Akhila Etikala
+1 (919) 410-6331
akhila.data36@gmail.com
Sr. Data Analyst /Business Analyst

PROFESSIONAL SUMMARY:

- Sr. Data Analyst with extensive experience spanning over 10+ years across various domains, including retail, healthcare, and Finance.
- Managed extensive data integration projects using SQL and Informatica, ensuring efficient data flow across multiple platforms.
- Developed predictive analytics models utilizing AWS Cloud and Amazon SageMaker, enhancing business forecasting capabilities.
- Automated report generation using Tableau, integrating data from AWS Redshift to deliver real-time insights.
- Designed and executed data cleansing processes with Alteryx, streamlining data accuracy and usability.
- Conducted complex data analyses using R, identifying key trends and opportunities for business optimization.
- Implemented machine learning algorithms with Apache Spark ALS, improving predictive accuracy and decision-making.
- Enhanced data workflows using AWS Recognition, optimizing facial recognition and image processing tasks.
- Leveraged SQL for database management, ensuring robust data storage, retrieval, and manipulation.
- Utilized Power BI to create interactive dashboards, providing stakeholders with actionable insights into organizational performance.
- Employed Excel for advanced data manipulation and analysis, supporting complex financial modeling and reporting.
- Orchestrated data quality improvements using Informatica Data Quality tools, elevating the integrity of healthcare data.
- Applied Ataccama for data governance, ensuring compliance with regulatory standards and internal policies.
- Adopted Agile methodologies to optimize data project management, enhancing team productivity and project outcomes.
- Built retail analytics solutions with Qlik Sense, driving enhancements in customer engagement strategies.
- Managed cloud-based data storage solutions using Azure SQL and Stream Analytics, ensuring scalable and secure data handling.
- Produced comprehensive retail reports using Data Studio, guiding marketing strategies and inventory management.
- Integrated MATLAB and Matplotlib in data analysis processes, refining statistical models for market analysis.
- Set up ETL processes with SSIS, facilitating efficient data extraction, transformation, and loading.
- Performed data transformations and cleansing using Oracle Data Integrator and XML, maintaining high standards of data quality.
- Crafted and maintained extensive documentation for data processes using Git and Linux, ensuring reproducibility and compliance.
- Implemented real-time data processing solutions with Apache NiFi and Kafka, improving operational efficiency and data timeliness.
- Analyzed business data using Microsoft Excel, providing foundational insights for strategic planning and resource allocation.
- Utilized Power BI and OBIEE for dynamic reporting and business intelligence, enhancing organizational decision-making and performance monitoring.

TECHNICAL SKILLS:

Query Languages & Database Management	PostgreSQL, NoSQL, MySQL, MS SQL server, Oracle, Azure SQL, Hive
Data Analysis & Reporting	Microsoft Excel, Power BI, SSRS, SAS, SAP Business Objects
Data Visualization	QlikView, QlikSense, Tableau, Power BI, Looker, Google Data Studio

Big Data Technologies	Apache Hadoop, Apache Spark, Snowflake, Databricks
Data Modeling	Erwin, Microsoft Visio, Star Schema, Snowflake Schema, Azure Synapse Analytics, Toad Data Modeler, MongoDB, Cassandra
Data Quality & Governance	Informatica Data Quality, Ataccama, Trifacta, Collibra, Open Refine
Programming/Scripting	SQL/PL-SQL, R, Python, Scala
Machine Learning	TensorFlow, PyTorch, Amazon SageMaker, Azure ML
Cloud Technologies	GCP(Beam, Data Proc, Big Query, Data Flow, composer), Microsoft Azure(Blob storage, ADLS, ADF, Synapse, Azure SQL Server), AWS (Lambda, S3, Redshift, Athena, Dynamo DB, RDS, Glue)
Version Control	Github, BitBucket, GitLab
ETL/ Data Integration	Informatica, Alteryx, OWB, SSIS, Abnitio
Devops Tools	Dockers, Jenkins, Terraform, Ansible, Kubernetes
Statistical Analysis	R, MATLAB, DAX, MDX
Data Formatting	CSV, Parquet, avro, json, xml, txt
Data Migration Tools	Talend, AWS DMS, Azure DMS, DB Convert, Fivetran, Apache Sqoop
Data Warehousing	Azure SQL, AWS Redshift, data warehouse management, Azure Data bricks
Image and Video Analysis	AWS Rekognition
Ticketing Tools	JIRA, ServiceNow, Rally, TFS, Azure Dashboard
Documentation & Reporting	SharePoint, Confluence, Lucid Chart
Agile and Scrum Methodologies	Scrum, Agile

PROFESSIONAL EXPERIENCE:
Client: PNC, Pittsburgh, PA   **Jan 2023 to till date**
Role: Sr. Data Analyst/Data Modeler
Roles & Responsibilities:
- Enhanced data analysis capabilities using SQL for insightful financial trend prediction and management reporting.
- Developed financial models in Python, utilizing libraries like NumPy and pandas for complex data processing and analysis.
- Created comprehensive data visualizations in Tableau, providing stakeholders with actionable insights for strategic decision-making.
- Managed AWS Lambda functions to automate financial data processing tasks, enhancing operational efficiency and accuracy.
- Integrated TensorFlow and PyTorch in financial predictive models, improving risk assessment and management practices.
- Utilized Azure ML to deploy scalable machine learning solutions, enhancing predictive capabilities in financial services.
- Analyzed financial data using SQL and Excel, providing comprehensive reports to stakeholders for informed decision-making.
- Employed MATLAB and TensorFlow for developing advanced financial models, improving prediction accuracy and reliability.
- Developed and deployed financial dashboards using Tableau, enhancing visibility into financial performance and opportunities.
- Utilized AWS S3 for efficient data storage and management, ensuring data availability and security for financial analysis.
- Implemented AWS Athena for querying large datasets, improving speed and efficiency in financial data analysis.
- Collaborated with team members using Agile methodologies, ensuring project deliverables met high standards and deadlines.
- Designed financial reports using Python and Tableau, integrating data from diverse sources for comprehensive analysis.

- Employed AWS Lambda and Azure ML for automating data workflows and deploying machine learning models, respectively.
- Analyzed market trends using advanced statistical tools, providing strategic insights to enhance financial decision-making.
- Developed predictive models using Python, employing libraries like NumPy and Pandas for detailed financial data analysis.
- Integrated financial data with AWS services, enhancing data workflows and analytics capabilities in financial services.
- Utilized Jupyter Notebook for data exploration and modeling, improving understanding and application of financial data.
- Enhanced data security practices using AWS technologies, ensuring the integrity and confidentiality of financial data.
- Employed MATLAB and PyTorch for complex financial simulations, enhancing predictive modeling and risk management.
- Facilitated requirement gathering sessions to understand and address specific financial data analytics needs effectively.

Environment: SQL, Python (NumPy, pandas, TensorFlow, PyTorch, MATLAB), Tableau, AWS Lambda, Azure ML, Excel, Jupyter Notebook, AWS S3, AWS Athena.

Client: HCSC, Chicago, IL   **Nov 2020 to Dec 2022**
Role: Data Quality Analyst/Business Analyst
Roles & Responsibilities:
- Implemented Informatica Data Quality tools to improve the accuracy and integrity of healthcare data, supporting critical care decisions.
- Collaborated with IT teams to deploy Ataccama for data governance, streamlining data management practices across healthcare systems.
- Conducted detailed data profiling and statistical analysis with R and SQL Server, identifying trends and anomalies in patient data.
- Managed and optimized Azure Databricks clusters to handle varying workloads and processing demands, ensuring optimal performance and resource utilization.
- Designed, developed, and implemented end-to-end data pipelines using Azure Data Factory to orchestrate and automate data workflows.
- Created efficient Extract, Transform, Load (ETL) processes within Azure Data Factory, integrating data from various source systems into target data warehouses or data lakes.
- Automated data workflows with Alteryx to accelerate healthcare data processing, reducing time-to-insight for patient care improvement.
- Led the adoption of Agile methodologies in data management projects, increasing team flexibility and project delivery speed.
- Enhanced data integrity checks using Trifacta, ensuring clean and reliable data for regulatory reporting and operational use.
- Deployed Collibra for metadata management, improving data discoverability and usage across various healthcare departments.
- Streamlined healthcare operations using Google Cloud technologies, improving efficiency in data management and storage.
- Integrated Snowflake into CI/CD pipelines for automated testing, deployment, and version control of database schema changes.
- Developed predictive models using R and Apache Spark ALS to forecast healthcare trends, enabling strategic decisions for patient care.
- Facilitated collaboration between data stewards and IT teams using SQL and collaborative tools, improving data quality initiatives.
- Oversaw the implementation of MDM (Master Data Management) systems, ensuring a unified view of patient and provider data.

- Designed data validation processes that comply with healthcare regulations, employing SQL and Informatica tools to ensure data quality.
- Documented healthcare data flows and processes, ensuring compliance with healthcare data privacy regulations such as HIPAA.
- Integrated data quality insights into clinical processes, using R and Python to develop predictive models that forecast patient outcomes.
- Established a comprehensive data lineage system with Informatica Enterprise Data Catalog, tracing data from source to consumption.
- Developed automated data quality metrics and reporting dashboards using SQL and Power BI, providing transparency into data quality levels.
- Utilized GCP services, including BigQuery and Data Studio, to enhance data visualization and reporting processes, enabling accurate, real-time insights for healthcare analytics and decision-making.
- Implemented GCP BigQuery for efficient data analysis and streamlined report generation, significantly improving data accuracy and reducing time-to-insight for healthcare quality assessments and operational analytics.
- Streamlined data reporting and analytics processes using Microsoft Excel and SQL, aiding in the quick generation of regulatory and internal reports.
- Used GitHub for version control to manage scripts and healthcare data pipelines, ensuring secure collaboration and compliance with healthcare regulations.

Environment: Informatica Data Quality, SQL, Ataccama, R, Python, SQL Server, Trifacta, Collibra, Open Refine, Power BI, Microsoft Excel, Azure Databricks, Azure Data Factory, ETL, GCP BigQuery, Datastudio, Alteryx, GitHub, Informatica Enterprise Data Catalog

Client: Safeway, Pleasanton, CA   **Feb 2018 to Oct 2020**
Role: Data visualization analyst/Business Intelligence Analyst
Roles & Responsibilities:
- Led the creation of interactive dashboards using Power BI, enabling real-time visualization of retail sales and inventory data.
- Analyzed customer purchase patterns using SQL and Excel, providing insights that informed marketing strategies and promotions.
- Utilized Qlik Sense to develop comprehensive reports, allowing stakeholders to track key performance indicators across various retail channels.
- Managed Azure Stream Analytics to process and analyze real-time data, ensuring timely insights into retail operations.
- Developed advanced data visualizations with Looker, facilitating better decision-making for management by presenting complex data in an intuitive format.
- Developed predictive models using Python libraries such as NumPy and pandas to enhance business trend forecasting.
- Leveraged Snowflake's cloud data warehouse to streamline data storage and retrieval processes, resulting in a 30% reduction in query execution time.
- Employed MATLAB for statistical analysis, enhancing the accuracy of sales forecasts and demand planning.
- Streamlined data integration processes using Azure SQL, ensuring seamless data flow from various sources into the reporting platforms.
- Created and maintained data-driven models using Excel and Power BI, providing actionable insights into customer behavior and sales trends.
- Utilized Matplotlib for advanced data plotting, enabling detailed analysis of retail metrics over time.
- Integrated data from multiple sources using Data Studio, producing comprehensive reports that supported strategic business decisions.
- Managed cloud-based data storage and processing using Microsoft Azure, ensuring data availability and security for analytics purposes.
- Leveraged R for conducting deep statistical analysis, improving the understanding of customer demographics and purchasing patterns.

- Enhanced retail data accessibility by developing self-service analytics solutions in Power BI, empowering users to generate their own insights.
- Leveraged Azure Synapse Analytics for large-scale data processing and real-time analytics, facilitating data modeling and reporting for financial operations.
- Created automated reporting pipelines using Azure Data Factory, reducing the time required to produce regular business intelligence reports.
- Implemented data visualization best practices in Power BI and Qlik Sense, improving the clarity and usability of reports.
- Conducted workshops to train team members on using Power BI and Looker, fostering a data-driven culture within the organization.
- Migrated financial data using Azure Database Migration Service (DMS), ensuring minimal downtime and data integrity during transitions from on-premise to cloud systems.
- Utilized SQL to query large retail datasets, extracting relevant information for ad-hoc analysis and reporting needs.
- Employed Azure Data Lake for storing and analyzing large volumes of retail data, supporting big data initiatives.
- Developed custom visualizations in Power BI to represent complex data sets, aiding in the communication of insights to non-technical stakeholders.

Environment: Power BI, SQL, Python, snowflake, Excel, Qlik Sense, Azure Stream Analytics, Looker, MATLAB, Azure SQL, Matplotlib, Data Studio, Microsoft Azure, R, Azure Data Factory, Azure Data Lake, Azure DMS.

Client: Ceequence Technologies, Hyderabad, India   **Mar 2016- Nov 2017**
Role: Data Analyst/Data Mining Specialist
Roles& Responsibilities:
- Engineered advanced analytics models using AWS Cloud to predict trends and behaviors, driving strategic business decisions.
- Created comprehensive dashboards and reports in Tableau, integrating data from AWS Redshift to provide actionable insights.
- Automated data extraction and analysis using Alteryx, reducing manual effort and enhancing process efficiency.
- Conducted in-depth data analysis with SQL, uncovering insights that shaped customer retention and acquisition strategies.
- Implemented machine learning models using Amazon SageMaker and Apache Spark ALS, improving prediction accuracy for business outcomes.
- Designed data validation processes with Excel and SQL to ensure the accuracy and reliability of business reports and analysis.
- Leveraged AWS Rekognition for image data processing, improving data accuracy and enhancing business operations.
- Managed AWS S3 buckets for efficient data storage, ensuring scalability and security of large datasets.
- Developed interactive Tableau visualizations, enabling real-time tracking of key performance indicators (KPIs).
- Applied R for statistical analysis to extract insights on business performance and customer demographics.
- Enhanced predictive analytics capabilities by refining forecasting models using AWS S3 and Amazon SageMaker.
- Automated data integration and workflow processes with AWS Lambda, optimizing data operations across multiple business functions.
- Utilized AWS CloudWatch for monitoring and optimizing performance of data processing jobs, ensuring efficiency and high availability.
- Employed Apache Spark ALS for advanced analytics, modeling interactions between millions of data points to generate valuable insights.
- Designed and implemented secure data-sharing environments using AWS IAM, ensuring compliance with privacy and security regulations.

- Led the migration of legacy data systems to AWS Redshift, optimizing data warehousing solutions for improved analytics.
- Automated business process models using Python and R, speeding up decision-making and improving operational efficiency.
- Leveraged AWS QuickSight for ad-hoc reporting and data exploration, improving business agility in responding to inquiries.
- Developed and implemented a data governance framework using AWS Glue, ensuring standardized data definitions and structures across the organization.
- Optimized data retrieval and analysis processes using SQL and Python, improving responsiveness across business operations.
- Implemented data encryption and security measures using AWS KMS to safeguard sensitive business data.
- Streamlined data pipelines using AWS Data Pipeline, ensuring timely and error-free data availability for reporting and analysis.

Environment: AWS Cloud, Tableau, Alteryx, SQL, Amazon Sage Maker, Apache Spark ALS, Excel, AWS Rekognition, AWS S3, R, AWS Lambda, AWS CloudWatch, AWS IAM, AWS Redshift, Python, AWS QuickSight, AWS Glue, AWS KMS, AWS Data Pipeline.

Client: Maistering B.V, Maharashtra, India   **Aug 2014- Feb 2016**
Role: Data integration analyst
Roles & Responsibilities:
- Managed ETL processes using SSIS and Informatica, ensuring efficient extraction, transformation, and loading of data across systems.
- Conducted data cleansing and transformations with Oracle Data Integrator, enhancing data accuracy for analytics and reporting.
- Implemented real-time data processing solutions using Apache NiFi, improving the timeliness of data availability for business use.
- Developed and maintained documentation for data integration workflows, promoting best practices and knowledge sharing.
- Utilized Kafka for stream processing, enabling real-time data analytics and decision-making capabilities.
- Performed data analysis using SQL and Microsoft Excel, providing insights that supported strategic business decisions.
- Integrated data from multiple sources using RESTful APIs, ensuring seamless and consistent data flows into the central database.
- Facilitated data migrations between systems with SSIS, ensuring data consistency and minimizing downtime during transitions.
- Developed automated data integration routines using Informatica, reducing manual intervention and errors in data processing.
- Managed data storage and retrieval in a Linux environment, ensuring high availability and performance of database systems.
- Utilized XML and JSON for data interchange between systems, ensuring compatibility and efficiency in data communications.
- Employed Git for version control, maintaining an organized and trackable history of data integration projects.
- Provided support and troubleshooting for data integration issues, leveraging deep knowledge of SSIS and Informatica to resolve challenges quickly.

Environment: SSIS, Informatica, Oracle Data Integrator, Apache NiFi, Kafka, SQL, Microsoft Excel, RESTful APIs, Linux, XML, JSON, Git.

Education:
- Bachelor of Technology (B.Tech) in Computer Science from Osmania University, Hyderabad, Telangana, India. - 2014