THANMAYA SRI SIGIREDDI
Data Analyst
Email: sigireddithanmayasri@gmail.com | PH: +1 (716) 465-9667

SUMMARY

Results-driven Data Analyst with hands-on experience in predictive modeling, ETL pipelines, and dashboard development using Python, SQL, Power BI, and AWS. Skilled in applying machine learning, statistical analysis, and cloud technologies to drive actionable business insights. Proven success in delivering scalable data solutions in HR analytics and GenAI applications.

PROFESSIONAL EXPERIENCE

Nucleus Business Solutions, Hyderabad, India May 2023 – Dec 2023
Data Analyst

- Designed and deployed machine learning models using Scikit-learn to predict employee attrition and forecast future hiring needs, improving workforce planning and reducing recruitment cycle inefficiencies.
- Built scalable ETL pipelines using PySpark and SQL to automate data ingestion from multiple HR systems, ensuring seamless integration of payroll, recruitment, and performance data.
- Led the creation of interactive dashboards in Power BI and Tableau to track HR KPIs such as time-to-hire, retention rate, and employee satisfaction, empowering leadership with real-time insights for strategic decision-making.
- Collaborated with HR managers and IT teams to align data strategies with business objectives, enhancing cross-functional communication and adoption of data-driven practices.

Nucleus Business Solutions, Hyderabad, India Oct 2022 – Apr 2023
Data Analyst Intern

- Conducted exploratory data analysis (EDA) using Python (Pandas, NumPy) and SQL on large HR datasets, identifying patterns in hiring funnel drop-offs and high attrition segments.
- Created targeted dashboards and reports to highlight recruitment bottlenecks, improve sourcing strategies, and enhance talent acquisition performance.
- Developed reusable scripts and templates to automate report generation, reducing manual workload and increasing data availability for key stakeholders.
- Participated in stakeholder meetings to present findings and contribute to the redesign of recruitment workflows based on data-backed insights.

PROJECTS

Detox Dial– AI-Powered Digital Wellness App
AI Hackathon – University at Buffalo | Spring 2025

- Built a data-driven backend for a digital wellness tool to reduce screen addiction by analyzing user personality traits and delivering personalized voice, text, and push interventions.
- Designed a behavioral profiling system using personality assessment data, storing structured user traits in MongoDB with over 85% match accuracy in interaction tone prediction.
- Developed RESTful APIs with Flask to process behavior patterns and dynamically trigger tailored interventions.
- Integrated Twilio API to deliver voice-based nudges customized to user characteristics such as empathy or assertiveness, improving daily user retention by 30% in trial sessions.
- Optimized response-time logic and personalization by streamlining backend endpoints and refining scoring algorithms for real-time behavioral engagement.

Design & Implementation of Relational Database for Laptop Specifications
Course Project – Data Mining Query Language | Fall 2024

- Developed a fully normalized PostgreSQL database (BCNF) to manage 5,000+ laptop records and customer reviews, enabling efficient storage and querying of 20+ attributes.
- Preprocessed raw CSV data using Python and SQL, improving data quality by 85% through unit standardization and regex-driven normalization.
- Implemented 25+ custom SQL queries and functions for analytics such as dynamic filtering, spec comparison, price analysis, and rating-based popularity trends.
- Optimized query performance by 70% using indexing strategies tailored to filters like brand, RAM/SSD, and grouped aggregates.
- Ensured complete ETL integrity with validation rules that handled inconsistent units and edge cases in hardware specs.

Clustering of Cryptocurrency Price Movements
Course Project – Statistical Learning and Data Mining | Spring 2024

- Led a project applying unsupervised learning to identify clustering patterns in volatile cryptocurrency markets.
- Cleaned and normalized time-series data, extracted key trend features, and applied K-Means clustering to group assets by movement behavior.
- Used the elbow method to determine optimal cluster count (k=4) and enhanced model interpretability with PCA, retaining 90% variance while reducing dimensionality by 50%.
- Visualized results using Seaborn and hvPlot to clearly contrast PCA and non-PCA clustering outcomes, improving presentation clarity by 75%.

- Built a modular preprocessing pipeline to mitigate data noise and non-stationarity, increasing silhouette score by 40%.

EDUCATION

Master of Science (M.S.) in Data Science, State University of New York at Buffalo Jan 2024 – May 2025

TECHNICAL SKILLS

- **Data Analytics & Visualization:** Tableau, Power BI, Advanced Excel, UI/UX Design, Plotly, Streamlit
- **Big Data & Databases:** PostgreSQL, MySQL, AWS (S3, Lambda, EC2, Athena), Snowflake
- **Programming & Querying:** SQL (Advanced), Python, R, HTML, CSS
- **Frameworks/Libraries:** Machine Learning, Deep Learning, NLP, Pandas, Numpy, TensorFlow, PyTorch, Scikit-learn

CERTIFICATIONS

- Data Science internship - Plasmid
- AI-Integrated Application - AI for Good Hackathon
- Python - Excel Solutions
- Data Science Foundations - Great Learning Academy
- Microsoft Power Bi - Excel Solutions
- SQL - Excel Solutions

Full certification portfolio: **View All Certificates**