**PROFESSIONAL SUMMARY:**

* Having around 8 Years of professional experience in Database development and **Data Analysis** design of OLTP and OLAP systems.
* Experienced in Designing Star Schema (identification of facts, measures and dimensions), **Snowflake** schema for Data Warehouse, ODS architecture by using tools like Erwin Data Modeler, Power Designer, Embarcadero **E-R Studio** and **Microsoft** .
* Worked on different **AWS** services such as **EC2** instances, IAM roles and policies, **Dynamo** **DB** and **EMR**.
* Extensive experience in languages like **SQL**, **Python** and **PL/SQL** and extensive experience in Relational and Dimensional Data modeling for creating Logical and Physical Design of Database and ER Diagrams using multiple data modeling tools like **Erwin**, **ER** **Studio**.
* Experienced in working with **Teradata** Utilities like Fast load, Multi load, and Fast Export **Teradata** Query submitting and processing tools like **Teradata** **SQL** Assistant.
* Experienced in Automating and Scheduling the **Teradata** **SQL** **Scripts** in **UNIX** using Korn **Shell** scripting.
* Worked on Bigdata platforms such as **Hive** and **spark**. Moving data from one environment to other using **HDFS** and **Linux**. Worked with Data Engineers to migrate data from Source systems such as **Oracle**, **DB2** to Big data **Hadoop** using **Pyspark**.
* Expertise with various **RDBMS** and No-**SQL** databases like Oracle, **SQL** Server, DB2 UDB, and **Teradata**, Netezza and **MongoDB**, **Cassandra**.
* Skilled in visualizing, manipulating, and analyzing large datasets and has ability of designing and developing effective reports. Proficient with **MS** **Office** Applicants, **Excel** (Pivot Table, Lookups, Index, Match, dashboards, etc.).
* Experience in developing rich interactive Tableau reports and Dashboards using various visualizations like Chart-Box (Drill Down, Drill Up& Cyclic Grouping), Heat Maps, Tree Maps, Bubble Charts, Reference Lines, Dual Axes, Line diagrams, Bar Graphs, Geographic Visualizations (Detailed and filled in maps), etc.
* Skilled in **Power** **BI** and visualization of the results using Tableau to import, analyze and create various filters and parameters for reports from various data sources such as **SQL** Server.
* Proficient in **Data Analysis**, System Analysis, Business Requirement Gathering and Data Warehousing concept.
* Knowledge of one or more Data Visualization & Reporting tools such as **Tableau**, **Power** **BI**, **Cognos** and **SSRS** are essential.
* Led strategic planning initiatives to align technology efforts with organizational goals in the context of **MMIS.**
* Developed and communicated a technology roadmap for **Data Preparation**, **ETL**, and **EDI** integration within the **MMIS framework.**
* Extensive experience working in Data Warehouse Environment and using **SQL** for data analysis including joins and Analytical functions. High level of efficiency in **MS** **Office** Suite including **MS** **Word**, **Excel** and **Powerpoint**.
* Strong knowledge of Microsoft **Azure** cloud platform and its data analytics tools, such as **Azure** **Data** **Factory**, **Azure** Data Lake, and **Azure** **SQL** Database. Proficient in leveraging Azure services to extract, transform, and load data from various sources for analysis and reporting.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Programming Languages:** | SQL, Python, PL/SQL |
| **Data Modeling:** | Erwin Data Modeler, Power Designer, Embarcadero E-R Studio, Microsoft SQL Server, Oracle |
| **UNIX Shell Scripting:** | Automating and scheduling Teradata SQL scripts |
| **Big Data:** | Hive, Spark, Hadoop, HDFS, Linux |
| **Data Manipulation:** | Stored Procedures, Triggers, Views, User Defined Functions, Sub-Queries, Joins |
| **RDBMS and No-SQL Databases:** | Oracle, SQL Server, DB2 UDB, Teradata, Netezza, MongoDB, Cassandra |
| **MS Office Suite:** | Excel (Pivot Table, Lookups, Index, Match, dashboards) |
| **Data Visualization:** | Tableau ,Power BI |
| **Cloud:** | Aws, Azure |
| **ETL Tools** | Informatica, DataStage IBM, SSIS, DTS |



AWS Data Analytics Fundamentals AWS Big Data Fundamentals

GE Digital Technology Enterprise Program Finance for Non-Financial Professionals Google Analytics for Beginners

Tableau Analyst

Google Project Management Certiﬁcation



The University of Texas at Dallas

Cummins College of Engineering, Pune University, India

**PROFESSIONAL EXPERIENCE**

**Client: AT&T, Dallas, TX Mar,2023 to Present**

**Role: Sr. Data Analyst**

**Responsibilities:**

* Developed and executed **Python** scripts for **ETL** processes, ensuring efficient data extraction, transformation, and loading. Developed and optimized complex **SQL** queries for efficient data retrieval and analysis.
* Implemented interactive and dynamic data visualizations using **D3**.**js** for effective representation and exploration of complex datasets.
* Designed and implemented **NoSQL** database solutions, optimizing data storage and retrieval for scalability and flexibility.
* Implemented and managed **Azure** services, including **Azure** **Blob** **Storage**, **Azure** **SQL** **Database**, **Azure** **Data** **Lake**, and **SQL** **Server** **DB**.
* Developed and maintained **ETL** packages in **Azure** Data Factory for data transformation and ingestion.
* Utilized **Azure** Key Vault for secure storage and management of sensitive information, such as authentication keys.
* Utilized **Jira** for agile project management, facilitating collaboration, issue tracking, and task prioritization in data analysis projects.
* Created interactive and insightful **Tableau** dashboards, supporting visualizations for complex datasets to aid decision-making.
* Implemented complex **Excel** formulas and functions, utilizing **Power** **Query** and **Power** **Pivot** for data cleaning and transformation.
* Collaborated with stakeholders to understand business requirements and translated them into effective **QlikView** solutions.
* Developed and maintained **SAS** programs for **ETL** processes and conducted statistical analysis to derive actionable insights. Designed and implemented **NoSQL** database solutions for optimal storage and retrieval of data.
* Utilized **D3.js** for interactive and dynamic data visualizations, effectively representing and exploring complex datasets. Implemented Data Exploration to analyze patterns and to select features using **Python** SciPy.
* Designed and created **Excel** dashboards for tracking key performance indicators (**KPIs**) and business metrics.
* Managed agile project management using **Jira**, ensuring collaboration, issue tracking, and task prioritization in data analysis projects.
* Developed and maintained **Power** **BI** reports and dashboards, optimizing data models for efficient storage and retrieval. Developed and designed **QlikView** applications for interactive and intuitive data visualization.
* Engineered **Hadoop** clusters, developing and executing **MapReduce** programs for data processing and analysis.
* Utilized **Apache** **Hive** for data warehousing and **SQL**-like queries on **Hadoop** Distributed File System (**HDFS**).
* Conducted data quality rules and transformations in **Power** **BI** to ensure high standards of data accuracy.
* Developed and executed data analysis workflows using **Jupyter** **Notebook**, integrating code, visualizations, and narrative documentation.
* Extracted, analyzed, and interpreted web data using **Google** **Analytics**, providing insights into user behavior and website performance.
* Conducted advanced statistical analysis and predictive modeling using **IBM** **SPSS** to derive actionable insights from structured data.
* Utilized **Azure** Data Factory for efficient **ETL** processes, ensuring data integrity and consistency for analysis.
* Collaborated with **SMEs** and stakeholders to determine requirements, identify entities, and build conceptual, logical, and physical data models for industry-specific offerings.

**Environment**: ER-Studio, Python, Azure, Teradata, Netezza, Oracle 12c, Cognos, MS Office (Word, Excel and PowerPoint), SQL Server, MS Project, MS FrontPage, MS Access, EDI, UML, MS Visio, Oracle Designer, SQL Server, Oracle SQL developer, Crystal Reports, **SSRS**, SSIS and Tableau.

**Capital One, Dallas, TX Dec, 2021 to Mar, 2023**

**Role: Sr. Data Analyst**

**Responsibilities:**

Created comprehensive risk assessments and workﬂow documentation for the EPIC Beaker system, driving process improvement recommendations.

Evaluated the effectiveness of current controls by analyzing and assessing risk, testing IT General Controls (ITGCs) for Change Management and Logical Access, and proposing enhancements to policies, procedures, and controls efﬁciency.

Thoroughly examined internal controls, focusing on areas like segregation of duties, authorization procedures, and access controls.

Collaborated in the development of audit reports, effectively summarizing audit discoveries and delivering actionable recommendations to the management team.

Developed and maintained ETL processes, creating mappings for seamless data ﬂow using SSIS.

Utilized Python packages like NumPy, matplotlib, and SciPy to generate graphical capacity planning reports. Designed ETL workﬂows in Informatica Power Center based on provided mapping sheets.

Optimized data sources for route distribution analytics dashboard runtime in PowerBI reports. Created UML use-case diagrams, Class diagrams, and diagrams using MS Visio.

Applied Python for predictive analytics, addressing various use-cases. Leveraged PowerBI to publish and share reports with business stakeholders.

Crafted action ﬁlters, parameters, and calculated sets for PowerBI dashboards and worksheets.

Managed Departmental Reporting systems, troubleshooting daily issues, and integrating Access databases with external data sources.

Designed and implemented multiple PowerBI dashboards, integrating custom visualization tools.

Orchestrated end-to-end systems for Data Analytics and Automation, integrating custom visualization tools through PowerBI.

Employed PowerBI and custom SQL features to create insightful dashboards and identify correlations. Developed R and Shell scripts for administrative task automation.

Enhanced performance of Informatica mappings through performance tuning efforts. Composed Teradata SQL Queries using SQL Assistant for Ad Hoc Data Pull requests.

Extracted data from various sources including Oracle tables, MS SQL Server, sequential ﬁles, and Excel sheets. Created Data Quality Scripts using SQL to validate data load success and data quality.

Conducted data analysis and proﬁling using complex SQL on diverse source systems. Categorized and generated reports on multiple parameters using MS Excel and PowerBI. Designed SQL scripts to identify data quality issues, keys, anomalies, and validation problems.

Contributed to logical and physical modeling of data marts and DW/BI architecture using Teradata.

 : Matplotlib, PowerBI, T-SQL, MS SQL Server, MS Excel, UML, MS Visio, ETL - SSIS, SSRS, Data Modelling - Star Schema, Snowﬂake Schema, Shell Script, Jira, Git, Teradata, Workﬂows

**Apilation.ai, Dallas, TX Sep, 2020 to Dec, 2021**

**Role: Sr. Data Analyst**

**Responsibilities:**

* Managed **AWS** services, including **EC2**, **SQS**, **Kenesis**, and **S3** bucket, configuring security groups for accessibility from on-premise environments.
* Administered **IAM** security policies, roles, and keypairs for Windows and Linux instances, implementing **NAT** gateways and **NAT** instances.
* Designed, built, and operationalized large-scale enterprise data solutions using **AWS** data and analytics services, including **Spark,** **EMR**, **DynamoDB**, **RedShift**, **Kinesis**, **Lambda**, **Glue,** and **Snowflake**.
* Utilized Python alongside **SQL** for seamless interaction with relational databases, ensuring optimal data retrieval.
* Employed **Rational** **Rose** for creating visual models and diagrams, enhancing understanding of complex data architectures and relationships. Implemented and managed **AWS** **EC2** instances, **S3** buckets, **SQS**, and **Kenesis** for efficient data storage and processing.
* Developed and executed complex **SQL** queries for data extraction, transformation, and analysis, providing actionable insights to support business decision-making.
* Executed data quality assurance processes using **Collibra**, defining and monitoring data quality metrics, and implementing corrective measures to ensure high-quality, reliable data.
* Developed and operationalized large-scale data solutions using **AWS** services such as **EMR**, **RedShift**, **DynamoDB,** and **Lambda**. Employed **Rational** **Rose** for creating visual models and diagrams, enhancing the understanding of complex data architectures and relationships.
* Developed and maintained entity-relationship **(ER**) **diagrams** to illustrate database structures and guide data modeling processes.
* Developed and designed **Tableau** dashboards and visualizations to convey actionable insights from complex datasets.
* Conducted **QlikView** training sessions for end-users to enhance data literacy and promote self-service analytics.
* Designed and executed **QlikView** scripting language for data cleansing, transformation, and aggregation.
* Designed and implemented interactive dashboards in **Sisense**, enabling intuitive exploration and visualization of complex datasets.
* Applied machine learning techniques in **Jupyter** **Notebook** for predictive modeling and pattern recognition.
* Conducted data validation and cleansing procedures to ensure accuracy and reliability of **Excel** datasets.
* Implemented data ingestion pipelines using **Apache** **Flume** to efficiently collect and transfer data into **Hadoop**.
* Managed and optimized **Hadoop** ecosystem components, including **HBase** for **NoSQL** storage and **Apache** **Pig** for data processing.
* Collaborated with cross-functional teams to understand business requirements and deliver effective **SAS** solutions. Designed and implemented scalable **Power** **BI** architectures to accommodate growing data volumes and evolving business requirements.
* Engineered data preprocessing pipelines using **Python** for machine learning models, including feature engineering and scaling.
* Implemented data blending in **Tableau** for merging disparate datasets and creating comprehensive visualizations.
* Analyzed, re-architected, and re-platformed on-premises **SQL** data warehouses to data platforms on the **AWS** cloud, utilizing **AWS** or third-party services.

**Environment**: Erwin 9.1, ORACLE, AWS, Teradata, SQL, T-SQL, PL/SQL, XML, Netezza Aginity, Teradata SQL Assistant, SQL Plus, Informatica, SSRS, SSIS, UML, Rational Rose, Power BI, Tableau Excel, MS Visio, Source Offsite (SOS), Windows 7, QC Explorer, Business Objects, Share point workspace 2010.

**Client: Verisk Analytics, Jersey City, NJ Dec, 2018 to Aug, 2020**

**Role: Data Analyst**

**Responsibilities:**

* Provided support to developers in mapping source attributes into **Teradata** Financial Services Logical Data Model (**FSLDM**) and interpreted business requirements for Data Marts.
* Led design discussions to determine the appropriate **Data** **Warehouse** at the lowest grain for each Dimension, designing **STAR** **Schema** for sales data using **Erwin** Data Modeler.
* Communicated and executed strategic plans for **data preparation, ETL, and EDI integration**.
* Developed a comprehensive technology roadmap aligning **MMIS** efforts with organizational goals.
* Conducted **A/B** **testing** and performance analysis using **Google** **Analytics** to optimize website functionality and user experience.
* Mentored **cross-functional teams, fostering collaboration** in MMIS data preparation projects.
* Provided clear guidance on **ETL** best practices and **EDI** standards for efficient data exchange.
* Implemented robust data governance policies, ensuring **MMIS data quality and integrity and** established metrics and monitoring mechanisms for proactive data quality management.
* Implemented performance optimization strategies, reducing **ETL** processing times.
* Introduced automation techniques for streamlined **data extraction, transformation, and loading**.
* Ensured **MMIS** compliance with **EDI standards**, facilitating seamless external communication.
* Successfully navigated **EDI implementation** challenges, resolving format and protocol issues.
* Conducted training sessions, empowering **MMIS** team members with the latest tools and techniques and established a knowledge-sharing culture, promoting continuous learning within the team.
* Implemented effective solutions for challenges in **ETL workflows** and **EDI interactions**.
* Conducted data profiling and analysis using **NoSQL** databases, ensuring data quality and consistency in analytics projects. Configured and optimized **Azure** **Data** Factory pipelines for efficient data processing and transfer.
* Conducted performance tuning in **Azure** **SQL** Database configurations to enhance query efficiency.
* Implemented data security measures using **Azure** Identity and Access Management (IAM**)** for access controls.
* Utilized **Jira** for tracking and managing data-related tasks, user stories, and project milestones throughout the development lifecycle. Utilized **SAS** Enterprise Guide for interactive data exploration and reporting.
* Created and maintained technical documentation for **QlikView** applications, including data models, scripts, and dashboard layouts. Designed and maintained **Excel** dashboards for tracking **KPIs** and business metrics.
* Integrated **QlikView** with various data sources, ensuring real-time data updates and accurate analysis.
* Created and maintained **SQL** scripts for data extraction, transformation, and loading (**ETL**) processes.
* Implemented data visualization solutions using **Python** libraries for effective communication of insights.
* Leveraged Apache **Spark** for in-memory data processing, **machine** **learning**, and real-time analytics on **Hadoop** clusters. Ensured data security in **Hadoop** environments through authentication and authorization measures.
* Implemented data validation checks within **Tableau** dashboards to ensure accuracy and quality.
* Conducted data profiling and analysis using **NoSQL** databases for consistency in analytics projects.
* Utilized **Jira** for tracking and managing data-related tasks, user stories, and project milestones.
* Conducted advanced statistical analysis using **Excel** features, including regression, correlation, and hypothesis testing. Designed and maintained **Tableau** data models for efficient storage and retrieval of information.
* Conducted data profiling and lineage analysis in **Power** **BI**, creating technical documentation for launching **Power** **BI** solutions and executing data-related tasks.

**Environment**: **SQL** Server, SSIS, SSRS, ETL, EDL, Windows 2000, AZURE, Micro strategy 9.2, Crystal Enterprise8.0, Erwin 9.1, CSV files, XML files, Business Objects, Oracle10g, Teradata, Netezza, SAS, SSIS, Informatica, XML.

**Client: Titan Group, Pune, India May, 2017 to Aug, 2018**

**Role: Data Analyst**

**Responsibilities:**

* Provided technical guidance for migrating **Teradata** warehouse operations to **Netezza**, optimizing query performance using **Teradata** Index techniques.
* Documented database structures, schema designs, and **SQL** queries to support transparent and replicable analyses.
* Integrated **Power** **BI** with diverse data sources, ensuring connectivity and real-time data updates, and implemented monitoring using **CloudWatch** for **AWS**-based deployments.
* Developed custom data preprocessing scripts in **Jupyter** **Notebook** and created data reports and visualizations using Sisense for informed decision-making.
* Conducted Excel training for team members and supported **Python**-based **API** development for seamless data interaction between applications. Applied exploratory data analysis (**EDA**) techniques in **Python**, revealing patterns and trends for data-driven decision-making.
* Applied **Rational** **Rose** for visualizing and documenting system architectures, fostering better communication among project stakeholders.
* Created and maintained comprehensive **ER** **diagrams** to depict database schema, relationships, and constraints, ensuring accurate data representation.
* Integrated **QlikView** with various data sources, ensuring real-time updates, and maintained technical documentation for **QlikView** applications.
* Visualized system architectures using **Rational** **Rose**, fostering effective communication among project stakeholders. Designed and maintained **Tableau** data models for efficient storage, retrieval, and analysis of information.
* Conducted performance tuning in **SQL** databases, improving query response times and overall system efficiency.
* Configured and utilized **AWS** CloudWatch for monitoring, logging, and alarming of **AWS**-based data solutions.
* Implemented data security measures using **AWS** Key Management Service (**KMS**) for cryptographic key management.
* Created and maintained **ER** **diagrams**, ensuring accurate data representation, and provided **Excel** training and support.
* Automated **GLUE** **ETL** jobs using **S3** events and **AWS** **Lambda** functions, developing Star and **Snowflake** schemas for Logical and Physical Data Models.

**Environment**: ER-Studio, Oracle 10g, SQL Server, AWS, Crystal Reports, MS Office, Teradata, Informatica, SSRS, SAS, SSIS, XML, Excel, Pivot Tables, Metadata, Cognos.