

Dilshod U.

Data Engineer

TechHive
IT's simple with us

Professional summary

Data Engineer with more than 7 years of experience in Database Design and Business Intelligence Development. Dilshod has strong expertise in building high performance data integration and workflow solutions, including extraction, transformation and loading (ETL) operations for data warehousing. His technical proficiency includes working with various database management systems, data modeling tools, and BI platforms, enabling him to create robust data architectures that support complex analytical needs. He is an articulate communicator who gets along well with clients and coworkers at all levels.

Education	National Research Nuclear University (MEPhI) Bachelor's Degree
Licenses& Certifications	Tools of Data Science IBM, Coursera; What is Data Science IBM, Coursera; Deep Learning Specialization, Coursera, Stanford University; Structuring Machine Learning Projects, Coursera, Stanford University
Skills	Python, C# AWS (Glue, Lambda, Athena, Redshift, S3) GCP, Azure Apache Airflow Pandas, PySpark SQL Power BI Tableau Snowflake DBT

Company	**UK fashion Retailer**
Project	NDA
Period	Nov 2022 - now
Role	Data Engineer/BI Developer
Technologies	Airflow, Data vault, Python, AWS (Redshift, S3, Glue, Athena),

Responsibilities

- Created DWH from scratch in Snowflake using data from transactional tables;
- Designed more than 50 business metrics at the request of the retention team;
- Developed ETL pipeline for receiving data from DynamoDB and put it on Snowflake;
- Created historical entities using the SCD2 method;
- Developed a three-level data storage on Snowflake;
- Configured usage and development settings on Snowflake using Terraform;
- Built analytical tables for usage and created reports on them;
- Configured an alerting system that checks data;
- Created ML to optimize business process;
- Migrated data from BigQuery to S3 bucket and then to Snowflake;
- Participated in the development of Data Vault and Data Mart using DBT and Snowflake;
- Created a DAG on Airflow to orchestrate processes.

Company	IT SVIT
Projects	Application is used to automate business operations of client companies and closely integrated with ERP systems such as Net Suit, JD Edwards; Solution for organization of tourist holidays for families who have an annual subscription, booking villas, houses, hotels.
Period	Oct 2020 – Nov 2022
Role	Data Engineer/BI Developer
Technologies	Airflow, Python (Pandas, Numpy and etc.), AWS (Redshift, S3, Glue, Athena), Airflow, SQL, PySpark, Power BI, SSRS, SSIS, Microsoft Azure (ADLS Gen 2, Azure SQL Database, Synapse, ADF), GitHub

Responsibilities

- Worked with different services such as S3 for main data lake in AWS;
- Used Redshift for the primary data warehouse;
- Used AWS Glue for ETL solutions and Athena for analysis and querying the data;
- Created and maintained different Lambda functions using PySpark in AWS;
- Designed Power BI reports on account information, research reports using ad-hoc reporting to be used by clients and advisors. Designed, developed, maintained and delivered additional new Power BI Dashboards and Reports with complex DAX query measures/columns, drill down/drill through/hyperlink options, dynamic slicing/dicing features;

- Used Spark-SQL to work with JSON, XML, Parquet as well as ORC data formats;
- Built Spark applications using PySpark and Spark SQL for extracting, transforming and aggregating data from several file formats for analysis;
- Built efficient ETL packages, using both SSIS and Azure Data Factory ADF2 pipelines for processing Fact and Dimension tables with complex transforms and SCD type 1/ type 2 changes;
- Did SQL performance monitoring and tuning of reporting data by optimizing indexes and stored procedures;
- Extracted JSON data from Rest API using Pandas in Python and parsed them into csv and text format;
- Created complex stored procedures, triggers, functions (UDF), indexes, tables, views to support the project;
- Did various performance tuning techniques by using parameterized queries, Indexing, SQL server profiler, query optimization, IO stats, database engine tuning advisor, execution plans etc;
- Scheduled ETL Packages to routinely deliver data into staging table and prepare the data before loading into Data warehouse;
- Created reports using Power BI Desktop and deployed to Power BI Service to share with end users via Workspaces and Apps which implemented Row Level Security and User Group management.

Company	Global Solutions
Domain	Finance
Period	Aug 2017 – Sep 2020
Role	Data Engineer
Technologies	Python (Pandas, Numpy), PySpark, DBT, Snowflake, MS SQL (T-SQL), SSIS, SSRS, SSAS, Power BI, some C# (in SSIS), Azure

Responsibilities

- Worked independently and as part of a team to accomplish critical business objectives and to make decisions under pressure;
- Analysed complex business requirements, created and modified complex stored procedures, also worked on performance tuning;
- Largely used T-SQL in supporting Finance Team and constructing User Profiles, Relational Database Models, Data Dictionaries and Data Integrity;
- Created complex T-SQL code such as Stored Procedures, functions, triggers, Indexes and views for the application;
- Created the optimal logics to load the data that business rules were based on flat file names or ragged hierarchy;
- Developed slowly changing dimension SCD-Type 0, 1, 2 in SSIS to update the old data in dimension tables when the new source data is loaded. Also automated the process by deploying the packages into SQL server and creating jobs in SQL Server Agent;

- Worked directly with Solution Analysts and Technical Architects to translate business requirements into conceptual and logical data models;
- Created ETL packages with different control flow options and data flow transformations such as Conditional Split, Multicast, Union all and Derived Column and etc;
- Converted old ETL packages to the new version (did optimization by applying more stored procedures instead of using control flow tasks);
- Designed, developed, and supported complex integration processes (including interfaces) using SQL Server technology, stored procedures and SQL program code;
- Involved in organizing conceptual data model, logical data model, physical data model, data mapping and data normalization;
- Performance Tuning of Stored Procedures and SQL queries using SQL Profiler and Index Tuning Wizard in SSIS;
- Created stored procedure using nodes method to load XML file data into SQL server tables;
- Created interactive, accurate, insightful, and well-organized Power BI reports/dashboards to support businesses in their data driven strategic planning;
- Developed filters, reports, dashboards and created chart types, visualizations and complex calculations to manipulate the data in Power BI.