Teja Swaroop Sayya

+1(980) 230 4200 | teja.sayya108@gmail.com | linkedin.com/in/teja-sayya/ | leetcode.com/Teja_Sayya/

EDUCATION

Masters of Science, Computer Science   **Dec 2024**
University of North Carolina at Charlotte | GPA: 4.0/4.0

PROFESSIONAL EXPERIENCE

Data Analyst – Outpatient Imaging Affiliates   **May 2024 - Present**

- Developed interactive Tableau dashboards for real-time patient turnaround time analysis by connecting to BigQuery datasets, enabling radiologists and hospital administrators to monitor imaging workflows efficiently.
- Optimized dashboard performance by leveraging BigQuery's partitioning and clustering strategies, ensuring sub-second load times for reports analyzing 10TB+ of imaging and operational data.
- Created advanced calculated fields to derive key performance metrics such as average scan duration, technician efficiency scores, and patient wait times, enhancing operational decision-making.
- Implemented LOD (Level of Detail) expressions to aggregate imaging volumes per modality (MRI, CT, X-ray) at different time granularities (daily, weekly, monthly), providing granular insights for workload distribution.
- Integrated Tableau parameters and dynamic filters to allow stakeholders to slice and dice data by facility, technician, and scan type, streamlining data exploration without requiring SQL knowledge.
- Designed heatmaps and trend analysis charts to visualize imaging bottlenecks across different centers, helping to optimize scheduling and reduce patient dela
- Automated report scheduling by publishing dashboards to Tableau Server and embedding BigQuery SQL queries with incremental data refresh, reducing manual reporting overhead.

Data Engineer, Loreal   **Oct 2022 – Dec 2023**

- Designed end-to-end GCP data pipelines using Cloud Composer (Apache Airflow) to orchestrate ETL workflows, integrating BigQuery, Cloud Storage, and Pub/Sub for real-time data ingestion, reducing manual effort by 95% and saving 35+ hours monthly.
- Optimized BigQuery performance by restructuring nested datasets, partitioning tables, and implementing materialized views, cutting query costs by 40% and reducing latency by 55% for 20TB+ datasets.
- Built scalable data ingestion frameworks with Python and PySpark to migrate 5TB+ of relational data into BigQuery and Cloud BigTable, transforming unstructured data into structured formats for BI consumption.
- Leveraged Python and PySpark to migrate 5TB+ of relational data into BigQuery and Cloud BigTable, transforming unstructured data into structured formats for BI consumption.
- Automated Tableau dashboards via integrated ETL workflows in Tableau Prep and GCP, reducing ad-hoc report generation time by 70% and improving dashboard load times by 45% through optimized data extracts and caching.
- Developed advanced calculated fields and interactive heatmaps in Tableau to derive key performance metrics and visualize regional sales and inventory trends, enabling granular analysis and actionable insights.
- Deployed Dataflow pipelines with Python to validate over 1TB of product datasets in BigQuery, ensuring 99.98% data quality and accuracy for global inventory tracking.
- Collaborated with cross-functional teams using Bitbucket for version control and Confluence for documentation, ensuring efficient deployment of Python-based ETL workflows and enhancing team productivity by 25%.

Data Analyst, CVS Health   **Sept 2021 – Oct 2022**

- Automated ETL workflows using Tableau Prep and Google Cloud Composer, reducing manual data preparation by 95% and enabling real-time analysis of patient claims and pharmacy operations.
- Developed scalable BigQuery pipelines to process 10TB+ of healthcare data daily, optimizing SQL queries to reduce latency by 60% for critical operational dashboards.
- Managed Google BigTable clusters for secure storage and retrieval of patient eligibility and prescription data, ensuring HIPAA-compliant access across teams.
- Built interactive Tableau dashboards with advanced calculated fields and heatmaps to track key healthcare metrics (e.g., patient adherence, drug utilization), empowering stakeholders to reduce costs by 15%.
- Automated recurring financial and inventory reports using SQL and Excel, reducing turnaround time by 25% for pharmacy inventory management.
- Orchestrated end-to-end GCP data pipelines (BigQuery, Composer, Pub/Sub) to process real-time streams of patient engagement data, enhancing predictive analytics for care interventions.
- Conducted root-cause analysis of data discrepancies in EHR systems using Python scripting and automated weekly/monthly KPI reporting via Tableau scheduling tools, resolving over 30 critical issues impacting patient billing cycles.

Data Analyst, Hewlett Packard Enterprise   **Jan 2019 – Sept 2021**

- Engineered AWS-based ETL pipelines using Python , SQL , and AWS Glue , automating ingestion of 5TB+ daily server logs into Amazon Redshift , reducing latency by 50%.
- Architected an Amazon DynamoDB cluster for IoT telemetry from 50K+ servers, enabling real-time predictive maintenance with Python-driven anomaly detection models.
- Slashed Amazon Redshift storage costs by 35% via tiered archiving to Amazon S3 and partitioning 15TB+ supply chain datasets for 200+ global users.
- Automated legacy Excel-based financial reports into Amazon Redshift and Tableau , improving HPE's quarterly revenue forecast accuracy by 30%.
- Designed 15+ interactive Tableau dashboards with advanced calculated fields and parameters for tracking patient outcomes, pharmacy inventory turnover, and claims processing efficiency, reducing manual reporting time by 70%.
- Collaborated with hardware teams to integrate device performance data into Tableau dashboards , identifying bottlenecks in server efficiency and driving a 25% reduction in support tickets.
- Streamlined HPE's IoT data ingestion by implementing Amazon Kinesis Data Streams and AWS Lambda pipelines, processing 2M+ events/hour with 99.9% reliability for real-time analytics.
- Built AWS QuickSight dashboards to provide self-service analytics for sales and operations teams, enabling real-time insights into customer usage patterns and operational metrics.
- Developed reusable Tableau templates with dynamic filters, drill-down capabilities, and advanced visualizations, improving dashboard usability and adoption by 40%.

TECHNICAL SKILLS

Programming:	**Data Structures and Algorithms (DSA)**, Java, Python, JavaScript
Cloud & Big Data:	GCP (BigQuery, Bigtable, Cloud Composer, Dataflow, Pub/Sub), AWS, Hadoop, PySpark, Hive, MySQL
BI/Visualization:	Tableau, Tableau Prep, PowerBI, AWS QuickSight, Excel
Developer Tools:	Bitbucket, Docker, Git, Jenkins, Terraform, Apache Airflow, JIRA, BitBucket, Jenkins, Google Analytics, Google Looker Studio, BigQuery, Composer, Jupyter notebook, colab
AI/ML Libraries:	Regression, Classification, Clustering, Predictive Modeling, ETL, Time Series Analysis, TensorFlow/Keras, PyTorch, Hugging Face, LangChain, LLM, RAG
Certifications:	AWS Academy Data Engineer (AWS), Google Data Analytics, Tableau

---

PROJECTS & RESEARCH

- **Chat with Teja Sayya | RAG(Retrieval Augmented Generation):** link
  Built a RAG-based AI agent using LLM functionality and custom data, enabling scalable and context-aware conversational AI Agent.
- **Connect4 with AI | Alpha-Beta Pruning Search Algorithm:** GitHub Repository
  Developed an AI-powered Connect Four game using Python and Pygame with an interactive GUI. Implemented various algorithms like Minimax and Alpha-Beta Pruning to create challenging AI opponents. Skills: Python programming, AI algorithms, and Search Algorithms via games.
- **Alzheimer's Disease Analysis | Machine Learning:** Github Repository
  Researched ML models to analyze the impact of various factors on Alzheimer's risk using Linear Regression, Random Forest, SVMs, and XGBoost.
- **MakeItTalk | Deep Learning models:** Hugging face spaces
  On research paper "MakeItTalk". Converts an Image & Audio file into Facial Animation Video. Used Convolution Neural Networks(CNNs), Recurrent Neural Networks (RNNs), Generative Adversarial Networks (GANs), LSTM (Long Short-Term Memory) networks
- **AI-Powered Agreement Risk Analyzer | Gemini AI:** link
  SaaS Tool to extract key clauses, assess risks, and simplify document analysis with cutting-edge AI
- **Pixo.ai | Text 2 Image:** link
  SaaS AI Tool to convert Text to Image
- **Fit Check | Fashion App (Currently Working):** link
  Personalized AI fashion recommendation system.