**Rafeeq**

**Sr. Data Architect / Modeler**

**Email:** **rafeeq0770@gmail.com**

**Ph: (619) 940-5312‬**

**Summary:**

* **9** years of Industry experienced in IT with solid understanding of **Data Modeling, Data Analysis**, **Data Architect** Evaluating Data Sources and strong understanding of **Data Warehouse/Data Mart Design, BI, OLAP, OLTP,** Client/Server applications.
* Strong Knowledge of Data Warehouse Modeleing and Star Schema, Snow flake Schema, FACT and Dimensional Tables.
* Strong Experience in **Big Data Hadoop Ecosystem** in **ingestion, storage, querying, processing** and analysis of big data.
* Experience in **Dimensional Data Modeling, Star/Snowflake schema, FACT** & **Dimension tables**.
* Experience with emerging technologies such **Big Data, Hadoop, and NoSQL**.
* Strong experience in analyzing/ Data Transformation of large amounts of data sets writing **Pig** scripts and **Hive, AWS EMR, AWS RDS**. Extensive knowledge in Hadoop stack components viz. **Apache Hive, Pig Scripting**, etc.
* Experience in analyzing data using **Hadoop** Ecosystem including **HDFS, Hive, HBase, Zookeeper, PIG, Sqoop, Flume**.
* Hands on experience in Normalization (**1NF, 2NF, 3NF** and **BCNF**) Denormalization techniques for effective and optimum performance in **OLTP** and **OLAP** environments.
* Experience in cloud development on **Amazon AWS, EC2, EC3,** Elastic Search, Redshift and Basic on **Azure.**
* Experience in BI/DW solution (ETL, **OLAP, Data mart**), Informatica, BI Reporting tool like **Tableau** and **QlikView** and also experienced leading the team of application, ETL, BI developers, Testing team.
* Good experience in working with different ETL tool environments like **SSIS, Informatica** and reporting tool environments like SQL Server Reporting Services **(SSRS)**, **Cognos** and **Business Objects.**
* Proficient in UML Modeling like Use Case Diagrams, Activity Diagrams, and Sequence Diagrams with **Rational Rose** and **MS Visio**.
* Experienced in Technical consulting and end-to-end delivery with **data modeling**, **data governance** and design - development - implementation of solutions
* Experience in writing expressions in **SSRS** and Expert in fine tuning the reports. Created many Drill through and Drill Down reports using **SSRS.**
* Solid knowledge of **Data Marts**, Operational Data Store **(ODS), OLAP, Dimensional Data Modeling** with Ralph Kimball Methodology (Star Schema Modeling, Snow-Flake Modeling for FACT and Dimensions Tables) using Analysis Services.
* organizing data as per the business requirements using **Erwin, ER Studio** in both **OLTP** and **OLAP** applications.
* Expertise lies in Data Modeling, Database design and implementation of **Oracle, AWS Redshift** databases and Administration, Performance tuning etc.
* Excellent experience in troubleshooting **SQL** queries, ETL jobs, data warehouse/data mart/data store models.Practical understanding of the Data modeling (Dimensional & Relational) concepts like **Star-Schema** Modeling, **SnowflakeSchema** Modeling, Fact and Dimension tables. Worked on data modeling using **ERWIN** tool to build logical and physical models.
* Skillful in Data Analysis using **SQL** on **Oracle, MS SQL Server, DB2 & Teradata**.
* Extensive experience in development of **T-SQL, Oracle PL/SQL Scripts**, Stored Procedures and Triggers for business logic implementation.
* Strong experience with architecting highly per formant databases using **PostgreSQL, PostGIS, MySQL** and **Cassandra.**
* Decode the **Teradata** and **SQL** queries to find all the data attributes involved and document for the purpose of development.
* Excellent understanding of Hub Architecture Style for **MDM** hubs the registry, repository and hybrid approach.
* Mapping the Risk Data elements to the Authoritative Data Source and documenting the Schema, Database, Table details for data modeling purpose.
* Good exposure on usage of **NoSQL** database.
* Experienced in understanding the **ETL** framework metadata to understand the current state ETL implementation.

**Technical Skills:**

**Data Modeling Tools:** IBM Info sphere Data Architect, ER Studio, Oracle Designer, Erwin R6/R9, Rational System Architect.

**Big Data:** HBase, PIG, Hive, Hadoop

**ETL/Data warehouse Tools**: Informatica 9.6/9.1/8.6.1/8.1, SAP Business Objects XIR3.1/XIR2, Web Intelligence, Talend, Pentaho.

**Cloud Platforms:** Azure, AWS EMR, AWS RDS, EC2, S3.

**Database Tools:**Oracle 12c/11G/10g/9i, Microsoft SQL Server12.0, Teradata 15.0, and MS Access

**BI Tools**: SAP Business Objects, Tableau 7.0/8.2, Tableau server 8.2, Tableau Reader 8.1, Crystal Reports.

**Tools & Utilities:**TOAD 9.6, Microsoft Visio 2010.

**Methodologies:** RAD, JAD, RUP, UML, System Development Life Cycle (SDLC), Waterfall Model.

**Packages:** SAP and Microsoft Visio, Share point, Microsoft Office 2010, Microsoft Project 2010.

**Operating Systems:** Windows, Centos, Sun Solaris, UNIX, Ubuntu Linux

**Version Tool:** VSS, SVN, CVS, SAP BO 4.1

**Work Experience:**

**Fannie Mae, Plano, TX Jan 2020 - Present**

**Sr. Data Modeler/ Architect**

**Responsibilities**

* Massively involved in **Data Architecture/ Modeler** role to review business requirement and compose source to target **data mapping** documents.
* Developed long term data warehouse roadmap and **Architectures, Designs** and builds the data warehouse framework per the roadmap.
* Specifies overall Data Architecture for all areas and domains of the enterprise, including Data Acquisition**, ODS, MDM, Data Warehouse, Data Provisioning** and ETL.
* Involved in end to end implementation of **Big** **data** design.
* Loaded data into **Hive Tables** from **Hadoop Distributed File System (HDFS)** to provide **Sql** access on **Hadoop** **data.**
* Created **HBase tables** to load large sets of **structured, semi-structured** and **unstructured** **data** coming from **NoSQL** and a variety of portfolios.
* Developed and configured on **Informatica MDM hub** supports the Master Data Management **(MDM), Business Intelligence (BI)** and **Data Warehousing** platforms to meet business needs.
* Developed **dimensional model** for **Data Warehouse/OLAP** applications by identifying required facts and dimensions.
* Interacted with Business Analyst, SMEs and other Data Architects to understanding Business needs and functionality for various project solutions.
* Play a key developer role and be part of building Data Infrastructure components such as Tracking, Messaging Queues, and Real Time Streaming & Batch **Data Pipelines.**
* Developing full life cycle software including defining requirements, prototyping, designing, coding, testing and maintaining software.
* Loaded and transformed large sets of structured, semi structured and unstructured data using **Hadoop/Big Data concepts**.
* Responsible for Big data initiatives and engagement including analysis, **brain storming, POC**, and architecture.
* Working on data profiling and analysis to create test cases for new **Architecture evaluation.**
* Translated business requirements into working **logical** and **physical** **data** **models** for **Data** **warehouse, Data marts and OLAP applications.**
* Develop and maintain our Big **Data pipeline** that transfers and processes several terabytes of data using Apache Spark, Python, Apache Kafka, Hive and Impala.
* Extensively used **Erwin r9.7** for **Data modeling**. Created Staging and Target Models for the **Enterprise Data Warehouse.**
* Defined facts, dimensions and designed the data marts using the Ralph **Kimball's** Dimensional Data Mart modeling methodology using Erwin.
* Extensively used **agile** **methodology** as the Organization Standard to implement the **data Models.**
* Established and maintained comprehensive **data** **model** documentation including detailed descriptions of business entities, attributes, and data **relationships**.
* Working on several ETL **Ab Initio** assignments to perform extract, transform and load data into Teradata and Oracle databases which had complex data models of Relational, Star and Snowflake schemas.
* Working on **SSIS** development using meta-data driven Architecture.
* Identified and streamlined complex queries which were causing iterations and effecting **database** and system performance.
* Experience in **Spark in Scala** Framework on Spark core, Spark SQL, and Spark Streaming
* Coordinated with the **Business Analyst** and prepared **Logical** and **Physical** **Data-models** as per the requirements involving multiple subject areas, domains and naming standards.
* Created **Data stage** jobs (ETL Process) for populating the data into the Data Warehouse constantly from different source systems like ODS, flat files, scheduled the same using **Data Stage** for System Integration testing.
* Created Data Movement standards for Micro Services and **ETL Talend with XML** Hierarchical Schema.
* Design and implementation of an **Azure data lake**.
* Design and development of the **data warehouse data ETL** Jobs sing **azure data factory**and **azure databricks.**
* Developed **data** **Mart** for the base data in **Star Schema, Snow-Flake Schema** involved in developing the **data** **warehouse** for the **database.**
* Design, build, and debug Hadoop, Hive, and Spark based data pipelines.
* Developed SQL Server Integration Services (SSIS, SSAS), Oracle and **Talend ETL jobs.**
* Design, develop, test, and deliver self-service infrastructure and business intelligence (BI) content including reports and dashboards to business professional using the **IBM Cognos BI** and **Qlikview**
* Deployed **SSRS** reports to Report Manager and created linked reports, snapshots, and subscriptions for the reports and worked on **scheduling** of the reports.
* Worked with **reverse engineering** Data Modelfrom Database instance and Scripts.
* **Metadata integration** across Enterprise Data dictionary, **Catalogs**, **Data Models** and **Informatica** establishing lineage and traceability.
* Generated parameterized queries for generating tabular reports using global variables, expressions, functions, and stored procedures using **SSRS.**
* Worked on **AWS** and architecting a solution to load data creates data models and run BI on it.
* Developed **Data Mapping, Data Governance**, and **Transformation** and **cleansing** **rules** for the **Master Data Management** Architecture involving **OLTP, ODS**.
* Developed and implemented **data cleansing, data security, data profiling** and **data monitoring processes.**
* Designed and developed architecture for data services ecosystem spanning Relational, **NoSql**, and **Big Data technologies**.
* Deployed and analyzed large chunk of data using **spark** in **Scala**.
* Azure logic app development to trigger the workflow on **Azure SQL DB.**
* Worked with the ETL team to document the **SSIS** **packages** for data extraction to Warehouse environment for reporting purposes.
* Documented **ER Diagrams, Logical** and **Physical models**, business process diagrams and process flow diagrams.
* Used **SSRS** to create reports, customized Reports, **on-demand reports, ad-hoc reports** and involved in analyzing multi-dimensional reports in **SSRS**.
* Presented the data scenarios via, **Erwin** logical models and excel mockups to visualize the data better.
* Responsible to create **conceptual, logical** and **physical** **data models**, of disparate information for report development.
* Designed both **3NF** **data** **models** for **ODS, OLTP** systems and **dimensional data models**
* Developed **data Mart** for the base data in **Star Schema, Snow-Flake Schema** involved in developing the **data warehouse** for the database.
* Developed **Python/Django** application for Google Analytics aggregation and reporting
* Conducted and participated in **JAD** **sessions** with the users, modelers, and developers for resolving issues.
* Developed and maintained **data** **dictionary** to create metadata reports for technical and business purpose.
* Developed automated data pipelines from various external data sources (web pages, API etc) to internal data warehouse (SQL server) then export to reporting tools like Datorama by **Python**.
* Worked with various **Teradata15** tools and utilities like **Teradata** **Viewpoint, Multi Load, ARC, Teradata Administrator, BTEQ** and other **Teradata** Utilities.
* Created one framework in **data stage** with all tables
* Created **DDL scripts** using **Erwin** and source to target mappings to bring the data from source to the warehouse.
* Hands on experience in analyzing data using **Hadoop Ecosystem including HDFS, Hive, Streaming, Elastic Search, Kibana, Kafka, HBase, Zookeeper, PIG, Sqoop, and Flume.**
* Proficient in **SQL** across a number of dialects (we commonly write **MySQL, PostgreSQL, Sql Server, and Oracle**)

**Environment**: Erwin 9.7, HBase 1.2, NoSQL, OLTP, OLAP, Azure, Python, Teradata 15, Talend, Kimball, Netezza, SQL Architect, MySQL, Oracle 12c, AWS, Hive 2.3, HDFS, Informatica, Snow- Flake, Star Schema, Sql, Hadoop, SSIS .

**Caterpillar, Peoria, IL Oct 2017 – Dec 2019**

**Sr. Data Modeler/ Analyst**

**Responsibilities**

* Designed and build relational database models and defines data requirements to meet the business requirements.
* Involved in developing Database Design Document including Data Model Conceptual, Logical and **Physical Models** using **ER studio.**
* Developing the Conceptual Data Models, Logical data models and transformed them to creating schema using **ER Studio**.
* Designing normalized and **star schema** data architectures using **ER Studio** and forward engineering these structures into Teradata.
* Working on **Amazon Redshift** and AWS and architecting a solution to load data, create data models and run BI on it.
* Actively involved in the Design and development of the **Star schema** data model.
* Implemented slowly changing and rapidly changing dimension methodologies; created aggregate fact tables for the creation of **ad-hoc** reports.
* Enforced referential integrity in the **OLTP** data model for consistent relationship between tables and efficient database design.
* Used **Star Schema** and **Snowflake Schema** methodologies in building and designing the **Logical Data Model** into Dimensional Models.
* Used **Python** to create **ETL pipelines**
* Created and maintained surrogate keys on the master tables to handle SCD type 2 changes effectively.
* Analyzed existing **SSIS package**, make changes to improve its performances, add standard logging and configuration system.
* Designed and implemented a Data Lake to consolidate data from multiple sources, using **Hadoop** stack technologies like **SQOOP, HIVE/HQL**.
* Created complex **stored procedures** to generate reports using **SSRS** and Extract Data.
* Written complex **SQL** queries for validating the data against different kinds of reports generated by Business Objects **XIR2**.
* Developed and deployed Data Warehouse infrastructure using SSIS for Data integration and SSRS to automate reports generations
* Implemented well-documented, well-architected, high-quality code of all enterprise applications **AWS S3**environment
* Designing Logical data models and Physical Data Models using **ER Studio**.
* Conducted numerous POCs to efficiently import large data sets into the database from **AWS S3** Bucket.
* Designed semantic layer data model. Conducted performance optimization for **BI** infrastructure.
* Involved in the creation, maintenance of Data Warehouse and repositories containing **Metadata**.
* Performed **Hive** programming for applications that were migrated to big data using **Hadoop**.
* As an Architect implement **MDM** hub to provide clean, consistent data for a SOA implementation.
* Installing and configuring the a 3-node Cluster in **AWS Linux Servers.**
* Used **javascripts** like plotly.js, D3.js and chart.js for creating data visualizations.
* Designed different type of STAR schemas like detailed data marts and Plan data marts, Monthly Summary data marts using **ER studio** with various Dimensions Like Time, Services, Customers and various **FACT** Tables.
* Developed and maintained data dictionary to create metadata reports for technical and business purpose.
* Extensive Data validation by writing several complex **SQL** queries and Involved in back-end testing and worked with data quality issues.
* Data Profiling, Mapping and Integration from multiple sources to **AWS S3**.
* Worked on AWS Redshift and RDS for implementing models and data on RDS and Redshift.
* Identify source systems, their connectivity, related tables and fields and ensure data suitably for mapping.
* Worked with **BTEQ** to submit **SQL** statements, import and export data, and generate reports in **Teradata**.
* Designed and documented Use Cases, Activity Diagrams, Sequence Diagrams, OOD (Object Oriented Design) using **UML** and **Visio**.
* Performed data cleaning and data manipulation activities using **NOSQL** utility.
* Designed and Developed **Oracle PL/SQL** Procedures and **UNIXShell** Scripts for Data Import/Export and Data Conversions.

**Environment:** ER Studio, Oracle 11g, MS-Office, SQL Architect, Hadoop, Hive, Pig, TOAD Benchmark Factory, Sqoop, SQL Loader, AWS S3, AWS RDS, PL/SQL, DB2,Javascript, Share Point, Python, MS-Office, SQL Server 2014,

**ATnT, San Diego, CA Jun 2016 – Sep 2017**

**Sr. Data Modeler/ Data Analyst**

**Responsibilities**

* Understood source business process and data model and do Data Analysis and map the requirement to source system.
* Created and documented the models, data Workflow Diagrams, Sequence Diagrams, Activity Diagrams and field mappings of all existing system.
* Create **Logical Data Model** for Staging, **ODS** and Data Mart and Time dimension as well.
* Generated comprehensive analytical reports by running **SQL** queries against current databases to conduct Data Analysis.
* Performed Data Analysis, Data Migration and data profiling using complex **SQL** on various sources systems including **Oracle** and **Teradata**.
* Worked with **Data governance, Data quality, data lineage**, Data architect to design various models and processes.
* Develop Logical and Physical data models that capture current state/future state data elements and data flows using **Erwin**.
* Designed Database with prominent activities like maintaining sequences, index, primary key, foreign key, manipulating columns and tables.
* Built reports using **SSRS** using complex **stored procedures**as a sources.
* Translated business requirements into working logical and physical data models for Data warehouse, Data marts and **OLAP** applications.
* Involved in preparing modeling estimates, **DFD**, **Volumetric**, defining enterprise standards for best practices.
* Performed Data modeling using **Erwin Data Modeler**. Identified objects and relationships and how those all fit together as logical entities, these are then translated into physical design using forward engineering
* Identified the entities and relationship between the entities to develop Conceptual Model using **ERWIN**.
* Enforced referential integrity in the **OLTP** data model for consistent relationship between tables and efficient database design.
* Worked with **BTEQ** to submit **SQL** statements, import and export data, and generate reports in **Teradata**.
* Gathering the business requirements from the respective application team (Data Architect's / Application Owner's).
* Extensively used **SQL**, **Transact SQL** and **PL/SQL** to write stored procedures, functions, packages and triggers.
* Used forward engineering to create a physical data model with **DDL** that best suits the requirements from the Logical Data Model.
* Worked with data compliance teams, Data governance team to maintain data models, **Metadata**, Data Dictionaries; define source fields and its definitions.

**Environment:** ERWIN, Windows, UNIX, Oracle11g, SQL SERVER 12.0, PL/SQL, DB2, Teradata SQL assistant, DQ analyzer, HDFS, HBase, YARN, IBM Optim Tool, IBM Rational Clear Case and Clear Quest, Toad 9.7,

**Citi Bank - New York, NY Feb 2015 – May 2016**

**Sr. Data Analyst**

**Responsibilities**

* Analyzed the business requirements by dividing them into subject areas and understood the data flow within the organization
* Attended and participates in information and requirements gathering sessions.
* Designed and developed **stored procedures**, queries and views necessary to support **SSRS reports**.
* Load data from **MS Access** database to **SQL Server** using**SSIS** (creating staging tables and then loading the data).
* Performed Data Analysis and Data validation by writing **SQL queries** and Regular expressions.
* Created reports using SQL Server Reporting Services **(SSRS).**
* Wrote standard **SQL** Queries to perform data validation and created excel summary reports **(Pivot tables and Charts).**
* Database Design (Conceptual, Logical and Physical) for **OLTP** and **OLAP** systems.
* Created several types of reports using **SSRS** like parameterized, drill down and drill through reports.
* Creating complex **SQL queries** and scripts to extract and aggregate Data to validate the accuracy of the data.
* Worked on developing the tool to extract the data from db2 database and conducted **Meta Data** and Data analysis.
* Created documents for technical & business user requirements during requirements gathering sessions.
* Turned **SQL** queries to make use of data base indexes, and analyzed the data base objects.
* Created Logical and Physical **EDW** models and **data marts**.
* Developed **SSIS Packages** to transfer the data between **SQL Server** database and files.
* Experienced in data migration and cleansing rules for the integrated architecture **(OLTP, ODS, DW)**.
* Created **SQL-Loader** scripts to load legacy data into Oracle staging tables and wrote **SQL queries** to perform Data Validation and Data Integrity testing.
* Performed drill down analysis reports using SQL Server Reporting Services.
* Managed all indexing, debugging and query optimization techniques for performance tuning using **T-SQL**.
* Developed the logical and physical model from the conceptual model developed using a tool **Erwin** by understanding and analyzing business requirements.
* Normalized the database up to **3NF** to put them into the star schema of the data warehouse.
* Worked with **BTEQ** to submit **SQL** statements, import and export data, and generate reports in **Teradata**.
* Handled data loading operations from flat files to tables using **NZLOAD** utility.
* Experienced in data cleansing for accurate reporting. Thoroughly analyzed the data and integrated different data sources to process matching functions.
* Applied data naming standards, created the data dictionary and documented data model translation decisions and also maintained **DW metadata**.
* Created DDL scripts for implementing Data Modeling changes. Created **ERWIN** reports in **HTML**, RTF format depending upon the requirement, Published Data model in model mart, created naming convention files, co-coordinated with DBAs' to apply the data model changes.
* Extensively used Normalization techniques **(up to 3NF)**.
* Writing complex queries using **Teradata SQL**.
* Worked with the **ETL** team to document the transformation rules for data migration from source to target systems.

**Environment:** MS Visio, Business Objects, Informatica, ERWIN r7.2, PL/SQL, SSRS, SSIS, MS SQL, Windows NT, Linux, Sybase Power Designer, Oracle 9i, SQL Server, Windows , MS Excel, MS Access.

**Cigna HealthCare, Nashville, TN Mar 2012 - Jan 2015**

**Data Analyst**

**Responsibilities**

* Responsibilities included source system analysis, data transformation, loading, validation for data marts, operational data store and data warehouse.
* Performed **Data Analysis**, Data Migration and data profiling using complex SQL on various sources systems including Oracle and **Teradata**.
* Generated comprehensive analytical reports by running **SQL queries** against current databases to conduct **Data Analysis**.
* Developed database objects such as SSIS Packages, **Tables, Triggers**, and Indexes using T-SQL.
* Worked in **Data Analysis**, data profiling and data governance identifying Data Sets, Source Data, **Source Meta Data,** Data Definitions and Data Formats.
* Extracted and analyzed data to provide data-driven product recommendations.
* Designed and implemented business intelligence to support sales and operations functions to increase customer satisfaction
* Developed Data Mapping, Data Governance, Transformation and Cleansing rules for the Master Data Management Architecture involving **OLTP, ODS** and **OLAP**.
* Worked in importing and cleansing of data from various sources like **Teradata, Oracle, flat files, SQL Server** with high volume data.
* Built reports using **SSRS.** Developed efficient reporting solutions. Set up subscriptions as needed for **SSRS.**
* Used **SQL queries** for Data verification and validation
* Created Drill Down, Drill Through, Sub and Linked reports using the **SSRS** as well as managed the subscription and authentication of these reports.
* Involved with all the phases of Software Development Life Cycle **(SDLC)** methodologies throughout the project life cycle.
* Created logical data model from the conceptual model and it's conversion into the physical database design using **Erwin**.
* Analyzed the data and provide resolution by writing analytical/complex **SQL** in case of data discrepancies.
* Tuning and code optimization using different techniques like **dynamic SQL**, dynamic cursors, and **tuning SQL** queries, writing generic procedures, functions and packages.
* Responsible for Relational data modeling **(OLTP)** using **MS Visio** (Logical, Physical and Conceptual).
* Create and execute test scripts, cases, and scenarios that will determine optimal system performance according to specifications.
* Reverse Engineered **DB2** databases and then forward engineered them to **Teradata** using **Erwin**.
* Tested the database to check field size validation, check constraints, stored procedures and cross verifying the field size defined within the application with metadata.
* Extensively worked on development of mappings with **BODS** Transformations like Map Operation, Table Comparison, History Preserving, Key Generation, Pivot, Reverse Pivot etc.
* Responsible for design of logical and physical Data model for client's investment management **ODS** using dimensional modeling.

**Environment:** Oracle 8i, Developer 2000 with Forms 5.0 and Reports 3.0, O, SSIS, SSRS, PL/SQL, MS-Access, Sql Server, MS Office, MS Visio, Informatica Power center 5.1, Erwin, Teradata SQL Assistant.