

Giri Sai Kumar Pragada | Data Analyst

📍 Binghamton, NY | ✉ girisaikumarpragada@gmail.com | ☎ (607) 245-9921 | 🔗 LinkedIn

SUMMARY

- Backed by **4+ years** of hands-on experience as a Data Analyst. Proficient in **Python (Pandas, NumPy, Matplotlib, Seaborn)** and **R (basic)**, leveraging scripting and data wrangling to automate processes, reducing manual efforts by 30% across multiple cross-functional analytics projects;
- Designed and built end-to-end data pipelines using **dbt, Apache Airflow**, and **ETL** tools, enhancing pipeline efficiency and delivering production-ready data workflows with 99.8% accuracy.
- Developed and maintained scalable data models using **Snowflake Schema** and **Star Schema** on **Snowflake, Redshift**, and **BigQuery**, improving reporting turnaround by 2x through enhanced data architecture.
- Created dynamic dashboards in **Power BI, Tableau**, and Looker, improving executive reporting accuracy by 35% and enabling real-time visibility into KPIs for 5+ business units.
- Hands-on experience with **AWS (S3, Redshift, Athena)**, **Google Cloud (BigQuery, Data Studio)**, and **Azure (Synapse Analytics)** for scalable data storage and analytics solutions across cloud environments.
- Built automated workflows using **Advanced Excel (VLOOKUP, Pivot Tables, VBA Macros)** and **Power Automate**, streamlining data transformation and reducing manual effort by 15+ hours weekly.
- Conducted **A/B Testing, Regression Analysis**, and **Predictive Modeling** to evaluate product performance and increase feature adoption by 22% across multiple digital platforms.
- Led time series forecasting and hypothesis testing for marketing and **finance teams**, achieving 90% forecast accuracy for quarterly sales projections.
- Applied AI/ML concepts and experimented with **LLMs** like **ChatGPT** to generate user insights and perform root cause analysis, cutting incident investigation time by 18%.
- Practiced **Agile methodologies (Scrum, Kanban)** with tools like **Jira** and **Confluence**, managing 8+ sprint cycles and ensuring on-time delivery of analytics deliverables.

TECHNICAL SKILLS

Languages:	Python (pandas, NumPy, matplotlib, seaborn), R (basic), SQL (PostgreSQL, MySQL, BigQuery, Oracle OTBI)
Data Visualization & BI Tools:	Tableau, Power BI, Looker, Google Data Studio, Oracle Fusion Data Intelligence (FDI)
Data Engineering:	Snowflake, Redshift, BigQuery, dbt, Apache Airflow, ETL Development, Data Modeling (Star & Snowflake Schema),
Cloud Platforms:	AWS (S3, Redshift, Athena), Google Cloud Platform (BigQuery, Data Studio), Microsoft Azure (Synapse Analytics), Oracle Cloud (ERP, EPM)
Statistical & Analytical Techniques:	A/B Testing, Hypothesis Testing, Regression Analysis, Predictive Modeling, Time Series Forecasting, Experimentation Analysis
Productivity & Automation:	Advanced Excel (Pivot Tables, VLOOKUP, Macros, VBA), Power Automate, PowerApps, SharePoint, Jupyter Notebooks, Git, REST APIs, Data Automation Scripting
Project Management & Collaboration:	Agile Methodologies (Scrum, Kanban, User Story Development, Sprint Planning), Jira, Confluence, Cross-Functional Team Collaboration, Stakeholder Management
Emerging Technologies:	AI & Machine Learning Basics, Large Language Models (e.g., ChatGPT), Prompt Engineering

EXPERIENCE

Data Analyst (Intern)

Northern Trust | Binghamton | Apr 2025 – Current

- Built automated **Python scripts** using **pandas**, **NumPy**, and **SQL (PostgreSQL)** to clean, merge, and transform multi-source financial data, reducing data prep time by 40% and improving model accuracy.
- Developed **Tableau** dashboards for internal investment teams, visualizing key metrics like **P&L trends, fund performance**, and **client flows**, which improved reporting turnaround by 3x.
- Created optimized **ETL workflows** using **dbt** and **BigQuery** to ingest structured & semi-structured data into Snowflake, handling over 5M transactions monthly with 99.8% data accuracy.
- Integrated **Jupyter Notebooks** and **Git** versioning into the team's analytics pipeline, enhancing collaboration and reproducibility across cross-functional data science efforts.
- Supported migration from on-prem PostgreSQL to **AWS Redshift** and **S3-based architecture**, contributing to 25% cost savings in cloud storage and compute efficiency.
- Collaborated with Risk & Compliance to design Looker dashboards using **Oracle** Fusion Data Intelligence, reducing regulatory reporting delays by 45%.

Environment: Python, pandas, NumPy, SQL, PostgreSQL, Tableau, dbt, BigQuery, Snowflake, R, Excel, A/B testing, hypothesis testing, Jupyter Notebooks, Git, AWS Redshift, S3, Looker, Oracle Fusion Data Intelligence.

Data Analyst

Tata Consultancy Services | India | Apr 2021 – Jul 2023

- Utilized **SQL (Oracle OTBI and MySQL)** and **Python (pandas, seaborn)** to analyze 20M+ transaction records, identifying customer churn segments and driving a 12% increase in retention.
- Built **Power BI** dashboards with row-level security for customer cohorts, reducing dependency on legacy Excel models and improving data access for 15+ sales teams.
- Created ETL pipelines in **Airflow** and **dbt** to integrate data from **CRM, ERP (Oracle Cloud)**, and payments systems into **BigQuery**, automating monthly report generation.
- Conducted regression analysis and customer lifetime value prediction using **R** and **Python**, delivering actionable insights for marketing campaign personalization.
- Leveraged **Jupyter Notebooks** and **GitHub** for reproducible notebooks that were used to train 25+ analysts across three departments on data wrangling best practices.
- Implemented **PowerApps** and **SharePoint** integrations to create internal tools for managing client feedback, reducing support resolution time by 35%.
- Participated in cloud data migration projects involving **Oracle ERP to Oracle Cloud**, contributing to seamless transition for 2 global clients.
- Mentored junior analysts on **Python scripting, data visualization standards**, and **Git workflows**, reducing onboarding time by 50%.
- Built prompt templates for **ChatGPT**-based internal ticketing support prototype, reducing manual triage time by 20% in pilot phase.

Environment: SQL (Oracle OTBI, MySQL), Python (pandas, seaborn), Power BI, Row-Level Security, Excel, Airflow, dbt, CRM, ERP (Oracle Cloud), BigQuery, R, Regression Analysis, Customer Lifetime Value Prediction, Jupyter Notebooks.

Data Analyst

Cybage Software | India | Jan 2019 – Mar 2021

- Used **Python (pandas, matplotlib)** and **SQL (PostgreSQL, BigQuery)** to automate data cleaning for client cloud usage datasets, reducing manual work by ~60%.
- Designed data models and **ETL** pipelines using **Snowflake** and **dbt** to support enterprise analytics for clients migrating to Azure Synapse and AWS Redshift.
- Created interactive **Tableau** and **Looker** dashboards for cloud cost optimization insights, leading to 18% savings across 4 enterprise clients.
- Implemented time series forecasting models for server load and capacity planning, boosting resource allocation efficiency by 21%.
- Built reusable **macros** and Excel-based dashboards with **VBA** for project tracking, enabling faster client reporting and improved transparency.
- Managed **Agile workflows (Scrum and Kanban boards)** via **Jira** and ensured 100% sprint goal delivery with support from cross-functional teams.
- Created automated reporting pipelines with **Power Automate** and **SharePoint** to monitor SLAs for IT operations teams across 3 continents.
- Conducted statistical hypothesis testing and root cause analysis on network latency issues, resulting in a 35% improvement in uptime.

Environment: Python, Pandas, Matplotlib, SQL, PostgreSQL, BigQuery, Snowflake, dbt, Azure Synapse, AWS Redshift, Tableau, Looker, Time Series Forecasting, Excel, VBA, Jira, Agile, Scrum, Kanban, Power Automate.

EDUCATION

 Master of Science in Information Systems Binghamton University, State University of New York Binghamton	Aug 2023 - May 2025
 Bachelor of Technology in Electronics and Communication Andhra University, India Visakhapatnam	June 2016 – Apr 2020

PROJECT

AI-Powered Business Insights Chatbot

Designed and developed an interactive GenAI-based chatbot that enables users to ask business questions in natural language and receive dynamic responses by executing real-time SQL queries on sales and operations data, with results visualized using charts and dashboards. The solution enhances stakeholder decision-making by bridging the gap between raw data and strategic insight through conversational analytics.

- Integrated OpenAI’s GPT model to translate user queries into optimized SQL code for PostgreSQL and MySQL databases.
- Built a clean frontend interface (HTML/CSS/JS) for interactive user input and visual feedback.
- Used Pandas & PySpark SQL for data processing and Seaborn/matplotlib for in-app visualizations.
- Embedded Power BI/Tableau dashboards for deeper executive-level analytics and reporting.
- Automated data refresh and ETL workflows using Apache Airflow and scheduled jobs.
- Followed Agile methodology (sprint planning, retros, user stories) and tracked tasks via JIRA & Git.