Kiranmayee Sai

**kiranmayeesai7799@gmail.com** **| +1 9809446336 |** [**www.linkedin.com/in/sai-bajantri**](http://www.linkedin.com/in/sai-bajantri) **|** [**Tableau Public**](https://public.tableau.com/app/profile/kiranmayee.sai/vizzes)

**SUMMARY**

Experienced Data Professional with over 5 years of expertise in designing scalable data infrastructures, pipelines, and analytical dashboards. Proficient in Python, SQL, PySpark, AWS, and tools like Power BI, Tableau Snowflake, and Databricks, delivering high-performance, data-driven solutions. Skilled in ETL/ELT processes, data modeling, and visualization, with strong business acumen, focus on transforming complex datasets into actionable insights to drive impactful business decisions.

**CERTIFICATIONS & TECHNICAL SKILLS**

**Professional Certificates:** AWS Academy for Data Analytics, Data camp on SQL, Relational Databases in SQL, Joining data to SQL

**Databases:** MS SQL Server, NoSQL, MS Access, MongoDB, PostgreSQL, Oracle

**Tools & Languages:** C, C++, Java, Python, SQL, SAS, VBA, macros, VLOOKUP, PL/SQL, Graphana

**Data Visualization & Reporting Tools:** Tableau, Looker, Power BI, Qlik Sense, Informatica, Alteryx, KNIME, Matplotlib, Seaborn, Excel

**Hadoop Ecosystem & Cloud:** Hadoop, Spark, MapReduce, Hive, HDFS, Pig, Snowflake, AWS S3, AWS RDS, AWS DynamoDB

**ML & AI:** Scikit-Learn, Pandas, NumPy, Exploratory Data Analysis, Predictive Modeling, Feature Engineering, GenAI, NLP, Time Series

**Operating Systems:** Microsoft Windows, Linux (Ubuntu), Microsoft Office Suite (Word, Excel, PowerPoint)

**Project Management and Version Control:** Jira, Gitlab, Bitbucket, Bamboo, Jenkins

**PROFESSIONAL EXPERIENCE**

**Data Analytics Engineer,** Blue Cross Blue Shield, Charlotte, NC 03/2024 – Current

* Developed automated data pipelines using Informatica PowerCenter and AWS Glue, processing over 300K healthcare records daily (claims, member, and provider data) with strict HIPAA compliance.
* Enhanced performance of complex SQL joins and window functions, reducing query latency for large-scale claims and eligibility reports by 60%.
* Implemented query optimization techniques, including indexing and partitioning, to reduce SQL query execution time by 40%.
* Created interactive Power BI dashboards to visualize KPIs such as claim denials, member utilization, and care costs, improving operational reporting and reducing manual tasks by 50%.
* Utilized Python (Pandas, NumPy) to perform data profiling, anomaly detection, and cleansing operations, improving data quality across key datasets.
* Developed streaming data pipelines using Apache Kafka to process log data in real-time.
* Implemented Elasticsearch to index and query large datasets, improving search performance for log and metrics data by 40%.
* Conducted data mining and segmentation analysis on historical claims data to identify patterns in treatment outcomes and support risk scoring models.
* Integrated Python scripts within Power BI for data manipulation and advanced analytics, leveraging machine learning results.
* Developed interactive Tableau dashboards to visualize key performance metrics, reducing manual reporting time by 40%.
* Developed ETL processes to ingest structured and unstructured data from AWS S3 to Google BigQuery.
* Implemented batch data ingestion and transformation using AWS Glue + PySpark, enabling efficient processing of multi-terabyte clinical datasets.
* Integrated AWS S3 and Lambda for secure data storage and serverless workflows, improving data accessibility across analytics and compliance teams.
* Supported regulatory and compliance reporting by validating enrollment and claims data accuracy and aligning reporting logic with CMS and NCQA standards.
* Collaborated with cross-functional teams (Data Governance, Actuarial, BI) to refine data lineage and ensure adherence to enterprise data quality standards.

**Healthcare Research Analyst,** UNCC, Charlotte, NC01/2023 – 12/2023

* Engineered web scraping scripts using Python's BeautifulSoup to extract comprehensive healthcare data, enriching data depth for analysis.
* Orchestrated data pipelines using PySpark to transfer data from AWS S3 buckets to Google BigQuery.
* Utilized natural language processing techniques to perform sentiment analysis on patient feedback from surveys and social media platforms.
* Created comprehensive Power BI dashboards to monitor healthcare claims, reducing data processing time by 30%.
* Integrated Tableau with SQL databases and cloud data sources (AWS Redshift, Google BigQuery) to create real-time visualizations.
* Automated data quality checks for missing values, duplicates, and schema mismatches during ingestion.
* Implemented Kafka Connect for integrating Kafka with various data sources, including MySQL and AWS RDS.
* Leveraged BigQuery’s SQL engine to perform advanced analytics on healthcare data, reducing processing time by 30%.
* Refined the performance of SQL databases through strategic index creation and comprehensive tuning of queries, achieving a reduction in average execution times by 25 seconds to boost overall data access speed.
* Designed interactive dashboards in Looker, integrating data from sources to provide actionable insights into key performance metrics, improving stakeholder efficiency by 30 minutes per report.
* Conducted A/B testing to evaluate the impact of process changes, analyzing results with Python to identify an improvement of 10 units in operational effectiveness for tested initiatives.

**Lead Data Analyst,** Tata Consultancy Services, India 03/2021 – 07/2022

* ﻿Enhanced SQL and Python query performance, reducing execution time by 30% and improving data retrieval efficiency.
* Orchestrated ETL processes Using Alteryx, KNIME and Informatica for customer data, integrating over 1 million records into Data Warehouses and Data Lakes. Built pipelines, resulting in a 30% enhancement in data loading speed and system reliability.
* Leveraged KNIME’s workflow automation to process and validate datasets from diverse client systems, identifying and correcting over 200 data discrepancies monthly ensuring the IT consulting reports delivered to clients were accurate.
* Built data pipelines to load and transform data into Redshift using AWS Glue and ETL jobs.
* Configured Oracle database instances to manage over 300 records, ensuring fast data retrieval for consulting teams and improving the delivery of IT solutions for enterprise clients across global networks.
* Generated over 30 interactive visualizations using Matplotlib, illustrating key performance indicators like system uptime and patient throughput, which empowered consulting teams to propose data-driven operational strategies to clients.
* Built a Data Warehousing solution storing 2 terabytes of operational logs, decreasing latency by 1 minute and providing a unified data source for IT service reports delivered to global consulting clients.
* Designed and customized Tableau reports for stakeholders to track KPIs like revenue growth, customer retention, and operational efficiency.
* Conducted SQL performance analysis using tools like EXPLAIN and ANALYZE, identifying and resolving bottlenecks.
* Employed Spark to process and transform massive datasets from IT systems, enabling rapid analysis of 5 million+ log entries monthly and supporting real-time consulting insights for technology improvements.
* Engaged in Agile workflows across 4 continents, conducting 50+ sprint planning sessions to align IT service deliverables with client requirements, resulting in the successful deployment of 10 custom data solutions per quarter.
* Led the migration to AWS Redshift, enhancing scalability and reducing data retrieval time by around 50%.
* Managed a team of 2 Data Analysts, tracking business KPIs and collaborating with cross-functional teams to optimize data-driven strategies.

**Data Analyst, AirBnb**, India 10/2020 – 02/2021

* Analyzed booking trends and user behavior using Python (Pandas, Matplotlib) and MySQL, delivering actionable insights that supported dynamic pricing models and targeted marketing strategies.
* Designed and deployed interactive dashboards in Excel and Tableau to monitor key KPIs such as occupancy rates, conversion metrics, and revenue trends across 100+ properties.
* Optimized SQL queries and indexing strategies on large-scale datasets (>1 TB), improving data retrieval efficiency for real-time property searches and user interactions.
* Maintained and supported Oracle databases, ensuring the accuracy and availability of daily transactional data including bookings, payments, and host settlements.

 **Financial Data Analyst,** Thomson Reuters, India 12//2019 – 09/2020

* Managed enterprise-level reporting by aggregating costs from sources like SAP and SharePoint, consolidating over 500 GB of data into SQL. Refreshed Tableau, Graphana dashboards for multi-dashboard overviews of IT expenses, serving more than 50 stakeholders.
* Crafted 10+ tableau analytical dashboards focusing on vital Billing and Sales KPIs, resulting in a 20% increase in data-driven decision-making efficiency.
* Integrated Elasticsearch with Python and Kafka to build a real-time data analytics pipeline.
* Established data governance protocols to ensure compliance and integrity across operational datasets, standardizing metadata for over 5,000 records monthly to support regulatory reporting.
* Configured Apache HTTP servers to optimize content delivery, decreasing page load times and improving access to news archives for 10,000+ daily users worldwide.
* Leveraged MongoDB (NoSQL) to store and inquiry datasets exceeding 6,000 records, minimizing data retrieval times by 30 seconds per query and enhancing reporting efficiency for real-time news analytics.
* Compiled and presented Excel-based audit reports, ensuring precise tracking of 1,000+ units of renewable energy resources and expenditures, enhancing financial accountability and transparency in energy optimization.
* Collaborated closely with stakeholders to identify reporting needs and deliver tailored solutions, improving satisfaction by 20%.
* Developed Robotic Process Automation (RPA) solutions using UI Path automating repetitive tasks and improving process efficiency.

**PROJECTS & PUBLICATIONS**

**Digi Farming Prediction (Python, SQL, Tableau)**

* Enhanced data quality 98% accuracy and prediction accuracy 25% using SQL, Python, and machine learning, while creating 10+ Tableau, PowerBI, Looker and Python visualizations for decision-making.

**Sentiment Analysis and Classification of Tweets using Deep Learning Models (Python)**

* Conducted sentiment analysis on Twitter data using Deep Learning models to predict sentiment scores comparable to NLTK's Vader Sentiment Intensity Analyzer.

**“Analyzing Transportation Data for Smart Parking System using Internet of Things”** - The International Journal of Analytical and Experimental Modal analysis, IJAEMA/4134, Volume XII

**EDUCATION**

**University of North Carolina at Charlotte, NC, USA Aug 2022 – Dec 2023**

Master of Science in Computer Science

#### Vignana Jyothi Institute of Engineering and Technology, India Jun 2016 - Apr 2020

Bachelor of Technology in Information Technology

**ACHIEVEMENT & LEADERSHIP COMMUNITY INVOLVED**

* Received **on the Spot** award for Reacting promptly to 95% of KPI delays and failures, deploying urgent business-critical changes within 24 hours with communication skills and analytical thinking.
* Awarded **Star of the Month** for spearheading a notable 90% reduction in critical production issues through pioneering solutions, concurrently elevating operational efficiency, and reliability.
* **Lead Technical Speaker,** Computer Society of India (CSI) where I Presented on advanced technical topics and industry trends at seminars and conferences; provided guidance and support to junior attendees on career development, coding, data analytics, and problem-solving during workshops and seminars.
* **Core Member,** National Service Scheme (NSS) I Participated in various camps addressing societal issues, contributing to initiatives aimed at community development and social awareness.