrating Azure Key Vault and Databricks Secret Scopes to ensure **secure credential management**.
- Collaborating with product owners, QA, and platform teams to deliver high-quality production data pipelines.
- Monitoring and tuning Spark jobs to **improve performance**, reduce compute costs, and ensure reliability.

### INFOSYS LTD | TECHNOLOGY ANALYST | MAY 2022 – JUL 2025

- Designed and implemented **end-to-end ETL pipelines in Azure Databricks** processing large-scale data, improving performance through PySpark optimizations and Delta Lake tuning.
- Developed ingestion workflows for **structured (Teradata)** and **unstructured (JSON, Parquet)** data with automated data quality checks.
- Engineered **CDC (Change Data Capture)** solutions for customer billing and sales order pipelines with <1% latency.
- Built automated reporting pipelines generating daily reports and Slack alerts, eliminating manual Tableau refreshes.
- Created optimized SQL queries and PySpark transformations, achieving significantly faster execution.
- Built and maintained **interactive Tableau dashboards** for marketing, operations, and financial insights.
- Worked closely with Apple's Data Operations team to design dashboards for anomaly detection, pipeline health, and KPI monitoring.

## Education

POST GRADUATE STATISTICS | MAY 2019 | PRESIDENCY COLLEGE, CHENNAI

## Skills

- **Languages:** Python, SQL
- **Big Data & Frameworks:** PySpark, Apache Spark, Delta Lake, Delta Live Tables, Snowpark
- **ETL & Orchestration:** Azure Data Factory, Informatica, Databricks Workflows
- **Databases:** Snowflake, Teradata, Oracle, Azure SQL
- **Cloud & Storage:** Azure (ADLS, Key Vault, AKV), Unity Catalog
- **Tools & DevOps:** GitHub, Azure DevOps, Slack Integration
- **Visualization:** Tableau Desktop, Tableau Server

## Certifications

- **Databricks Certified Data Engineer Associate**
- **Microsoft Azure Data Engineer Associate**
- **Infosys Certified PySpark Developer**
- **Infosys Certified Python Programmer**
- **Infosys Certified Tableau Developer**

## Projects

### Snowpark-Based Dynamic Rule Processing Framework

Data Engineer | Humana Inc.

- Designed and developed a **config-driven Snowpark framework** that dynamically reads transformation and validation rules from JSON, enabling flexible, code-light processing.
- Implemented **dynamic rule parsing** for dataset mappings, join logic, transformation sequences, and parameters.
- Engineered Snowpark pipelines to **load, combine, and transform** data using joins, unions, and rule-based logic.
- Built a **sequential rule execution engine** ensuring deterministic and fully auditable workflows.
- Developed an **audit logging module** capturing rule name, datasets, parameters, timestamps, initiator, and execution status.
- Implemented robust **error logging for exceptions, validation failures**, and diagnostic information.
- Enabled full **extensibility** by making rule updates driven through JSON without code changes or redeployment.
- Ensured operational transparency and governance compliance through end-to-end traceability.

## Databricks – Snowflake Dynamic Extract Framework

### Data Engineer | Humana Inc.

- Built a **parameter-driven, config-based PySpark framework** to dynamically process any system + file ID combination without code changes.
  - Automated metadata retrieval from **Snowflake control tables** (ADLS paths, target tables, transformation SQL).
  - Developed ingestion pipelines to extract Snowflake data into **Delta format on ADLS**.
  - Automated **Unity Catalog External Table creation** for seamless SQL access.
  - Implemented ACID-compliant **INSERT...SELECT** pipelines loading into managed Delta tables.
  - Developed a centralized **Audit & Error Logging module**.
  - Secured pipelines using **Azure Key Vault / Secret Scope**.
  - Deployed using **Databricks Jobs & Workflows** ensuring reliability.
- 

## Heritage EMEA Billing & Sales Order Pipelines

### Data Engineer | Nike Inc.

- Gathered, cleansed, and processed customer billing and sales order data from **Teradata** for EMEA analytics.
  - Built PySpark pipelines for **ingestion and transformation**, performing multi-table data blending.
  - Designed **Informatica workflows** for EMEA pipeline objects.
  - Created and tested **CDC SQL queries enabling** accurate incremental data updates.
  - Increased ETL processing efficiency by 50% for datasets exceeding 8M records.
  - Implemented multilayer validation with automated email-based verification reports.
  - **Achieved <1% latency and 99.9% EMEA pipeline uptime.**
- 

## AMP Data Science & Platform Operations – Apple Inc.

### Data Analyst / Data Engineer

- Developed PySpark pipelines to ingest and blend data from Teradata for analytics and reporting.
  - Designed anomaly detection dashboards aligned with Apple's Data Ops portal using Tableau.
  - Automated report generation and Slack notifications using Databricks, minimizing manual intervention.
  - Analysed KPIs, pipeline status, anomaly metrics, and record-level trends.
- 

## Personal Details

- **Date of Birth:** 29/07/1996
- **Nationality:** Indian
- **Gender:** Male
- **Location:** Chennai, Tamil Nadu, India