

HARISHKUMAR R

Data Engineer

+91 79044 76829 • hk.de13@gmail.com • Chennai

Summary

4 years of experience as a Data Engineer specializing in Snowflake, AWS, SQL, dbt, Python, and cloud-based ETL solutions. Experienced in designing and developing scalable data pipelines using Azure Data Factory (ADF) and Apache Airflow for batch and automated data processing. Strong hands-on experience in Snowflake data warehousing, data ingestion, transformation, data modeling, and performance optimization. Proficient in Python for data processing, automation, validation, and pipeline development. Experienced in building dbt-based transformation frameworks, implementing incremental models, testing, documentation, and reusable SQL transformations. Skilled in integrating data from multiple sources into Snowflake using AWS services and ETL workflows, with a strong focus on data quality, reliability, and automation.

Skills

Snowflake • DBT • ETL • AWS S3 • Snow SQL • Python • ADF • Airflow • SQL • SQL Server 2019 • IAM • PL/SQL • Oracle RDBMS (19c,21c) • JIRA • UNIX • SQL Server • Windows

Education

Madurai Kamaraj University	2021
Post Graduate in Journalism and Mass Communication	
Thiagarajar College	2019
Bachelor's of Arts in English Literature	

Experience

Creativeye Technologies Pvt Ltd	06/2025 - Present
Project: RCI Bank	
- Designed and developed scalable data pipelines using Snowflake, AWS, dbt, Python, ADF, and Apache Airflow. - Developed ETL/ELT pipelines to ingest data from databases, APIs, CSV, JSON, and other source systems into Snowflake. - Built AWS S3-based data ingestion workflows and integrated S3 with Snowflake using stages, file formats, and COPY commands. - Created Snowflake databases, schemas, tables, views, stages, streams, tasks, and stored procedures. - Developed complex SQL transformations using joins, CTEs, window functions, MERGE, aggregations, and analytical functions. - Developed dbt models for data transformation, cleansing, data modeling, testing, documentation, and incremental processing. - Implemented incremental and CDC-based loading processes to process new and changed records efficiently. - Designed and developed scalable data ingestion and ETL pipelines using Azure Data Factory (ADF) and Apache Airflow. - Implemented dbt snapshots to maintain historical records and support SCD Type 2 requirements. - Developed Python scripts for data extraction, transformation, validation, automation, file processing, and error handling.	
LIDAVI Technologies	10/2022 - 03/2025
Project: Lion Energy	
- Designed and developed scalable cloud data warehouse solutions using Snowflake and AWS for enterprise data processing and analytics. - Developed complex Snowflake SQL queries using joins, CTEs, subqueries, window functions, aggregations, MERGE, and CASE expressions. - Built AWS S3 to Snowflake data ingestion pipelines for structured and semi-structured data. - Implemented full load and incremental load strategies for processing large-volume datasets efficiently. - Created and managed Snowflake databases, schemas, tables, views, stages, file formats, streams, tasks, and stored procedures. - Worked extensively with JSON, CSV, and Parquet files and implemented parsing, transformation, validation, and data-quality checks. - Implemented Streams and Tasks for incremental/CDC-based data processing and automated data pipelines. - Developed Stored Procedures in Snowflake Scripting for reusable ETL and business transformation logic.	
Project: SCORENXT	
- Designed and developed database solutions using Oracle PL/SQL for enterprise data processing and reporting requirements. - Developed complex SQL queries using joins, CTEs, subqueries, window functions, aggregations, and correlated subqueries. - Developed and maintained stored procedures, functions, packages, triggers, and cursors to implement business and data-processing logic. - Implemented ETL data loading processes for extracting, transforming, validating, and loading large-volume datasets. - Developed full load and incremental load processes based on business requirements. - Implemented data validation and reconciliation between source and target tables to ensure data accuracy. - Used explicit and implicit cursors for processing and handling multiple records within PL/SQL programs. - Implemented exception handling using predefined and user-defined exceptions to handle runtime and business-related errors.	