

PAVAN SAI REDDY CHITTEMREDDY

Dallas, TX | +1 (603) 512-5505 | pavansaireddychitemreddy@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Data Analyst with 3+ years of experience across technology, consulting, and banking domains, skilled in designing **SQL/Oracle SQL queries**, developing **ETL pipelines with SSIS and PySpark**, and creating interactive dashboards in **Power BI** and **Tableau**. Proficient in **Python (Pandas, NumPy, scikit-learn, NLTK, spaCy)** for **ML and NLP applications**, including forecasting, anomaly detection, and text mining. Experienced in **data modeling, migration, and governance**, with hands-on exposure to **cloud platforms (AWS, Azure)** for scalable data integration and reporting. Adept at translating complex datasets into actionable insights, preventing financial risk, and improving decision-making efficiency across enterprise systems.

TECHNICAL SKILLS

Programming & Scripting: Python (Pandas, NumPy, scikit-learn, NLTK), SQL, R

Data Engineering & ETL: PySpark, AWS S3, AWS Glue, Lambda, SSIS, GitHub Actions

Cloud & Data Platforms: AWS (EC2, Redshift, S3, Lambda, Glue), SQL Server, Oracle, PostgreSQL, Snowflake, Azure

Data Visualization: Power BI, Tableau, Alteryx, Vlookup, Macros

Machine Learning & NLP: Regression Models (scikit-learn), Time-Series Forecasting, Anomaly Detection, Topic Modeling (spaCy), Text Mining (Python NLP)

Data Modeling & Governance: Data Governance Frameworks, Data Reconciliation, Auditing, Data Warehousing

Reporting & Automation: SQL Procedures, Power BI Dashboards, Tableau Dashboards, Git, GitHub, Excel, Pivot Table

Business Intelligence & Analytics: Predictive Analytics, Forecasting, Sales & Portfolio Analysis, Cross-Sell Analytics

Collaboration & Agile: Cross-functional Team Collaboration, Agile Methodologies

PROFESSIONAL EXPERIENCE

UBER

USA

Data Analyst

Dec 2024 – Present

- Engineered **SQL Server** pipelines processing 2M+ daily financial transaction records, optimizing indexing and queries, reducing report generation time by 40 seconds and enhancing operational efficiency.
- Programmed anomaly detection models in **Python (Pandas, NumPy)**, validating 5K+ settlement and reconciliation errors monthly, ensuring accuracy and compliance for financial reporting.
- Designed and delivered 12+ **Power BI** dashboards to track financial KPIs, including revenue, transaction volumes, and portfolio performance, reducing executive review cycles by 2 days.
- Automated **ETL** workflows using **PySpark**, **SSIS**, and **GitHub Actions**, streamlining monthly financial data refreshes and eliminating 20+ hours of manual work, improving reporting accuracy.
- Applied **ML regression models** in **scikit-learn** to forecast regional financial demand and transaction trends, supporting reallocation of 15K+ resource hours quarterly for optimized financial operations.
- Built **NLP** pipelines with **NLTK** to analyze 100K+ client feedback entries, identifying 5 recurring patterns impacting financial product satisfaction, informing actionable strategy improvements.
- Orchestrated cloud-native **ETL solutions** on **AWS** using **S3**, **Glue**, and **Lambda** to scale ingestion of 500K+ new financial records daily, ensuring seamless pipelines and high data throughput.
- Implemented scalable data storage solutions in **AWS S3** and integrated **Snowflake** for efficient querying, improving accessibility and performance for 10+ finance teams.
- Led cloud migration initiatives, transitioning on-prem financial systems to **AWS** using **EC2**, **Redshift**, and **Glue**, enhancing data processing speed, reliability, and reporting consistency.
- Collaborated with finance and audit teams to ensure **data governance**, **compliance**, and **security** across cloud-based financial systems, supporting regulatory requirements and risk mitigation.

ICICI Bank

India

Data Analyst/Service Manager

Mar 2022 – Dec 2022

- Queried and reconciled 1M+ daily transactions in **Oracle SQL/SQL Server**, ensuring settlement compliance and reducing reconciliation errors.
- Built interactive **Tableau** dashboards for loan portfolio performance, shortening executive review time by 16 hours each month.
- Deployed **Python** classification models to flag suspicious accounts, detecting anomalies across 7K+ cases and preventing **\$2M+** potential fraud.
- Developed 6+ **Power BI** dashboards to analyze cross-sell opportunities across 15+ product categories, supporting targeted sales campaigns.
- Applied text mining techniques on 50K+ service records using **Python NLP**, reducing escalation investigation cycles by 12 hours weekly.
- Automated reporting schedules with **SQL** procedures, enabling recurring delivery and saving 10 analyst hours per week.
- Partnered with **IT** on data migration of 20M+ customer records, validating accuracy to ensure zero operational downtime.

- Supported regulatory audits with **data governance frameworks**, preparing 12 quarters of validated reports with zero audit exceptions.
- Utilized **Alteryx** to streamline data workflows, improving processing efficiency for large datasets and reducing time spent on manual data preparation.

KPMG

Data Analyst

India

Mar 2020 – Feb 2022

- Extracted and transformed over 500,000 records from multiple data sources using **SQL** and **Python (Pandas)**, creating audit-ready datasets for various advisory projects, ensuring accurate and timely reporting.
- Designed and developed interactive client dashboards in **Tableau**, enabling detailed tracking of revenue leakage trends, which helped uncover **\$1.5M** in potential savings across four enterprise clients.
- Built comprehensive CFO review packs in **Power BI** by consolidating data from 10+ sources into intuitive, drill-down reports, enhancing the ability to make data-driven strategic decisions at the executive level.
- Validated and reconciled over 20 **Excel**-based datasets, ensuring full compliance with external reporting requirements, and maintaining the integrity of the data for auditing and regulatory purposes.
- Applied time-series forecasting models in **Python** to predict anomalies, significantly reducing the time taken for audit variance reporting by three days, enhancing efficiency and timely decision-making.
- Developed and implemented topic modeling pipelines using **spaCy** and **scikit-learn** to categorize audit notes, thereby accelerating client review cycles by 12 hours, improving turnaround time for feedback and adjustments.
- Migrated reporting workflows to **GitHub**, embedding version control practices across the consulting teams, which improved collaboration and reduced the occurrence of rework across five consulting teams.
- Standardized data governance and modeling frameworks in collaboration with consultants, leading to more consistent and reliable deliverables across five enterprise clients, improving overall project efficiency and quality.

PROJECT

Customer Feedback Analysis & Service Optimization

- Developed NLP pipelines in Python (using NLTK, spaCy) to process and analyze over 200,000 customer feedback entries from a banking dataset.
- Implemented sentiment analysis and topic modeling to classify feedback and identify key areas for service improvement.
- Created Power BI dashboards to present feedback insights alongside operational KPIs, enabling easy visualization of service issues.

Cloud-Based Data Migration & Reporting Automation

- Led the migration of 50 million financial records from an on-prem SQL Server, ensuring data accuracy with validation scripts written in SQL and Python (Pandas, PySpark).
- Automated ETL processes using SSIS and PySpark, incorporating basic checks to flag data anomalies.
- Developed self-service Tableau dashboards to replace manual Excel reports, saving 30 analyst hours per month.

EDUCATION

Master’s in Information Technology

2023-2024

Southern New Hampshire University, Hooksett, NH.

Bachelor of Technology

2016-2020

Hindustan University, Chennai, India.

CERTIFICATION

Google Data Analytics Professional Certificate - Coursera

Intermediate Python - DataCamp

Introduction to SQL - DataCamp