


# Rekhaa Mallam

+1-757-412-8185 ✉ [rekha.mallam@gmail.com](mailto:rekha.mallam@gmail.com)  <https://linkedin.com/in/rekha-mallam-024427204>

## PROFESSIONAL SUMMARY

Data Engineer with 14 years of experience delivering robust data integration and scalable ETL solutions. Proficient in architecting dimensional data warehouses and real-time pipelines using Python and SQL to optimize performance both on-prem and in the cloud. Demonstrated ability to mentor teams and collaborate effectively with cross-functional partners to drive innovative data solutions.

## TECHNICAL SKILLS

- **ETL Tools:** IBM DataStage (Parallel & Server jobs), Informatica PowerCenter
- **ETL Development:** Joins, Lookups, Filters, Aggregators, Transformers, Funnel, Merge, Sequence Generators, SCD (Type 1 & 2), Data Cleansing, Parameterization
- **Cloud & Big Data:** Azure Data Factory (ADF), Azure Data Lake, Cosmos DB, Azure SQL DW
- **Languages & Scripting:** SQL, Python, PL/SQL, Unix Shell Scripting
- **Data Engineering:** Apache Iceberg, Spark, Snowflake
- **Databases:** Oracle, DB2, SQL Server, Teradata, PostgreSQL
- **Data Warehousing:** OLTP, OLAP, Dimensional Modeling, Star & Snowflake Schema
- **Testing Tools:** QuerySurge, SQL Developer, TOAD
- **Version Control & Agile:** GitHub, Bitbucket, Jira, Confluence

## PROFESSIONAL EXPERIENCE

### IBM

**Jun 2021 - Oct 2024**

*Data Engineer*

*Client: State of Colorado, USA*

- Developed and streamlined ETL workflows utilizing IBM DataStage and Azure Data Factory to efficiently process data.
- Engineered complex transformation logic employing joins, funnels, pivot/unpivot operations, and conditional aggregations.
- Automated ETL test validation with QuerySurge, cutting test cycle time by 60% and identifying unmatched row counts.
- Leveraged DataStage Sequencers and parameters to orchestrate reusable and maintainable jobs.
- Optimized pipeline performance through parallel processing and systematic tuning techniques.
- Adjusted APT\_BUFFER\_SIZE and APT\_NO\_TRANSFORM\_COMBINE settings to enhance memory management and transformation efficiency.

### IBM Watson

**Jan 2019 - Feb 2021**

*Senior Data Warehouse Engineer*

*USA*

- Designed multi-stage ETL pipelines to migrate legacy systems to Azure Data Warehouse, employing PolyBase for high-performance bulk loads.
- Implemented robust transformation logic with multi-source joins, router conditions, and surrogate key generation in ADF using Conditional Split transformations.
- Optimized data processing by tuning SQL queries and indexes, reducing ETL run time by 35% through early-stage pushdown filters.
- Engineered Azure Data Factory pipelines with scheduled triggers and integrated monitoring via Azure Monitor to proactively track job runs and failures.

### Freddie Mac

**Jan 2017 - Sep 2018**

*ETL Developer*

*USA*

- Developed DataStage jobs integrating flat files, Oracle, and DB2 sources.
- Created reusable mapplets and parameter files in Informatica to ensure consistency across ETL processes.
- Validated ETL data loads and executed regression testing using QuerySurge and SQL queries.

### Avery Dennison Corporation

**Jan 2010 - Sep 2012**

*ETL Consultant*

- Led enterprise-level ETL development using Informatica PowerCenter and IBM DataStage to support data warehouse operations.
- Constructed complex mappings featuring update strategies, aggregators, and lookup overrides.
- Automated ETL testing frameworks with shell scripts, streamlining quality assurance processes.
- Utilized Erwin Data Modeler to design and maintain conceptual, logical, and physical data models for the data warehouse.
- Collaborated on dimensional modeling initiatives to align data structures with BI reporting requirements.

### Tata Consultancy Services

**May 2008 - Dec 2009**

*ETL Developer*

*General Motors*

- Designed and built data workflows in Informatica for processing XML, flat file, and Oracle sources.
- Constructed STAR schema designs, developed corresponding dimension and fact tables, and optimized mapping performance.
- Conducted performance tuning using session parameters, partitioning strategies, and pushdown optimization techniques.

### Tata Consultancy Services

**Jan 2007 - Apr 2008**

*ETL Developer & QA Analyst*

*Bank of America*

- Developed DataStage jobs incorporating joins, filters, and transformers tailored for banking applications.

- Executed backend testing with advanced SQL queries and PL/SQL procedures to ensure data integrity.
- Engaged in quality assurance processes including test planning, execution, and defect tracking.

## EDUCATION

---

**Jawaharlal Nehru Technological University (JNTU), India**

*Master of Science, Computer Science*

## CERTIFICATIONS

---

- **Certified in QuerySurge**
- **Government Insights and Solutions (Bronze): 2023**