Sai Rahul | Data Engineer |Email: [sai1exist@gmail.com](mailto:sai1exist@gmail.com)

Phone: (516) 653-1736

*References: Official References are provided up on request.*

**------------------------------------------------------------------------------------------------------------------------------------------**

**PROFESSIONAL SUMMARY:**

* Around 5+ years of experience in **Analyzing, Designing, Development, Implementation**, Maintenance and **Documentation** of Database and Client Server applications.
* Experience on Migrating SQL database to Azure data Lake, Azure data lake Analytics, Azure SQL Database, Data Bricks and Azure SQL Data warehouse and Controlling and granting database access and Migrating On premise databases to Azure Data lake store using Azure Data factory.
* Extract Transform and Load data from Sources Systems to Azure Data Storage services using a combination of Azure Data Factory, T-SQL, Spark SQL and U-SQL Azure Data Lake Analytics . Data Ingestion to one or more Azure Services - (**Azure Data Lake, Azure Storage, Azure SQL, Azure DW**) and processing the data in In **Azure Databricks.**
* Strong skills in visualization tools **Tableau, Power BI, Confidential Excel - formulas**, Pivot Tables, Charts and DAX Commands.
* Expert in generating on-demand and scheduled reports for business analysis or management decision using SQL Servers Reporting Services (SSRS), **Tableau** and **Power BI**. Periodic reporting is done on a daily, weekly, monthly and quarterly basis which helps the client.
* Created various types of charts like Heat Maps, Geocoding, Symbol Maps, Pie Charts, Bar Charts, Tree Maps, Gantts, Circle Views, Line Charts, Area Charts, Scatter Plots, Bullet Graphs and Histograms in Table Desktop, Power BI and Excel to provide better data visualization .
* Expert in generating on-demand and scheduled reports for business analysis or management decision using SQL Servers Reporting Services (SSRS), Tableau and POWER BI.
* Solid knowledge of Power BI and Tableau Desktop report performance optimization.
* Expert on maintaining and managing Tableau and POWER BI driven reports and dashboards.
* Good understanding of **Spark** Architecture including Spark Core, Spark SQL, Data Frames, Spark Streaming, Driver Node, Worker Node, Stages, Executors and Tasks..
* Experience in Database Design and development with Business Intelligence using **SQL Server 2014/2016, Integration Services (SSIS), DTS Packages, SQL Server Analysis Services (SSAS), DAX, OLAP Cubes, Star Schema and Snowflake Schema.**
* Excellent communication skills with excellent work ethics and a proactive team player with a positive attitude.
* Experienced in managing users, groups, sites, permissions and roles on Tableau server.
* Expertise in various phases of project life cycles (Design, Analysis, Implementation and testing).
* Experience in database administration and database performance tuning efforts to provide scalability and accessibility in a timely fashion, provide 24/7 availability of data, and solve end-user reporting and accessibility problems.
* Extensive experience in Designing Relational database model and Implementation as per business needs.
* In-depth knowledge of T-SQL, DBCC, DCL, DDL, DML Scripts and established relationships between tables using Primary Keys & Foreign Keys.
* Expertise in scripting T-SQL queries, complex stored procedures, user defined functions (UDF), Database triggers using tools like SQL profiler, and database tuning Advisor (DTA).
* Extensively used Joins and Sub-Queries to simplify complex queries involving multiple tables.
* Very strong T-SQL skill in creating and maintaining database objects like Indexes, Views, Stored Procedure, User Defined Functions, Triggers in SQL Server 2012.
* Created Complex ETL Packages using SSIS to extract data from staging tables to partitioned tables with incremental load.
* Experience working with Active Directory issues such as: logon failures, account lockouts, and authentication problems.
* Worked on Spark Machine Learning library for Recommendations, Coupons Recommendations, Rules Engine.
* Expertise in Debugging, Error logging, Error Handling and Production support for SSIS.
* Experience in managing and automating Control flow, Data flow, Events and **Logging** programmatically using Microsoft **.NET framework** for SSIS packages.
* Used **ProClarity** for accessing and analyzing **cubes** and **dimensions**, and also worked on integration of **SharePoint** with **SSRS**.
* Experience using Spark - SQL in Databricks for data extraction, transformation and aggregation from multiple file formats for analyzing & transforming the data to uncover insights into the customer usage patterns.
* Expertise in SQL Server Analysis Services (**SSAS**) **CUBE** design, Development and **Deploying.**
* Optimized large cubes by creating **partitions**, usage based optimization, and using proper storage modes like **MOLAP, HOLAP and ROLAP**.
* Experience on performing Data loads to Dimension, Fact tables, Data warehouse, Data marts and OLAP cubes.
* Involved in preparing and reviewing BRD, FRD documents to serve development and reporting
* Strong skills in requirements gathering, analysis and ETL design, coding, debugging
* Excellent communication and report writing skills, with the ability to effectively communicate results of analytics to diverse audience

#### TECHNICAL SKILLS

**Technical Databases:**Oracle, Teradata, SQL Datawarehouse, Azure, Databricks, Google BigQuery

**Skills Programming Languages:**SQL, Spark SQL, PLSQL, C, Python, Shell

**ETL Tools:**ADF, Alteryx Designer, Informatica, SSIS

**BI Visualization Tools:**Tableau 2020.1 , Power BI

**Database Access Tools:**SQL Developer, SQL \*Plus, Toad, SSMS, SQL Workbench, Oracle Forms

**Version Control:**Tortoise SVN, GIT

**Operating Systems:**Windows, Mac, Linux

**Cloud Technologies:**Azure, Databricks.

**Business Tools:**Atlassian Jira, Serena Business Manager, Confluence

**WORK EXPERIENCE**

**Mizuho bank, New York City May 2018 – Till date**

**Data Engineer**

Mizuho Bank, Ltd. is the integrated retail and corporate banking unit of Mizuho Financial Group, the third largest financial services companies in Japan. It is a major provider of financial services. Mizuho offers a range of financial services, including banking, securities, trust and asset management service.

**Responsibilities:**

* Gathering business requirements, business analysis and design various data products.
* Developing, testing and maintaining pipelines by connecting various data sources and building the final products.
* Creating Databricks notebooks using SQL, Python and automated notebooks using jobs.
* Creating Spark clusters and configuring high concurrency clusters using Azure Databricks to speed up the preparation of high-quality data.
* Experience using Spark - SQL in Databricks for data extraction, transformation and aggregation from multiple file formats for analyzing & transforming the data to uncover insights into the customer usage patterns.
* Extract Transform and Load data from sources Systems to Azure Data Storage services using a combination of Azure Data factory, T-SQL, Spark SQL, and U-SQL Azure Data Lake Analytics. Data ingestion to one or more Azure **services (Azure Data Lake, Azure Storage, Azure SQL, Azure DW)**and processing the data in **Azure Databricks**
* Designed and developed **Power B**I graphical and visualization solutions with business requirement documents and plans for creating interactive dashboards.
* Experience in building Tableau dashboards to provide the effectiveness of weekly campaigns and customer acquisition. Have built the trends over weekly, monthly, quarterly and yearly.
* Extensively used **Azure Data Lake** and worked with different file formats csv, parquet, delta.
* Written Oracle SQL, PLSQL Procedures and packages and automated procedures using DBMS scheduler. Performance tuning the packages and dashboards.
* Created procedures in **Azure SQL Datawarehouse** and built the final aggregate tables for dashboards
* Worked on Big Query projects for extracting the data required for web analytics.
* Worked extensively on the migration of different data products from Oracle to Azure.
* Worked on creating Functional Specification Document
* Created Requirement Document, Approach Paper, Design Specification Document
* Created Test Cases Document, Testing & Code Review Document, Delivery Release Notes, Deployments, etc.
* Generated ad-hoc reports in Excel Power Pivot and shared them using Power BI to the decision makers for strategic planning.
* Developed various solution driven views and dashboards by developing different chart types including **Pie Charts, Bar Charts, Tree Maps, Circle Views, Line Charts, Area Charts, Scatter Plots in Power BI**.
* Developed various dashboards to know the **school climate, school safety, risk factors, and risk behaviors.**
* Created dashboard for **school system financial information**which includes **revenue and expenditure by district.**
* Developed custom **Georgia district map visual,** **the SVG**is written inline within a Bootstrap 2 container.
* Written some Java script plugin that makes the polygons work like hyperlinks.
* Specialized in transforming data into user-friendly visualization to give business users a complete view of their business using **Power BI**.
* Used various sources to pull data into **Power BI**such as SQL Server, oracle.
* Involved in installation of **Power BI Report Server**.
* Have worked on building **Tableau** self-service data sources and dashboards which have critical KPIs for Business and Finance teams.
* Effectively built Alteryx workflows to compare the data sets from Oracle and Azure and have built the pipelines. Scheduled the workflows on the Alteryx server and published the output to Tableau server.
* Worked on list pulls and data encryption for sending targeted audience data to digital media platforms
* Writing and **Debugging** T-SQL, stored procedures, Views and User Defined Functions
* Optimized the performance of various SQL scripts, stored procedures and triggers by using **embedded UDFs**, **CTEs** and **System stored procedures**.
* Created SSIS packages for importing the Excel layouts into the database using SSIS.
* Used BIDS to design SSIS packages to transfer data between servers, load data into database, and archived data file from different DBMS. Scheduled the job to do this task periodically.
* Loaded data from ODS to Data Warehouse using SSIS packages.
* Experience in SSIS script tasks, Look up transformations and Data flow tasks using T- SQL, C# and Visual Basic scripts.
* Wrote **Teradata code** / **SQL** to enable the **ETL** processes to meet demanding business requirements, including **BTEQs, Fast Loads, Multi Loads,**and**Fast Exports**
* Deploying and scheduling Reports using **SSRS** to generate all **daily, weekly, monthly** and quarterly Reports including current status.

**JP Morgan Chase, Columbus, OH                                                 April 2016 – May 2018**

**Data Engineer**

**JPMorgan Chase & Co.** is a U.S. multinational banking and financial services holding company headquartered in New York City. It is a major provider of financial services.  JPMorgan Chase is one of the Big Four banks of the United States. The Chase brand is used for credit card services in the United States and Canada, the bank's retail banking activities in the United States, and commercial banking.

**Responsibilities:**

* Experience on Hadoop and **Spark** clusters to build the On-Demand Data warehouse and provide the Data to Data scientist
* Migrate SQL Server and Oracle database to **Microsoft Azure Cloud.**
* Worked Azure SQL Database Environment
* Experience working with Windows Hyper-V Server, Azure, Windows Clustering including
* Worked closely with Business analysts and architects to gather the requirements and implements the engineering work to create final data sets.
* Knowledge in creating real time data streaming solutions using **Apache Spark Core**, Spark SQL & Data Frames, Spark Streaming, Apache Storm, Kafka.
* Worked on critical change requests and quarterly EOC runs to provide the data to end users for data mining and business intelligence reporting.
* Created **Power BI reports**by using Joins in multiple tables from multiples database using complex SQL queries.
* Designed a **Power BI**data model with multiple fact tables and dimensions depending on the business requirements.
* Experience in custom visuals and groups creation in **Power BI**.
* Experience in **scheduling refresh of Power BI reports**, hourly and on-demand.
* Used **Tableau** Desktop Data Visualization tool for reports, integrated tableau with Alteryx for Data & Analytics.
* Involved in creating data blending in case of merging different sources, used Alteryx for Data blending for Tableau.
* Prepared the deployment plan and steps for each release by working along with DBAs and environment administrators. Release support and post deployment validations.
* Created External Tables to load data from the external files to Oracle source tables.
* Have re-rewritten the procedures to reduce the redundancy and improved the performance of the procedures by tuning the code. Using Explain Plan, Correct Joins, Indexing.
* Created global temporary tables to improve the query performance by eliminating the repetitive reference of the main tables in the procedures and run time improvements.
* Effectively partitioned the tables by creating suitable range and list partitions, helped the data reads and deletes to run faster by using partitions.
* Identify potential problems and recommend alternative technical solutions
* Participating in Technical Architecture Documents, Project Design and Implementation Discussions
* Azure Automation through Runbooks Creation, Migration of existing .PS1 scripts, Authorizing, Configuring, Scheduling
* Created Complex ETL Packages using SSIS to extract data from staging tables to partitioned tables with incremental load.
* Worked on SSIS Package, DTS Import/Export for transferring data from Database (Oracle and Text format data) to SQL Server.
* Automated the packages and procedures using Autosys and implemented the dependencies
* Worked on a critical project to purge and archive the deactivated loans from the OLTP application and archived the data into history database. Closely worked with DBAs and architect to implement the desired solution for the entire project including the POC to code deployment to production and support.
* Worked on deploying the latest changes in the data model to dev environments by taking the snapshot subversion code and pushing it to dev. Tested the deployments and have run the packages ahead of release into UAT and PROD environments.
* Scheduled Cube Processing from Staging Database Tables using SQL Server Agent.
* Generated reports using SQL Server Reporting Services 2005/2008 from OLTP and OLAP data sources.
* Extensively worked on Dimensional Modeling.
* Designed and deployed reports with Drill Down, Drill Through and Drop down menu option and Parameterized and Linked reports in BIDS. Deploying and scheduling Reports using SSRS/Crystal to generate all daily, weekly, monthly and quarterly Reports including current status.
* Expert in automating process by Using SQL Server Agent in creating jobs and error report using Alerts, SQL Mail Agent and SMTP.
* Handling user’s permissions and space issues on Production and Semi-Production Servers and handling maintenance Jobs, Backups and restores.
* Successfully tested all task in staging and stored all versions of data using team foundation server (TFS).