GEETANJALI DHOKE

business analyst

Business Requirements Document (BRD)

Amazon RDS to Snowflake Data Migration with DBT

Project Title: Amazon RDS to Snowflake Data Migration with DBT
Date:
Prepared By: Geetanjali Dhoke

Executive Summary:

This document outlines the business requirements for migrating structured data from Amazon RDS (PostgreSQL/MySQL) to Snowflake, leveraging DBT (Data Build Tool) for transformation and modeling. The goal is to enhance data scalability, performance, and analytics capabilities while modernizing the data platform.

Business Objectives

- ✓ Improve performance and scalability of analytics queries.
- ✓ Centralize reporting and analytics on Snowflake.
- ✓ Enable transformation and modeling using DBT.
- ✓ Establish a reliable and automated ETL/ELT pipeline.
- ✓ Minimize downtime and ensure data integrity during migration.

Scope of Work:

- ✓ Assessment of existing RDS schema and data volume
- ✓ Data extraction from Amazon RDS
- ✓ Data loading into Snowflake
- ✓ DBT setup for data transformation and modeling
- ✓ Data validation between source (RDS) and target (Snowflake)
- ✓ Scheduling via orchestration tools (e.g., Airflow, dbt Cloud, etc.)

Out of Scope:

- ✓ Real-time data sync (focus is batch or scheduled)
- ✓ Migration of unstructured data
- ✓ Building dashboards or BI tools (unless specified)

Functional Requirements:

Requirement	Description
Schema Mapping	Map RDS schema to Snowflake equivalent (data types, constraints).
Initial Data Load	Full load of existing data into Snowflake.
Incremental Loads	Set up DBT models for incremental refresh.
DBT Transformations	Build staging, intermediate, and mart layers in DBT.
Data Validation	Row counts, checksums, and spot checks post-load.
Logging & Monitoring	Set up logs and alerting for job failures.

Assumptions:

- ✓ Access to source (RDS) and target (Snowflake) databases is provided.
- ✓ Snowflake warehouse and storage are pre-configured.
- ✓ DBT profiles and credentials are set up.
- ✓ Business logic for transformation is documented or available.
- ✓ RDS data changes during migration window are minimal or managed via change capture.

Timeline:

Phase	Duration
Requirement Analysis	3–5 days
Extraction and Schema Mapping	4 days
Initial Data Load	2–3 days
DBT Transformation Development	7–10 days
Validation & Testing	3–4 days
Go Live & Monitoring	2 days

Risks & Mitigation:

Risk	Impact	Mitigation
Data Volume Causing Long Load Times	Medium	Use partitioned/batch loading, parallelization
Incomplete Schema Mapping	High	Conduct thorough schema review before migration
DBT Model Performance Issues	Medium	Optimize models using Snowflake best practices
Data Loss or Corruption	High	Take backups and validate with check sums and counts

Approval:

Name	Role	Signature	Date
[Client Name]	Client		
Geetanjali Dhoke	Business Analyst		