**PAVAN SHEELAM**

 pavansheelam27@gmail.com || 732-353-2353

**OBJECTIVE:**

Motivated and curious data analyst with around 8 years of experience. Proficient in Python, SQL, R, Advanced Excel and various visualization tools, I bring a passion for creative innovation and a knack for solving complex problems. Eager to apply my extensive experience and to further develop my data analysis skills within a team. Excited about contributing to dynamic projects and embracing opportunities for continuous learning and professional growth.

**Summary**

* A Data Warehouse developer, Data modeler, Data Analyst with more than 6 years of experience in all phases of Software development life cycle (SDLC) including System Analysis, Design, Data Modelling, Implementation and Support and maintenance of various applications in both OLTP and OLAP systems.
* Possess strong Documentation skill and knowledge sharing among Team, conducted data modeling sessions for different user groups, facilitated common data models between different applications, participated in requirement sessions to identify logical entities
* Extensive Experience working with business users as well as senior management.
* 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.
* Strong experience in Database such as Microsoft Access, Oracle 11g/10g/9i/8i, SQL Server, DB2, Teradata, on Windows 2000/NT/98/95 platforms and latest.
* Expert in writing SQL queries and optimizing the queries in Oracle, SQL Server 2008 and Teradata.
* Excellent Software Development Life Cycle (SDLC) with good working knowledge of testing methodologies, disciplines, tasks, resources and scheduling
* Excellent knowledge in Data Analysis, Data Validation, Data Cleansing, Data Verification and identifying data mismatch.
* Performed data analysis and data profiling using complex SQL on various sources systems including Oracle and Teradata.
* Excellent experience on Teradata SQL queries, Teradata Indexes, Utilities such as Mload, Tpump, Fast load and Fast Export.
* Strong experience in using Excel and MS Access to dump the data and analyse based on business needs.
* Excellent knowledge on Perl & Unix.
* Experienced working with Excel Pivot and VBA macros for various business scenarios.
* Strong experience in Data Analysis, Data Migration, Data Cleansing, Transformation, Integration, Data Import, and Data Export through the use of multiple ETL tools such as Ab Initio and Informatica PowerCenter Experience in testing and writing SQL and PL/SQL statements - Stored Procedures, Functions, Triggers and packages.
* Experience in automating and scheduling the Informatica jobs using UNIX shell scripting configuring Korn-jobs for Informatica sessions.
* Excellent experience in analysis huge volumes of data in industries such as Finance, Healthcare & Retail.
* Excellent knowledge on creating reports on SAP Business Objects, Webi reports for multiple data providers.
* Excellent knowledge in preparing required project documentation and tracking and reporting regularly on the status of projects to all project stakeholders
* Experienced in writing Toad and PL/SQL Stored Procedures, Triggers and Functions.

**TECHNICAL SKILLS**

Data Warehousing Informatica 9.1/8.6/7.1.2 (Repository Manager, Designer, Workflow Manager, and Workflow Monitor), SSIS, Data Stage 8.x
Reporting Tools Business Objects6.5, XIR3, Cognos 8 Suite
Data Modeling Star-Schema Modeling, Snowflake-Schema Modeling, FACT and dimension tables, Pivot Tables, Erwin
Testing Tools Win Runner, Load Runner, Test Director, Mercury Quality Center, Rational Clear Quest
RDBMS Oracle 10g/9i/8i/7.x, MS SQL Server, UDB DB2 9.x, Teradata, MS Access 7.0
Programming SQL, PL/SQL, UNIX Shell Scripting, VB Script
Environment Windows (95, 98, 2000, NT, XP), UNIX
Other Tools TOAD, MS-Office suite (Word, Excel, Project and Outlook), BTEQ, Teradata SQL Assistant

**SKILLS**

**Python and Libraries, R programming, SAP APM, Salesforce, SQL, PostgreSQL, Tableau, Power BI, Looker Studio, Alteryx, Advanced Excel, ETL, Big Data Ecosystem, NLP, Azure, AWS, GCP, MS Office, Power pivot, GIT, Jupyter Notebook, SSIS, SSRS, SSAS, Snowflake, SAS, JIRA, Service now.**

**PROFESSIONAL EXPERIENCE**

**Arkansas Blue Cross Blue Shield Sr. Data Analyst May 2024 – Present**

* Maintain, manage and continuously improve the systems and processes associated with the contracting, payments and reporting of compensation for internal and external agents and brokers.
* Responsible for quantitative quarterly and year-end group service fee compensation audit analysis of $4.6 million in transactional data across multiple entities using **SAP APM**.
* Developed **SQL** queries to model financial analysis-based recommendations capturing over $500,000 in retro commissions.
* Worked extensively on **Kythera data** (voluminous **raw medical claims data**) by building several schemas to break it into various dimensions and facts.
* Point of contact for various Payer data feeds like **MMIT, DRG & CMS Medicare** and Claims datasets like **SHA** & **IQVIA LAAD**. Implemented various business solutions using these datasets right from summarizing the raw data.
* **Incentive Compensation (IC**) implementation using Payer data for the **Salesforce** team
* Designed relational databases for in-depth analysis, guaranteeing data accuracy and integrity. Processed, cleansed, and prepared **Medicaid claim**s data for analysis using technologies like **SQL and Python.**
* Ensuring compliance with healthcare data privacy standards, such as **HIPAA**, while analyzing and sharing data.
* Designed and built dynamic reports and dashboards in **Tableau** and **Power BI**, leveraging advanced **SQL** and **ETL tools** (such as **Alteryx** and **Informatica**) to consolidate data from multiple sources for real-time decision-making.
* Used **Python** and **R** for predictive modeling and **Tableau** for visualization to support data-driven project outcomes.
* Extracted data from **Microsoft SQL Server** databases and performed initial assessments in **Excel** and **Tableau** to ensure alignment with project goals.
* Reduced manual work by 60% building strategic sales reports accurately using **Power BI** calculating monthly production sales of $186,000 and prepared compensation files sent for payroll.
* Streamlined complex processes by identifying trends, detecting systematic inefficiencies from application inception to compensation and recommended process improvements.
* Familiarity with the **Software Development Life Cycle (SDLC).**
* Worked closely with finance, marketing, and underwriting divisions.
* Trained junior associates on **Power BI** and **SQL** Star schema concepts using **SHA** **claims** **datasets**.
* Implemented **row-level security (RLS)** in **Power BI** to monitor access to sensitive patient data.
* Applied **DAX** functions and **FHIR** to create measures and calculated fields for analyzing data, enabling logical visualization.

**Environment: - MS SQL Server 2008 client & SERVER, MS office, Legacy - Mainframes, Titanium, Rational Clear Quest, Clear Case.**

**Swiss Re Armonk, NY**

**Senior Data Analyst (Lead) June 2023 – April 2024**

* Collaborated closely with marketing team, focusing on refining audience targeting for ad campaigns using **Python, SQL, and data visualization tools (Tableau, Looker Studio, Power BI)**.
* Managed and analyzed complex data sets using **BigQuery**, developing sophisticated Python-based data pipelines for efficient data integration, aggregation, and reporting processes, enhancing the speed and accuracy of data-driven insights.
* Managed data storing and processing using big data tools such as **Hadoop,** **HDFS,** **Hive**and **Spark**
* Analyzed consumer segments and behaviors, deriving insights that led to more efficient marketing campaigns and increased **ROI**.
* Using **R**, create and validate predictive models that include regression analysis, time series forecasting, and machine learning methods.
* Actively involved in designing **A/B tests,** defining metrics to validate new user interface features, calculating sample size and checking statistical assumptions for tests.
* Created real-time dashboards in **Tableau**to visualize and monitor key metrics and A/B test processing using both external and internal data.
* Developed and implemented a comprehensive **Tableau dashboard** to support the 3-year audience and targeting strategy for team, focusing on audience generation, targeting, and measurement.
* Automated the **ETL (Extract, Transform, Load**) process for a Tableau dashboard, enhancing the campaign audience performance and modeling analysis. Developed Python scripts to merge and map extensive datasets, significantly improving data accuracy and efficiency.
* Managed Python scripts for complex data analysis procedures, harmonized multi-source datasets, and ensured high data quality with context-driven unit tests **(Pandas, NumPy).**
* Developed and maintained data workflows and integrated external **API’s** using **Python and SQL**.
* Reduced manual data visualization time from 3 days to 1-2 minutes by automating with **Python (Seaborn, Matplotlib, Plotly)**.
* Conducted sentiment analysis on customer feedback using **NLP**, transforming qualitative data into quantifiable insights to inform product improvements and strategies.
* Experience working with **Shell**Scripting in operating systems such as Linux and version-control tools such as **Git.**
* Solid experience working with cloud platforms such as **AWS**and **Google Cloud.**
* Delivered comprehensive reports highlighting key trends, patterns, anomalies, presenting findings to senior management for informed **decision-making** purposes.

**Environment: Quality Center 9.2, MS Excel 2007, PL/SQL, Business Objects XIR2, ETL Tools Informatica 8.6, Oracle 10G, Teradata V2R12, Teradata SQL Assistant 12.0**

**University of New Mexico Sep 2022- May 2023**

**Data Analyst (Graduate Assistant) Albuquerque, NM**

* Designed relational databases for in-depth analysis, guaranteeing data accuracy and integrity. Processed, cleansed, and prepared **Medicaid claims** data for analysis using technologies **like SQL and Python**.
* Exhibited proficiency in overseeing data, including database creation and upkeep, data entry, quality control, and query execution, guaranteeing data correctness and adherence to guidelines.
* Used statistical software and tools like **Excel** to analyze data, both quantitatively and qualitatively, assisting in the creation of evidence-based decisions.
* Managed program metrics, reporting, and data compliance; efficient reporting and monitoring were achieved by using technologies like **Tableau, Power BI and Excel**.

**Exceed Tek Jan 2017 – July 2022**

**Data Analyst INDIA**

* Implemented quantitative market research, user segmentation marketing, and some predictive modeling using **Machine Learning** algorithms.
* Designed and implemented robust **Python**-based data pipelines, facilitating the efficient processing and analysis of large and complex data sets.
* Developed a Python and Spark-based framework in Azure to integrate data sources like **MongoDB, Splunk, and PostgreSQL** for loading into Power Bl.
* Ensured data privacy compliance **(GDPR)** while performing complex data analysis using Python scripts.
* Use data analytics to **optimize inventory management**, reduce stockouts, and improve warehousing processes, leading to a 10% reduction in overall supply chain costs.
* Participate in cross-functional teams to execute process changes, using approaches like **DMAIC** (Define, Measure, Analyze, Improve, Control) and **DMADV** (Define, Measure, Analyze, Design, Verify) to create long-term change.
* Utilized **SQL** and **Python** to streamline data extraction processes, improving efficiency by 25%.
* Built **ETL pipelines** and spearheaded the creation of real-time dashboard reports for R&D leaders, enabling data-driven insights for informed decision-making and resulting in a 15% improvement in product conversion rate.
* Extensively used **ETL** to load data using Power Center/from source systems like Flat Files and Excel Files into staging tables and load the data into the target database Oracle.
* Managed program metrics, reporting, and data compliance; efficient **reporting and monitoring** were achieved by using technologies like **Tableau, Power BI and Excel**.
* Streamlined data analysis workflow by leveraging Python libraries **like NumPy, pandas, pipeline, SciPy, scikit-learn, NLTK, and statsmodels**, resulting in a 30% saving in analysis time and improved efficiency in generating actionable insights.
* Developed data models, dashboards and **KPI Metrics** and orchestrated seamless integration of automated outputs into **Power Bl dashboards**, enabling sophisticated and insightful analysis for informed decision-making.
* Implemented several DAX functions for various fact calculations for efficient data visualization in **Power BI.**
* Created **Power BI reports**by using Joins in multiple tables from multiples database using complex SQL queries.
* Used **DAX** functions to create measure and calculated fields to analyze data for logical visualization and used Power Queries to transform the data.
* **Deployment of SSIS packages**and **schedule SQL Server Agent**jobs to automate the run of SSIS packages on a weekly basis.
* Creating **JIRA** tickets for the defects identified and tracking them to closure and creating the process flow documentation for the KPI metrics in confluence including the calculation of each metric and data mapping.
* Leveraged advanced **Excel** formulas like **VLOOKUP, INDEX and MATCH, SUMIF, and COUNTIF** to automate data analysis, ensuring real-time insights and facilitating strategic decision-making using pivots and charts.
* Resolved data issues and discrepancies in collaboration with business teams.
* Collaborated with business and product teams to identify data-driven insights and trends, providing actionable recommendations to inform data-backed product strategies.

**Environment: Informatica 8.1, Data Flux, Oracle 9i, Quality Center 8.2, SQL, TOAD, PL/SQL, Flat Files, Teradata**

**CERTIFICATIONS:**

Council for Six Sigma Certification – Green Belt

**EDUCATION:**

B. Tech in Computer Science JNTU Hyderabad May 2018

MS in Business Analytics University of New Mexico May 2023