**Name: Hrithik K**

**Sr. Data Analyst**

**PROFESSIONAL SUMMARY:**

* Overall, **10** years of professional experience in IT and Analysis, Design, Development, Documentation, Deployment, and Integration using **SQL** and **Big Data** technologies
* Experience in **Data analysist** with large sets of Structured and Unstructured data, **Data Acquisition**, Data Validation, Data Visualization and Data Warehousing.
* Proficient in **SQL** Querying, Stored Procedures, Data Analysis, Validation, and generating complex reports using a varied range of development tools.
* Proficiently analyze and transform data using **SQL Server** and **Excel**, adeptly develop and execute SQL scripts to generate data extracts and reports.
* Generate functional specification documents and conduct **Data Mapping Analysis** for seamless System Integrations.
* Develop, modify, and execute **SSIS** packages to efficiently load client's raw data from Excel to SQL Server.
* Perform **ETL/SSIS** imports and exports meeting specified requirements, using SSIS control flow, Data Task, SQL Tasks, and data conversions.
* Collaborate with users and report developers to develop functional specifications and **data mapping** for report development and system enhancements.
* Experienced in troubleshooting and performance tuning of **reports**, ensuring prompt issue resolution.
* Developed **Tableau** dashboards and reports from multiple data sources, incorporating maps, density charts, pie, and bar graphs.
* Experience in designing data models in **QlikView** and **Qlik Sense**, implementing its **ETL** abilities.
* Acquired knowledge of big data technologies, including **Hadoop**, **Spark**, **Hive**, and **Tableau** Framework.
* Demonstrated expertise in creating comprehensive Test Plans, Test Scripts, and Test Cases for various testing phases, including Functional, System, Integration, **UAT**, and Regression Testing.
* Experienced in cloud technologies with focus on AWS and Azure, particularly proficient in data engineering services such as **AWS EMR**, **S3**, **DMS**, **Redshift**, **Cosmos DB**, and **Azure Data Factory** (ADF).
* Expertise in leveraging the **Databricks** platform and utilizing **Apache Spark** to construct robust data lake solutions, including **Databricks-managed** Delta tables, while employing Scala for efficient ingestion code development.
* Experience in using **Azure Data** Factory to orchestrate linked services, **data sets**, activities, pipelines, and triggers, with a strong background in working with **Azure** integration runtimes.
* Strong interpersonal skills, with demonstrated proficiency in verbal and written communications to maintain effective work relationships.
* Experience in using **SQOOP** for importing and exporting data from RDBMS to HDFS and Hive.
* Experience on MS SQL Server, including **SSRS, SSIS**, and **T-SQL.**
* Extensive **ETL tool** experience using **IBM** Infosphere/**WebSphere** DataStage, Accentual DataStage.
* Worked on DataStage tools like DataStage Designer, **DataStage** Director and DataStage Administrator.
* Strong understanding of the principles of **Data** **Warehousing** using fact tables, dimension tables and star/**snowflake** schema modeling.
* Comprehensive experience in **Pharmaceutical**/Biotech/Dietary Industry in versatile areas of Project Management; Quality Assurance; Qualifications; Facilities, Utilities and Equipment Validation.
* Strong understanding of the principles of **Data warehousing**, **Fact Tables**, **Dimension Tables**, Star and **Snowflake** schema modeling.
* Strong experience with **Database** performance tuning and optimization, query optimization, **index tuning**, caching and buffer tuning.
* Extensive experience in relational **Data modeling**, Dimensional data modeling, **logical/Physical** Design**, ER Diagrams** and **OLTP and OLAP** System Study and Analysis.
* Extensive experience in Enterprise Information Management and Architecture technologies including Information **Lifecycle, Master Data Management** and Business Intelligence.
* Extensively used **ERWIN and Power Designer** to design Logical and Physical Data Models, to forward and reverse engineering data models and publishing data model to acrobat **PDF files**.
* Excellent leadership, organization skills and requirement gathering skills such as **JAD**.
* Experienced in writing Toad and **PL/SQL** Stored Procedures, **Triggers** and **Functions**.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **BI Tools** | Tableau Desktop, MS Excel, MS SQL Server, Business Intelligence Development Studio (SSIS, SSRS, SSAS), SAS visual analytics, QlikView/Qlik Sense |
| **ETL tools** | DataStage, Talend, Informatica, Alteryx |
| **Databases** | Oracle 10g/11g/12c Teradata, MS SQL Server 2016/2017/2019, DB2, PostgreSQL |
| **Tools** | MS Excel, MS Word, MS PowerPoint, MS project, MS Front, MS Expressions Web, TOAD, SQL Developer, SQL Star Plus. |
| **Cloud Platform** | Amazon AWS, Azure Data Factory, Azure SQL Database, Azure SQL Data Warehouse, Azure Analysis Services, Azure Data Lake, Data Factory |

**PROFESSIONAL EXPERIENCE:**

**SciPlay - Austin, TX Nov 2021 – Till Date**

**Role: Sr. Data Analyst**

**Responsibilities:**

* Coordinated with business users and supervisors to collect business requirements effectively.
* Converted user requirements into business requirements document, as well as **functional** and technical requirements documents.
* Collaborated on designing and implementing tables, views, indexes, stored procedures, **triggers**, queries, and **macros**.
* Spearheaded partitioning of tables with bulk inserts, deletes, and **updates**, resulting in significant performance enhancements.
* Consulted with application development **business analysts** to translate business requirements into data design requirements used for driving innovative **data designs** that meet business objectives.
* Involved in **information-gathering** meetings and **JAD** sessions to gather business requirements, deliver business requirements document and preliminary logical data model.
* Performed installation and proficiently managed **MS SQL Server**, ensuring optimal database performance and reliability.
* Extracted data from **XML** and flat files utilizing SSIS.
* Created parameterized/Drilldown, Drill through & Cascaded reports utilizing **SSRS** and **Crystal Reports**.
* Collaborated on the development and debugging of **Informatica** mappings, mapplets, sessions, and workflows.
* Spearheaded the creation, maintenance, modification, and optimization of **SQL Server databases**, while ensuring robust database security and implementing auditing protocols.
* Conceived the design, deployment, and thorough testing of requested reports.
* Designed Data Flow Diagrams to **visualize** and document the data flow from legacy systems into the application database tables, incorporating key checkpoints for **testing** and verification.
* Conducted comprehensive evaluation of **database** performance and perform maintenance operations such as tuning, backup and restoration.
* Designed and implemented **data profiling** and data quality improvement solution to **analyze**, **match**, cleanse, and consolidate **data before loading** into **data warehouse**.
* Extensively used **ERWin** to design Logical, Physical Data Models, and **Relational database** and to perform forward/reverse engineering.
* Responsible in designing new **FACT** or Dimension Tables to existing Models.
* Implemented stored procedures, functions, **views**, **triggers**, packages in **PL/SQL**.
* Conducted analysis of long-running **queries**, employing tuning complex queries to optimize application and system performance.
* Contributed to test planning and execution, creating a comprehensive test plan for **User Acceptance Testing (UAT).**
* Executed a well-planned Backup and Restoration strategy to safeguard **databases** against data loss, mitigating the impact of potential failures caused by various issues.

**Environment:** MS SQL Server 2019(15.0), MS Access 2019(17.0), Oracle, SQL Server Integration Services (SSIS), SSRS Import & Export Data, SQL Reporting Services, XML, Informatica.

**SCL Health Group – Denver, CO Feb 2019 – Oct 2021**

**Role: Data Engineer**

**Responsibilities:**

* Designed and executed data **pipelines** by deploying multiple **Spark** clusters in a multi-cloud environment in **AWS** and **Azure**.
* Utilized the Databricks environment to construct a data lake in **S3** and blob storage, using **Apache Spark** in Scala.
* Implemented and managed various file formats such as **TEXTFILE**, **BINARY**, **CSV**, and **PARQUET**, using them for ingestion into **delta** tables and conducting querying and processing.
* Successfully ingested **3500 tables** from legacy systems into delta tables.
* Managed and updated the metadata configurations in **Cosmos DB**.
* Executed Scala code from Databricks notebooks using Databricks Notebooks Activity in **Azure Data Factory**, while concurrently designing and implementing pipelines.
* Implemented **Talend ETL** on both single and multi-server environments, including installation and configuration.
* Ingested all tables using **DM** and **Spark** **streaming**, populating delta tables.
* Transformed SQL Server scripts/Components of the **ETL** process into **Spark SQL**.
* Utilized Spark Streaming APIs to execute necessary transformations and actions in **real-time**, building the common learner data model by ingesting data from **Kafka**.
* Wrote scripts and indexing strategy for a migration to Confidential **Redshift** from **SQL Server** and **MySQL databases**.
* Implement software enhancements to port legacy software systems to **Spark** and **Hadoop** ecosystems on **Azure Cloud**.
* Used Pig as **ETL tool** and performing data fusion to do transformations, event joins, filters and some pre-aggregations before storing the data onto **HDFS**.
* Experience in resolving **on-going** maintenance issues and **bug fixes**; monitoring Informatica sessions as well as performance tuning of mappings and sessions.
* Installed and configured **Apache air flow** for workflow management and created workflows in **python**.
* Developed python code for different tasks, dependencies, **SLA** watcher and time sensor for each job for workflow management and automation using **Airflow** **tool**.
* Develop marketing and communication materials for data governance.
* Build servers using **GCP**, importing volumes, launching **EC2, RDS**, creating security groups, auto-scaling, **load balancers**(ELBs) in the defined virtual private connection.
* Development of test framework in **python**.
* Understanding **python** files in open stack environment and make necessary changes if needed.
* Collaborated with cross-functional teams to understand their analytical needs and trained colleagues on effectively using **QlikView** and **Qlik Sense** to independently explore data and generate reports.
* Designed and implemented quarterly and annual reports for diverse programs, while developing analytical models for customer engagement programs utilizing **SAS EG**, **Hadoop**, and **SQL** Server.

**Environment:** Apache Spark, SQL server, Hadoop, Azure Data Factory, Azure SQL Database, Azure SQL Data Ware house, Azure Analysis Services, Apache Kafka, QlikView/ Qlik Sense.

**KPMG – Santa Ana, CA May 2017 – Jan 2019**

**Role: Data Analyst**

**Responsibilities:**

* Implemented the project using **agile** methodology.
* Created a framework model to generate reports.
* Performed requirements gathering and **design analysis** for customer contact history for customer relationship management.
* Worked as business requirements **analysts** with subject matter experts to identify and understand requirements.
* Managed the conversion and migration of **Tableau** and **Cognos** reports from the **OLTP** system to the OLAP system and handled various versions of **Tableau** and **Cognos**.
* Collaborated with business users to facilitate **data querying** based on specific requirements.
* Implemented **data modeling** techniques to resolve dimension-dimension join challenges.
* Executed partitioning and indexing strategies to enhance performance **tuning**.
* Executed self-service requests by developing reports.
* Implemented Package-level and Object-level security on the **FM package**.
* Revised the data integrity of models considering the business.
* Directed the migration of reports from **Cognos 10.1** to **Cognos 11** environment and conducted through testing to ensure accuracy and functionality.
* Designed interactive dashboards using Cognos Analytics and Tableau within the **Hadoop Hive** environment.
* Employed **QlikView**'s **ETL capabilities** to efficiently extract, transform, and load data from various sources, ensuring data accuracy and consistency across **visualizations** and **reports**.
* Utilized ETL extensively to efficiently load data from a wide range of sources, including flat files, Oracle databases, and XML, into **Teradata**.
* Executed the project by extracting data from diverse sources, performing data transformation from these files, and loading the processed data into target (warehouse) **Teradata tables**.
* Demonstrated proficiency in creating mappings using transformations, such as Source qualifier, Aggregator, Expression, **Lookup**, **Router**, Filter, Joiner, and Update Strategy.
* Managed the Asset Management domain, handling data related to **investment funds** (Exchange Traded **Funds, mutual funds**, etc.) and **brokerage data**.
* Chaired as a Workflow Manager to create and manage sessions, monitoring, editing, scheduling, copying, aborting, and deleting sessions.

**Environment:** SQL, SQL server 2008, Quality center, MS Excel, MS project, MS Office suite, SSMS, Tableau Desktop 8.3, Tableau Server, QlikView/Qlik Sense.

**First Republic Bank – San Francisco, CA Jan 2016 – April 2017**

**Role: Big Data Analyst/ Developer**

**Responsibilities:**

* Designed and developed an end-to-end **Data warehouse** infrastructure from ground up on **Redshift**, enabling seamless management of large-scale data.
* Executed and managed **ETL** solutions, **streamlining** and automating operational processes.
* Implemented performance **optimizations** and tuned the Redshift environment, resulting in query performing up to **100x** faster for Tableau and **SAS** Visual Analytics.
* Executed migration of on-premises database structure to Redshift data warehouse.
* Performed ETL processes and data validation using **SQL Server Integration Services (SSIS).**
* Executed the implementation of **Workload Management (WLM)** in Redshift to prioritize basic dashboard queries over more complex longer-running ad hoc queries, resulting in a more reliable and accelerated reporting interface, achieving sub-second query responses for basic queries.
* Developed and published interactive **data visualization** dashboards, reports, and workbooks using **Tableau** and **SAS Visual Analytics**.
* Designed logical and physical data modeling on Redshift.
* Created and engineered ETL jobs to extract data from Salesforce replica and seamlessly load it into Data Mart in Redshift.
* Implemented and deployed monitoring metrics and logging systems in AWS.

**Environment:** SQL Server Integration Services, SQL Server, AWS Data Migration Services, Redshift, Airflow, AWS Data Pipeline, S3, Python, SAS Visual Analytics and Tableau.

**Client: Axis Bank, Hyderabad, India June 2012 – June 2014**

**Role: Jr. Data Analyst**

**Responsibilities:**

* Implemented data collection and qualification procedures to enhance efficiency.
* Generated spreadsheets in excel using advanced functions, **macros,** and **pivot tables**.
* Designed and developed **Tableau** dashboards to create data visualizations, effectively transforming complex datasets into actionable business insights.
* Generated daily data **queries** and reports on a daily, weekly, monthly, and quarterly basis.
* Supported data warehouse projects with logical and physical data modeling in **Oracle** **environment** and assured accurate delivery of **DDL** to the development teams.
* Documented the data requirements and system changes into detailed functional specifications. Took ownership of Source to Target Mapping and tracked and maintained changes.
* Performed Extraction. Transformation and Loading (**ETL**) using Informatica power center.
* Tracking trends, patterns, and correlation in case of **complex data** sets.
* Extracted the data provided by the users, **allocated** it to team members, and conducted a quality check prior to the final submission.
* Provided **SQL support** to front end users.
* Provisioned instances and databases as needed on **PostgreSQL**.
* Collaborated with team members to successfully accomplish special projects and meet project deadlines.
* Revised **data-related** issues including, **data quality** assurance, **consolidating data**, and **evaluating** existing data sources.

**Environment:** Tableau Desktop 8.2, Tableau server / Administrator, PostgreSQL, Oracle 12c, Microsoft Excel 2013, MS Access, Windows server 2012, Microsoft office 2013.

**EDUCATION:**

**Bachelor’s Degree** Computer Science BITS PILANI, India 2008 - 2012

**Master’s Degree** Data Science Texas A&M University, USA 2014 - 2015