SHUMETIE TEFERA
DATA ANALYST/DATA ENGINEER

E-Mail: tshumetie5@gmail.com   **Tel:** 559-216-3852
LinkedIn: https://linkedin.com/in/shumetie-a-tefera527
Website: s3shumetie23.com   **Address:** Los Angeles, CA

ROFESSIONAL SUMMARY

Results-driven Data Analyst with 5+ years of experience in analyzing complex datasets to extract meaningful insights. Proficient in data manipulation, statistical analysis, data visualization techniques, spreadsheet, and Machine Learning Algorithms. Skilled in Excel, SQL, Python, Power BI and Tableau for data extraction, transformation, analysis, and visualization. Data Cleansing and Data Migration. Experience in Implementing High Availability, expertise in Data Warehousing, Reporting and BI solutions. Strong background in designing, developing, and implementing data integration solutions using SSIS, MuleSoft Anypoint Platform and Apache Kafka. Proficient in carrying out secure, scalable, and cost-effective sets of data analytics tasks in the data pipeline on the cloud. Strong communication skills and collaborative with cross-functional teams. Excited to leverage analytical expertise to contribute to data-driven decision-making and enhance business performance. Flexible and versatile to adapt to any new environment, technology, and work on any project.

KEY SKILLS

• **Data Analysis** data collection, cleansing, transformation, and exploratory analysis to identify trends, patterns, & correlations, data modeling, data mining, data warehousing, data wrangling and extracting meaningful insights in large datasets.
• **Statistical Analysis:** hypothesis testing, Time Series Analysis, and classification
• **Machine Learning:** Classification, Regression, Clustering, Feature Selection, Model Evaluation.
• **Data Visualization:** interactive and visually appealing dashboards and reports using Excel, Power BI, Tableau, Matplotlib and Seaborn.
• **Programming Languages:** SQL, T-SQL and Python for data extraction, manipulation, and analysis.
• **Data Tools:** Pandas, NumPy, Excel, Jupyter Notebook, Google Collab, Kaggle kernel and Git for data analysis and version control.
• **Problem-Solving:** analyze complex problems, break them down into manageable components.
• **Cloud Techs:** *Cloud Platforms* (Amazon Web Services (AWS), Microsoft Azure), *Cloud Storage & Warehousing (*Amazon S3, Azure ADF, Data Lake, Amazon Redshift, Azure Synapse Analytics), *Data Processing technologies (*AWS Glue, Azure Data Factory), *Data Security and Compliance (*OATH2, JWT, Open ID, IAM), *Machine Learning and AI services (AWS* SageMaker)*.
• **Source Control:** Git, TFS, Azure DevOps, p4merge and Jira
• **Communication:** Excellent verbal and written communication skills to present data findings and recommendations to both technical and non-technical audiences using tools like power point and Google Slides. Also wise in using communication tools like Teams, Slack, Zoom, GoToMeeting, and Google Hangouts, Google Meet
• **Collaborative Team Player:** Demonstrated success in working effectively with cross-functional teams and stakeholders to meet project objectives.

TOOLS, ENVIRONMENTS & TECHNOLOGIES

RDBMS: *MSSQL Server, MySQL, PostgreSQL, MS Access*
Reporting & Visualization Tools: *power BI (Power Query, DAX), Tableau (MDX), Excel, SSIS, SSRS, Crystal Report, SPSS*
Languages: Python, *T-SQL, HTML, CSS, Bootstrap, JavaScript, React js, Node js*
IDE: *Jupyter Notebook, VS code, Google Collab, Kaggle kernel, PyTorch, Visual Studio, SSDT, Eclipse, Sublime text*
Operating Systems: *Windows Server, Linux Ubuntu*
Methodology: *Agile, Waterfall*

PROFESSIONAL EXPERIENCE

McKesson Corporation – Chino, CA *Jul 2021-Present*
Sr. Data Analyst
Responsibilities:
- Perform data cleaning, transformation, and validation on large datasets using SQL and Python, ensuring data integrity and accuracy.
- Conduct statistical analysis to identify trends, patterns, and correlations.
- Partnered with Solutions Architect and Technical Consultants early in the project to discuss key points of interest for design and analysis.
- Fully automated existing analytics solutions using AWS Glue.
- Identified opportunities for process improvement and data quality enhancement.
- Developed automated reports and visualizations using Tableau, providing actionable insights to key stakeholders and senior management.
- Collaborated with cross-functional teams to define data requirements and design analytical frameworks to support business objectives.
- Conducted ad-hoc data analysis to address specific business questions and provide actionable recommendations for process improvements.
- Developed reports and presentations, to communicate analytic insights to management.
- Used MLLP Protocol in transmitting Health Level Seven (HL7) messages via TCP/IP writing code programs to help capture and organize relevant data.
- Identified problem spots by running performance recordings, interacted with the Workbook, interpreted it and took specific actions to optimize performance in Tableau.
- Provided guidance and support to junior team members, assisting with technical challenges and problem-solving.
- Stay updated with industry trends and emerging technologies in data analysis.

DigiSure Insurance Solutions Inc –Burlingame, CA *Sep 2020 - June 2021*
Insurance Data Analyst
Responsibilities:
- Conducted data collection (including web scraping), cleaning, and analysis ensuring high data quality and integrity for various projects.
- Developed visual reports, KPI scorecards, tooltips, filters, parameters, and dashboards using Power BI desktop for analysis of claims amount, number, status, assured amount by policy, productivity trend, top sales by producer and sales by region.
- Conducted market research and competitor analysis to identify growth opportunities and inform business strategies.
- Interpreted data and analyzed outcomes to ensure accurate research results.
- Used advanced Microsoft Excel and the XLOOKUP function to efficiently create pivot tables and pivot reporting.

- Implemented Advanced Analytics Models and Machine Learning algorithms along with data visualization to support critical operational and analytical applications.
- deconstructed existing reports and rebuilt on a new data stack.
- Led data analysis initiatives, working closely with cross-functional teams to identify business requirements and translate them into data-driven solutions.
- Conducted exploratory data analysis (EDA) to identify trends, patterns, and outliers, resulting in process optimization and business growth.
- Built and maintained predictive models using machine learning algorithms to forecast sales, customer behavior, and demand.
- Collaborated with stakeholders to define KPIs, track performance metrics, and generate data-driven recommendations for business improvement.
- Presented analytical findings and insights to non-technical audiences in a clear and concise manner with power point enabling informed decision-making.
- Imported data from flat files into excel, implemented pivot tables and charts in excel for summarizing data.

First General Bank- Los Angeles, CA   *Aug 2019- Jul 2020*
SQL/ETL Developer
Responsibilities:
- Performed Data Extraction, Transforming and Loading (ETL) using SQL Server Integration Services (SSIS), Bulk Insert and BCP for Enterprise Data Warehouse.
- Involved in Planning, Defining and Designing databases based on business needs for insights & analytics.
- Studied in-house requirements and conducted Joint Requirement Development (JRD) session to analyze the database requirement in detail.
- Maintained Alerts, notifications, Jobs, Job schedules.
- Developed Shell Scripting to automate repetitive database processes.
- Designed, Created, and managed primary objects such as tables, triggers, database links, indexes, and privileges, based on logic design model.
- Created SSIS packages to automate maintenance, database backup and restoration, update statistics, re-build indexes as well as create error log for event during data load.
- Performed performance tuning and optimization of SQL queries, enhancing database efficiency, and improving query response times.
- Used Power Query, DAX to create Dimensions, Measures and Calculated fields.
- Designed paginated reports and vivid visualizations using SSRS.
- Defined Key Performance Indicator metrics to create scorecards.

City National Bank of New Jersey– Newark, NJ   *Feb 2018- May 2019*
MuleSoft Data Integration Engineer
Responsibilities:
- Designed, developed, and implemented end-to-end data integration solutions using MuleSoft Anypoint Platform.
- worked closely with cross-functional teams to gather requirements, analyze data integration needs, and develop robust and scalable integration solutions.
- ensured smooth data flow and integration across various systems and applications.
- Built and configured MuleSoft flows, connectors, and APIs to enable seamless data integration between systems and applications.
- Developed data mappings, transformations, and validations to ensure data accuracy and consistency.
- Monitored, troubleshooted, and resolved integration issues to ensure uninterrupted data flow.
- Used Mule Batch processing to handle the large data and used the batch processing capabilities of DataWeave to transform large payloads.
- Worked on Mule connectors like HL7 MLLP, FILE, HTTP, MySQL database, Salesforce.

- Developed and maintained technical documentation, including architecture diagrams, design specifications, and integration process documentation.
- Handled Message Transformation between XML, JSON, CSV
- Used CI / CD tools like Jenkins, Maven, and GitHub during the development and deployment.
- Defined integration points, data flow, integration strategy and developed functional and technical specifications to support cross-functional, multi-system solutions.
- Conducts API modelling and validation with clients using RAML and mocking MuleSoft API implementation.

EDUCATION & CERTIFICATIONS

- MSc in Applied Mathematics-Numerical Analysis
- AWS Certified Solutions Architect-Associate
- Salesforce Certified MuleSoft Developer
- Certified in SQL Server DBA, MS BI
- Certified in Full Stack Developer (MERN Stack)

RECENT PROJECT

Drug Reimbursement Analysis:
As Reimbursement Dashboards are interactive data visualization tools that provide insights into the *reimbursement processes* and *financial aspects* of healthcare services, this intensive analysis is used to aggregate and present data related to claims, payments, reimbursements, and financial metrics. It served the purpose to
- give self-service analysis to closely monitor expenses related to the reimbursement program and improve profitability by displaying critical metrics (claim volume, reimbursement amounts, denial rates, average payment times and revenue by payer).
- drill down into specific time periods, payers, providers, and service categories to identify patterns, outliers, and trends that impact reimbursement processes.
- make comparisons of payer performance (reimbursement rates, payment accuracy, and timeliness.)
- advance Revenue Forecasting and Compliance Monitoring (tracking reimbursement changes due to new policies and organizational structures)