**Spurthi Shetty**

**Data Analyst ||spurthishettyy26@gmail.com || 857-334-1054**

**PROFESSIONAL SUMMARY**

* Over 6+ years of strong experience in Business and Data Modeling/ Data Analysis, Data Architect, Data Profiling, Data Migration, Data Conversion, Data Quality, Data Governance, Data Integration, MDM, NoSQL and Metadata Management Services and Configuration Management.
* Expertise and Vast knowledge of Enterprise Data Warehousing including **Data Modeling, Data Architecture, Data Integration** (ETL/ELT) and **Business Intelligence**.
* Skilled in implementing SQL tuning techniques such as Join Indexes(JI), Aggregate Join Indexes (AJI's), Statistics and Table changes including Index.
* Proficient in utilizing analytical applications like **Python** and **R** to identify trends and relationships between different data points to draw appropriate conclusions and translate analytical findings into risk management and marketing strategies that drive value.
* Experienced in using various Teradata Utilities like **Teradata Parallel Transporter** (TPT), **Mload, BTEQ, FastExport,**and**Fast load.**
* Experience in business intelligence and data analysis, Reporting, Data Preparation, Data Validation, Data warehousing, ETL techniques, Visualizations, software and predictive model design and development using Tableau, RDBMS tools, IBM Cognos, Power BI, Excel, python.
* Experience in Documentation tools like **MS Visio, MS Project**, and **MS Office** to create reports required for the projects and client submissions.
* Extensive knowledge in **Machine learning, Data mining, SQL Queries,** and Databases.
* Proficient in writing queries and subqueries to retrieve the data using **SQL** from various servers including **Microsoft SQL, Oracle,** and **MySQL**.
* Worked on generating ad hoc reports and analyzing the data for fixing the errors and expert taking initiatives in taking Business decisions.
* Extensive experience in the development and designing of **ETL** methodology for supporting data transformations and processing in a corporate-wide environment using Teradata, Mainframes, and UNIX Shell Scripting.
* Experienced in Dimensional Data Modeling experience using Data modeling, Relational Data modeling, ER/Studio, Erwin, and Sybase Power Designer, **Star Join Schema/Snowflake modeling, FACT & Dimensions tables, Conceptual, Physical & logical data modeling.**
* Experienced in handling all the domain and technical interaction with application users, analyzing client business processes, documenting business requirements.
* Experienced in writing Design Documents, System Administration Documents, Test Plans & Test **Scenarios/Test Cases** and documentation of test results.
* Extensive experience in the development of **T-SQL, OLAP, PL/SQL, Stored Procedures, Triggers, Functions, Packages**, performance tuning and optimization for business logic implementation.
* Responsible for architecture design, data modeling, and implementation of Big Data platform and analytic applications.
* Experience developing algorithms to create Artificial Neural networks to implement AI solutions to optimize business processes and minimize costs.
* Experienced in Data Scrubbing/Cleansing, Data Quality, Data Mapping, Data Profiling, Data Validation in ETL.
* Extensive experience in business intelligence (and BI technologies) tools such as OLAP, Data warehousing, reporting and querying tools, Data mining and Spreadsheets.
* Experienced in Creating and Maintaining documentation such as **Data Architecture/Technical/ETL** Specifications.

**EDUCATION:**

 Bachelor of Computer Science.

**TECHNICAL SKILLS:**

|  |  |
| --- | --- |
| **Programming Languages** | Python, SQL, R. |
| **ETL and Big Data Technologies** | SSIS, Talend, Alteryx, Hadoop, Sqoop, Hive, Apache Spark |
| **DataModeling** | Sybase Power Designer / IBM Data Architect  |
| **Web Analytics** | Google Analytics, Adobe Analytics |
| **MS-Office Package** | Microsoft Office (Windows, Word, Excel, PowerPoint, Visio, Project). |
| **Data Analytics and BI Tools** | Tableau, Power BI, Looker, QlikView, MS Excel (Pivot Tables, VLOOKUP) |
| **ETL Tools / Tracking tool** | Informatica, SSIS, SSAS, SSRS / JIRA. |
| **Database technologies** | MySQL, MS SQL Server, Oracle, PostgreSQL, MongoDB |
| **Testing and defect tracking Tools** | HP/Mercury (Quality Center, Win Runner, Quick Test Professional, Performance Center, Requisite, MS Visio & Visual Source Safe  |

**Professional Experience:**

**Client: BBVA Bank Birmingham Jul 2020 - Present**

**Role: Data Analyst.**

**Description:** BBVA Bank is leading the transformation of banking worldwide, united in pursuing our goal of bringing the age of opportunity to everyone. Firmly focused on the future, our on-going digital transformation is already producing disruptive innovations that power our vision of banking.

**Responsibilities:**

* Constructed advanced SQL scripts to extract data from **Oracle SQL server** which saw 23% increase in query performance in **AWS Workspace**
* Perform data discovery, profiling and data quality using Oracle SQL server and various other tools, like ETL Tool **Informatica** for running SQL queries in Agile methodology using JIRA stories. Resolved data integrity issues and developed data mapping templates which resulted in improvement in data consistency
* Translated the business needs into analytic and reporting strategies through the development, execution of data, implemented **Python**, scripts to web scrapping and data wrangling processes for structured as well as unstructured data
* Automated strategic business planning process by deploying centralized Reporting/ Data Catalogue from source to target data
* Optimized query performance and data quality testing to ensure actionability and reliability of data.
* Regularly monitor the counts of the files in **DATIO** and requesting the Re-runs for any mismatches from source system
* Updated/Managed data and queries using **PLSQL** and retrieve data related to equipment’s performance and quality reports.
* Generated ad-hoc **SQLqueries** using joins, database connections,and transformation rules to fetch data from legacy**DB2** and **SQL Server** database systems.
* Developing visualization in **Tableau** for better understanding of the customer experience
* Developed **Data Mapping, Data Governance, Transformation** and Cleansing rules for the Master Data Management Architecture involving **OLTP** and **OLAP**.
* Performed data analysis and data profiling using complex **SQL** on various sources systems and answered complex business questions by providing data to business users.

**Environment:**Oracle, SQL, Oracle SQL Server, AWS WorkSpace, SQL, SSAS, XML, SDLC, PL/SQL, Informatica IDQ

**Client:Prolifics,Nashville.** **Apr 2018-Jul 2020**

**Role: Data Analyst.**

**Description:**Our experience across multiple industries means we understand – and will help you meet your customers’ demands and expectations. We use our knowledge of Cloud, DevOps, Data, and Quality to give you fast, complete solution delivery experiences that you will find nowhere else.

**Responsibilities:**

* Created Logical & Physical Data Modeling on Relational (**OLTP**), Dimensional Data Modeling (**OLAP**) on Star schema for Fact & Dimension tables using **Erwin**.
* Assisted to create an algorithm to detect anomaly/outlier detection on abnormal events for the month to month sales of medical devices to distributors.
* Generated Monthly Global Metric/ KPI reports for the Global Service Performance to US, CA, Europe and Asian Regions supporting the **PLM** project.
* Performed **Data collection, Data cleaning and Data visualization** using RStudio, and feature engineering using python libraries such as pandas and NumPy, performed Deep feature synthesis and extracted key statistical findings to develop business strategies.
* Worked on weekly **GSP BI** reports and supported Ad-hoc reports with the request by Users/Engineers using Infor **Nexus** and **Cognos** 11.
* Worked parallelly on different regions using Oracle PLSQL and gained knowledge on various domains in databases
* Automated tasks for Ad-Hoc, metric reports, and excel spreadsheets to reduce the workload, quality and to enhance time redundancy
* Worked extensively on trials to analyze the sales data production for predictive models using Machine Learning techniques like **KNN** and distance algorithms.
* Automated Access with **Excel** to extract information from various Excel files and imported them into MS Access.
* Held Walkthroughs with **QA, BA, SMEs,** and Stakeholders for a better understanding of Data Analysis with Reports.
* Gathered business requirements, working closely with business users, project leaders,and developers. Analyzed the business requirements and designed Conceptual and Logical Data Models.
* Prepared **ETL** technical Mapping Documents along with test cases for each Mapping for future developments to maintain **SDLC**.
* Involved in Data flow analysis, Data modeling, Physical database design, forms design and development, data conversion, performance analysis, and tuning.
* Implemented Normalization Techniques and build the tables as per the requirements are given by the business users.
* Experienced in advanced **Tableau** features including calculated fields, parameters, table calculations, row-level security, **joins**, data blending, and dashboard actions.
* Performed data collection, data cleaning in a huge dataset that had many missing data & extreme outliers from **Hadoopworkbooks** and explored data to draw relationships and correlations between variables.
* Built model and algorithm templates on python using and python package for deployment on the entire data set using **HDFS** and **MapReduce**.
* Extensively worked on early-stage business projects, discovery efforts, and engagements initiated by Business Relationship Managers to provide appropriate architecture deliverables, such as stakeholder analyses, **capability analyses, risk/value analyses**, or technical analyses.
* Worked with Database Administrators, Business Analysts,and Content Developers to conduct design reviews and validate the developed models.

**Environment:**Oracle, SQL, Teradata,IBM Cognos,SQL, SSAS, MDM, XML, SDLC, PL/SQL, Informatica IDQ,ErwinRalph Kimball.

**Client: Caterpillar,Peoria,IL. Jun 2017-March 2018**

**Role: Data Analyst.**

**Description:**Caterpillar Inc. is an American Fortune 100 corporation that designs, develops, engineers, manufactures, markets and sells machinery, engines, financial products and insurance to customers via a worldwide dealer network.

**Responsibilities:**

* Developed full life cycle software including defining requirements, prototyping, designing, coding, testing and maintaining software.
* Extensively used the **Agile** methodology as the Organization Standard to implement the Data Models.
* Developed an **OLTP** system by designing Logical and eventually **Physical Data Model** from the Conceptual Data Model.
* Strong understanding of the principles of Data warehousing using Fact Tables, Dimension Tables, star schema modeling and snowflake schema modeling, Slowly changing dimensions, foreign key concepts, referential integrity.
* Extracted data from various sources (SQL Server, Oracle, text files and excel sheets), used ETL load scripts to manipulate, concatenate and clean source data.
* Involved in the data transfer creating tables from various tables, coding using PL/SQL, Stored Procedures and Packages.
* Performed data analysis, statistical analysis, generated reports, listings, and graphs using **SAS** tools, SAS Integration Studio, **SAS/Graph, SAS/SQL, SAS/Connect** and **SAS/Access.**
* Responsible for checking the daily inbound, outbound data activities and update appropriate **EMR** data in the back end of the production server.
* Supported and worked closely with the Data Engineering team using **Azure Data Factory** creating pipelines and following **ETL** processes during migration to the cloud
* Extensively worked on **MS SQL server** to retrieve, validate and generate the clinical data using **SQL** queries.
* Worked with senior programmers and **QA** personnel in performing unit and user acceptance tests to ensure best performance and compatibility of **MS Access database** or applications on various platforms as well as providing regular system updates as required
* Maintained Referential Integrity by Introducing foreign keys and normalized the existing data structure to work with the **ETL** team and provided a source to target mapping to implement incremental, full and initial loads into the target data mart.
* Created conceptual, logical and physical data models, data dictionaries, DDL and DML to deploy and load database table structures in support of system requirements.
* Ability to document activities and communicate with Data Architects as to the status of Physical Data Modeling activities.
* Implemented Forward Engineering by using **DDL** scripts and generating indexing strategies to develop the logical data model using **Erwin**.
* Implemented **ETL** techniques for **Data Conversion**, **Data Extraction** and **Data Mapping** for different processes as well as applications.
* Maintained **Data Consistency** by evaluating and updating logical and physical data models to support new and existing projects.
* Developed and maintained the data dictionaries, **Naming Conventions, Standards**, and Class Words Standards Document.
* Handled importing of data from various data sources, performed transformations using **Hive**. (External tables, partitioning).

**Environment:**OLTP, DBAs, ETL, DDL, DML, Erwin 9.6, UML, diagrams, Snow-flak schema, SQL, Data Mapping, Metadata, OLTP, SAS, Informatica 9.5.

**Client: Staples Inc,Framingham,MA. Oct 2016-May 2017**

**Role: Data Analyst.**

**Description:** Staples Inc. is an American office retail company. It is primarily involved in the sale of office supplies and related products, via retail channels and business-to-business-oriented delivery operations.

**Responsibilities:**

* Developed **data access** related to sorting, retrieving or acting on housed data.
* Predominately responsible for creating new, modifying existing, and analyzing data within our reporting system.
* Involved in gathering data from a different team, along with business intelligence team to provide reports, coordinating the deliverables, gathering and documenting the requirements.
* Used **ETL** to **develop** jobs for extracting, cleansing, transforming and loading the data from various sources.
* Experienced **Python** in data wrangling, cleansing, preparation for analysis.
* Performed in-depth analysis of data and prepared daily reports by using **SQL, MS Excel, Share Point**.
* Managed to scheduled data refresh on the **Tableau**Server for weekly and monthly increments based on business changes, which updated on the **dashboard**.
* Experienced in translating requirements into actionable reports and providing consulting support to clients that is data-based, analysis-driven, and a strong understanding of customer relationship management.
* Performed various ad-hoc analysis by extracting data from multiple source systems and creating comprehensive reports for end-users.
* Define data needs, evaluate data quality, and extract/transform data for analytic projects and research.
* Analyzing data and prototype models for targeting and personalization, work on analytical or experimental requirements to devise data solutions.
* Effectively communicate and document technical analyses and results.
* Involved closely with the **Marketing** and Operations team to understand/define requirements, domain knowledge/models, and data needs.
* Ensure analysis and solutions drive business decisions.
* Developed a solution that will aid in the data capture, data cleansing, data monitoring and reporting of customer data.
* Helped define key business problems to be solved; analyze data to solve those problems.
* Generated Heat maps to identify the risk and flaws in the business.
* Validated Data to check for the proper conversion of the data. Data cleansing to identify unnecessary data and clean, data profiling for accuracy, completeness, consistency.
* Assisted and produced standard reports, charts, graphs, and tables from a structured data source by querying data repositories using **Python** and **SQL**.
* Developed and produced a dashboard, key performance indicators and monitor organization performance.

**Environment:** Python, SQL, Jupyter, NumPy, Tableau, MS Office Suite, Visio, PowerShell, Windows XP,7,10..**.**

**Client:Sutter Health,Sacramento,CA. Aug 15 - Sep 2016**

**Role: Data Services Analyst.**

**Description:**Sutter Health is a not-for-profit health system in Northern California, headquartered in Sacramento. It includes doctors, hospitals and other health care services in more than 100 Northern California cities and towns.

**Responsibilities:**

* Collecting and evaluating business requirements, functional specifications, project schedules, documentation,Performed ETL data integration using Alteryx and developed Tableau applications to analyze **KPI’s** and trends to improve the performance of the firm
* Perform data discovery, profiling and data quality using **SQLServer** (**SSMS**) and various other tools, like Alteryx **ETL Tool** for running **SQL** jobs.
* Built and developed data analytical databases from complex underwriting, claims, finance, and external source data.
* Automated in migrating reports from **Alteryx** to **Tableau** server and successfully archived reports using **Snowflake** cloud platform
* Develop data warehouse and business intelligence architecture, technical standards, centralized technical metadata, data quality scorecards, dashboards, and ad-hoc reporting. Maintained and documented **ETL** mappings from various sources for applications deployed in production
* Gained strong working experience using Waterfall and Agile software development and project management methodologies.
* Used **ETL** to develop functions for extracting, transforming, cleaning, loading data from various sources to a different destination.
* Used **Python**, **Tableau,** and **Excel** to analyze the number of products per customer and sales in a category for sales optimization.
* Visualized data by using advance tableau functions like action filters, context filters, and Level-of-Detail (**LOD**) expressions.

**Client:Radiant Technologies,Mumbai. June 2013 - March 2015**

**Role: Business Data Analyst.**

**Description:** Radiant Technologies Vision is to become the preferred choice of electrical utilities across the country and thus make a significant contribution to the industrial progress of the country.

**Responsibilities:**

* Worked with Analytics team to evaluate & developed organizational change business models with the requirements, also gathered, analyzed and documented business and functional requirements and created UML diagrams for systems analysis
* Acted as a liaison between business leaders, technical managers and internal stakeholders to translate use cases into requirements and developed implementation strategies for new technology and by prototyping tools like Azure RP, Balsamiq
* Translated the business needs into analytic and reporting strategies through the development, execution of data, implemented Python, R to web scraping and data wrangling processes for structured as well as unstructured data
* Migrated data from OLTP to OLAP by using big data technology Sqoop for ETL process and HIVE to fetch data and performed data visualization to develop BI reports consisting of interactive dashboards, charts to perform trend analysis, using Tableau
* Deployed SQL queries cloud platform Microsoft Azure & AWS Platform for overall better performance & efficiency.