**Abdul**

**Sr. Data Analyst/Modeler**

**Email:** [abdul.analyst90@gmail.com](mailto:abdul.analyst90@gmail.com) **Phone: 3092015755**

**SUMMARY:**

* About 9+ years of experience in Data analysis and design of Data processing, Data warehousing, Data Quality, Data modeling & Business Intelligence development projects with complete end-to-end SDLC process.
* Expert in writing SQL queries and optimizing the queries in Oracle 10g/11g/12c, DB2, Netezza, SQL Server 2008/2012/2016 and Teradata 13/14.
* Performed data analysis and data profiling using complex SQL on various sources systems including Oracle and Teradata and worked on Teradata SQL queries, Teradata Indexes, Utilities such as Mload, Tpump, Fast load and Fast Export.
* Hands on experience Hadoop framework and its ecosystem like Distributed file system (HDFS), Hive, Sqoop, Spark and Proficient in Hive Query language and experienced in hive performance optimization using Static-Partitioning, Dynamic-Partitioning, Bucketing and Parallel Execution concepts.
* Good experience in Data Profiling, Data Mapping, Data Cleansing, Data Integration, Data Analysis, Data Quality, Data Architecture, Data Modelling, Data governance, Metadata Management & Master Data Management.
* Expertise in Data Modeling, created various Conceptual, Logical and Physical Data Models for DWH projects. Created first of a kind unique Data Model for an Intelligence Domain.
* Experienced working with Excel Pivot and VBA macros for various business scenarios and involved in data Transformation using Python scripts in AWS EMR, AWS RDS and AWS Glue.
* Experienced in designing & implementing many projects using various set of ETL/BI tools involving the latest features & product trends, experience in technologies such as Big-data, Cloud Computing (AWS) & In-memory Apps.
* Experienced in continuous Performance Tuning, System Optimization & many improvements for BI/OLAP systems & traditional Databases such as Oracle, SQL, DB2 and many high-performance databases.
* Well versed in Normalization / De-normalization techniques for optimum performance in relational and dimensional database environments and implemented various data warehouse projects in Agile Scrum/Waterfall methodologies.
* Expertise in writing SQL queries and optimizing the queries in Oracle, SQL Server 2008/12/16 and Teradata and involved in developed and managing SQL, Python code bases for data cleansing and data analysis using Git version control.
* Excellent Software Development Life Cycle (SDLC) with good working knowledge of testing methodologies, disciplines, tasks, resources and scheduling.
* Have good exposure on working in offshore/onsite model with ability to understand and/or create functional requirements working with client and also have Good experience in requirement analysis and generating test artifacts from requirements docs.

**Technical Skills:**

* **Analysis and Modeling Tools:** Erwin 9.6/9.5/9.1, Oracle Designer, ER/Studio and SAS
* **Languages:** SQL, Python, T-SQL and PL/SQL
* **Database Tools:** Microsoft SQL Server 2016/2014/2012, Teradata 15/14, Oracle 12c/11g/10g, MS Access, Poster SQL, Netezza, DB2, Snowflake, HBase, MongoDB and Cassandra
* **ETL Tools:** SSIS, Informatica Power 9.6/9.5 and SAP Business Objects.
* **Cloud:** AWS S3, Redshift and Glue
* **Operating System**: Windows, Dos and Unix.
* **Reporting Tools**: SSRS, Business Objects, Crystal Reports and Tableau
* **Tools & Software's**: TOAD, MS Office, BTEQ, Teradata SQL Assistant and Netezza Aginity
* **Big Data Technologies**: Hadoop, HDFS, Hive, HBase, Sqoop, Oozie and MongoDB

**Work Experience**

**Target- Chicago IL**

**Sr. Data Analyst/Modeler**

**Jan 2020 to Present**

**Responsibilities:**

* Performed System Analysis & Requirements Gathering related to modeling, ETL, Data Quality, MDM, Dashboards and Reports and captured enhancements from various entities and provided Impact analysis.
* Designed the Logical Data Model using ERWIN 9.64 with the entities and attributes for each subject areas and involved in Dimensional modeling (Star Schema) of the Data warehouse and used Erwin to design the business process, dimensions and measured facts and developed and maintained an Enterprise Data Model (EDM) to serve as both the strategic and tactical planning vehicles to manage the enterprise data warehouse.
* Used Hive Context which provides a superset of the functionality provided by SQLContext and Preferred to write queries using the HiveQL parser to read data from Hive tables (fact, syndicate) and involved in creating Hive tables, and loading and analyzing data using hive queries
* Selecting the appropriate AWS service based on data, compute, database, or security requirements and defined and deployed monitoring, metrics, and logging systems on AWS and responsible for full data loads from production to AWS Redshift staging environment.
* Designed both 3NF data models for ODS, OLTP systems and dimensional data models using star and snow flake Schemas and creating dimensional data models based on hierarchical source data and implemented on Teradata achieving high performance without special tuning.
* Analyzed, transformed, and contextualised a variety of ingested data - social data, GIS data, POI& AOI data, and some consumer behaviour data for building direct marketing predictive models.
* Clean data and processed third party spending data into manoeuvrable deliverables within specific format with Excel macros and python libraries such as NumPy etc.
* Involved in working on converting SQL Server table DDL, views and SQL Queries to Snowflake and applied data analysis, data mining and data engineering to present data clearly and reverse engineered some of the databases using Erwin.
* Used python script to connect to oracle pull the data or make the data files and load those data to snowflake and worked on RDS for implementing models and data on RDS.
* Developed several Python Scripts to consume, massage and refine data from different data sources before loading into RedShift and Snowflake tables.
* Used BTEQ for SQL queries and TPump for loading data from source into Teradata tables and used Teradata utilities such as Fast Export, MLOAD for handling various tasks.
* Integrated SAS datasets into Excel using Dynamic Data Exchange, using SAS to analyze data, statistical tables, listings and graphs for reports and using SAS to read, write, import and export to another data file formats including delimited files, spreadsheet, Microsoft excel and access tables.
* Prepared high level & qualitative analysis to observe patterns and relationships in data and produce reports using with SAS and Tableau visualization.
* Generated, wrote and run SQL script to implement the DB changes including table update, addition or update of indexes, creation of views and store procedures.
* Created data dictionary, Data mapping for ETL and application support, DFD, ERD, mapping documents, metadata, DDL and DML as required.

**Technology:** Erwin 9.64, Python, Oracle 12c, Sqoop, Hive, H-Base, PySpark, Teradata, MS SQL, AWS Redshift, S3, Snowflake, Apache Hadoop, Informatica Data Quality, Snowflake, XML Files, SAS, SAP BO, Tableau, PL/SQL.

**Sr. Data Analyst/Modeler**

**BMO Harris- Chicago IL**

**Oct 2016 to Dec 2019**

**Responsibilities:**

* Performed as a Data Analysis, Data Modeling, Data Migration and data profiling using complex SQL on various sources systems including Oracle and Teradata.
* Managed Logical and Physical Data Models in ER Studio Repository based on the different subject area requests for integrated model. Developed Data Mapping, Data Governance, and transformation and cleansing rules involving OLTP, ODS.
* Written Python Scripts, mappers to run on Hadoop distributed file system (HDFS) and performed troubleshooting, fixed and deployed many Python bug fixes of the two main applications that were a main source of data for both customers and internal customer service team.
* Enforced referential integrity in the OLTP data model for consistent relationship between tables and efficient database design and performed/automated many ETL related tasks including data cleansing, conversion, and transformations to load Oracle based Data Warehouse.
* Load data into Amazon Redshift and use AWS Cloud Watch to collect and monitor AWS RDS instances and designed and developed ETL/ELT processes to handle data migration from multiple business units and sources including Oracle, Postgres, MSSQL, Access and others.
* Created SQL scripts and analyzed the data in MS Access/Excel and Worked on SQL and SAS script mapping and successfully optimized codes in Python to solve a variety of purposes in data mining and machine learning in Python.
* Designed and implemented multiple dashboards using Power BI - PowerPivot & Power Query tools for in house metrics.
* Developed, managed and validated existing data models including logical and physical models of the Data Warehouse and source systems utilizing a 3NF model.
* Implemented logical and physical relational database and maintained Database Objects in the data model using ER Studio and used Star schema and Snowflake Schema methodologies in building and designing the Logical Data Model into Dimensional Models.
* Involved in debugging and Tuning the PL/SQL code, tuning queries, optimization for the Oracle, and DB2 database.
* Developed automated processes for flattening the upstream data from Cassandra which in JSON format. Used Hive UDFs to flatten the JSON Data.
* Developed several behavioural reports and data points creating complex SQL queries and stored procedures using SSRS and Excel and generated periodic reports based on the statistical analysis of the data using SQL Server Reporting Services (SSRS) and generated reports using Global Variables, Expressions and Functions using SSRS.
* Validated the data of reports by writing SQL queries in PL/SQL Developer against ODS and involved in user training sessions and assisting in UAT (User Acceptance Testing).
* Perform troubleshooting, fixed and deployed many Python bug fixes of the two main applications that were a main source of data for both customers and internal customer service team.
* Involved in developing advanced ANSI SQL queries to extract, manipulate, and/or calculate information to fulfil data and reporting requirements including identifying the tables and columns from which data is extracted.

**Technology:** ER Studio 16.5, Oracle 12c, Python, Shell Scripting, UNIX, Hadoop, Spark, Hive, MongoDB, Cassandra, Windows7, XML, SQL, PL/SQL, T-SQL, Datastage, AWS S3, Redshift, UNIX, Agile, Informatica, MDM, Teradata, MS Excel, MS Access, Metadata, SAS, SQL Server, Tableau, Netezza, SSRS, Teradata SQL Assistant, Power BI, DB2, Netezza Aginity.

**Data Analyst/Modeler**

**CVS, Hartford CT**

**Dec 2014 to Sept 2016**

**Responsibilities:**

* Gather and analyze business data requirements and model these needs. In doing so, work closely with the users of the information, the application developers and architects, to ensure the information models are capable of meeting their needs.
* Performed Business Area Analysis and logical and physical data modeling for a Data Warehouse utilizing the Bill Inmon Methodology and also designed Data Mart application utilizing the Star Schema Dimensional Ralph Kimball methodology.
* Designed and Developed logical & physical data models and Meta Data to support the requirements using Erwin and designed the ER diagrams, logical model (relationship, cardinality, attributes, and, candidate keys) and physical database (capacity planning, object creation and aggregation strategies) for Oracle and Teradata as per business requirements using Erwin
* Worked on multiple Data Marts in Enterprise Data Warehouse Project (EDW) and involved in designing OLAP data models extensively used slowly changing dimensions (SCD).
* Designed 3rd normal form target data model and mapped to logical model and involved in extensive DATA validation using ANSI SQL queries and back-end testing and generated DDL statements for the creation of new ERwin objects like table, views, indexes, packages and stored procedures.
* Design MOLAP/ROLAP cubes on Teradata Database using SSAS and used SQL for Querying the database in UNIX environment and creation of BTEQ, Fast export, Multi Load, TPump, Fast load scripts for extracting data from various production systems
* Developed automated procedures to produce data files using Microsoft Integration Services (SSIS) and performed data analysis and data profiling using complex SQL on various sources systems including Oracle and Netezza
* Developed mapping spreadsheets for (ETL) team with source to target data mapping with physical naming standards, data types, volumetric, domain definitions, and corporate meta-data definitions.
* Used CA Erwin Data Modeler (Erwin) for Data Modeling (data requirements analysis, database design etc.) of custom developed information systems, including databases of transactional systems and data marts.
* Creation of multiple highly interactive Power BI and SSRS reports for supervisors and managers, which enables them to see the current status of the work, shift allocation for the next day at location, district and borough level.
* Identified and tracked the slowly changing dimensions (SCD I, II, III & Hybrid/6) and determined the hierarchies in dimensions.
* Worked on data integration and workflow application on SSIS platform and responsible for testing all new and existing ETL data warehouse components.
* Designing Star schema and Snow Flake Schema on Dimensions and Fact Tables and worked with Data Vault Methodology Developed normalized Logical and Physical database models.
* Transformed Logical Data Model to Physical Data Model ensuring the Primary Key and Foreign key relationships in PDM, Consistency of definitions of Data Attributes and Primary Index considerations.
* Generated various reports using SQL Server Report Services (SSRS) for business analysts and the management team and wrote and running SQL, BI and other reports, analyzing data, creating metrics/dashboards/pivots/etc.
* Working along with ETL team for documentation of transformation rules for data migration from OLTP to warehouse for purpose of reporting.
* Involved in writing T-SQL working on SSIS, SSRS, SSAS, Data Cleansing, Data Scrubbing and Data Migration.

**Technology:** SQL Server 2012, Erwin9.1, Oracle, PowerBI, Informatica, RDS, JDBC, NOSQL, Python, MySQL, PostgreSQL, Teradata, SSRS, SSIS, SQL, DB2, Shell Scripting, Tableau, Excel, MDM, Agile.

**Data Analyst**

**TYSYS, Columbus GA**

**June 2012 to November 2014**

**Responsibilities:**

* Attended and participated in information and requirements gathering sessions and translated business requirements into working logical and physical data models for Data Warehouse, Data marts and OLAP applications.
* Performed extensive Data Analysis and Data Validation on Teradata and designed Star and Snowflake Data Models for Enterprise Data Warehouse using ERWIN.
* Created and maintained Logical Data Model (LDM) for the project includes documentation of all entities, attributes, data relationships, primary and foreign key structures, allowed values, codes, business rules, glossary terms, etc.
* Integrated data from various Data sources like MS SQL Server, DB2, Oracle, Netezza and Teradata using Informatica to perform Extraction, Transformation, loading (ETL processes) Worked on ETL development and Data Migration using SSIS and (SQL Loader, PL/SQL).
* Created Entity/Relationship Diagrams, grouped and created the tables, validated the data, identified PKs for lookup tables.
* Created workspace and content packs for business user to view the developed reports and scheduled automatic refresh and scheduling refresh in power BI service and wrote calculated columns, Measures query’s in Power BI desktop to show good data analysis techniques.
* Involved in Designed and Developed logical & physical data models and Meta Data to support the requirements using ERWIN.
* Involved using ETL tool Informatica to populate the database, data transformation from the old database to the new database using Oracle.
* Involved in modeling (Star Schema methodologies) in building and designing the logical datamodel into Dimensional Models and Performance query tuning to improve the performance along with index maintenance.
* Involved in the creation, maintenance of Data Warehouse and repositories containing Metadata.
* Wrote and executed unit, system, integration and UAT scripts in a Data Warehouse projects and Created components to extract application messages stored in XML files.
* Wrote and executed SQL queries to verify that data has been moved from transactional system to DSS, Data Warehouse, and data mart reporting system in accordance with requirements.
* Responsible for Creating and Modifying T-SQL stored procedures/triggers for validating the integrity of the data.
* Worked on Data Warehouse concepts and dimensional data modelling using Ralph Kimball methodology.
* Used Power BI power query to extract data from external sources and modify data to certain format as required for consumption in the dashboards.
* Developed separate test cases for ETL process (Inbound & Outbound) and reporting.

**Technology**: Oracle , MS Visio, PL-SQL, Microsoft SQL Server, Power BI, T-SQl, Rational Rose, Data warehouse, OLTP, OLAP, ERWIN , Informatica 9.x, Windows, SQL, PL/SQL, SQL Server, XML files, Talend Data Quality, Oracle 9i/10g, Flat Files, Windows , SVN.

**Education: Bachelor of Science in Computer Science**