



PHANI VISHNU ADDEPALLI

Data Analyst



408-368-8479 | phanivishnuaddepalli@gmail.com | www.linkedin.com/in/phaniaddepalli | <https://github.com/galaxyinvader69>

SUMMARY:

Results-driven Data Analyst with a strong foundation in designing, developing, and operating data-driven applications to support process automation and decision-making. Proficient in **Microsoft Power Platform (Power BI, Power Automate, Power Apps)**, SQL, and modern data platforms. Proficient in **Python** and **R** for **data wrangling, automation, and machine learning (regression, classification, clustering, forecasting)**. Experienced in developing and maintaining SQL-based data pipelines and creating actionable Power BI dashboards and reports. I have a solid understanding of the **software lifecycle**, including code reviews and testing, and a proven ability to collaborate with cross-functional teams to deliver user-friendly solutions. I am committed to continuous improvement and a growth mindset, with a strong focus on using solid data to support development decisions.

TECHNICAL SKILLS

Data Analysis & Visualization: Advanced **Microsoft Excel** (PivotTables, VLOOKUP/XLOOKUP, Power Query, dashboards, macros), Data visualization using **Tableau, Power BI, and Looker**, Exploratory Data Analysis (**EDA**), hypothesis testing, **A/B testing**

Programming & Scripting: Proficient in **Python** and **R** for data manipulation and analytics, including data wrangling, transformation, and visualization

Libraries: **pandas, NumPy, Matplotlib, Seaborn, SciPy**, Automation and scripting for reporting and data transformation

Databases & Querying: Proficient in writing complex **SQL** queries (joins, subqueries, window functions, CTEs), Hands-on experience with **Snowflake, Google BigQuery, and AWS Redshift**

Cloud Services: **AWS** (S3, Redshift), **Google Cloud Platform (GCP)** (BigQuery), **Azure** (Data Lake)

ETL & Data Integration: ETL tool experience with **Informatica, Alteryx**, and building custom ETL pipelines, Data wrangling and cleansing for analytics readiness

Machine Learning & Predictive Analytics: Supervised and unsupervised learning models: **linear/logistic regression, decision trees, clustering (K-Means, DBSCAN)**, Model evaluation techniques (**cross-validation, ROC-AUC**, confusion matrix), Feature engineering, dimensionality reduction (**PCA**), **time-series forecasting**

Tools: **Scikit-learn, GitHub, Statsmodels, XGBoost, TensorFlow** (basic knowledge)

Soft Skills: Strong communicator, detail-oriented, collaborative, analytical thinker, experienced in **project management** and **regulatory compliance** environments

EDUCATION

The George Washington University
Master of Science, Data Science - GPA: 3.91/4
Recipient of the Global Leaders Award: *Awarded as part of the Columbian College's Global Leaders Fellowship, covering tuition for top candidates from selected countries based on merit and achievement.*

Washington, DC
Aug 2023 – May 2025

PROFESSIONAL SUMMARY

- Proficient in conducting **Exploratory Data Analysis (EDA)** and creating **complex SQL queries** to extract and manipulate data from various databases and sources.
- Skilled in developing **interactive visualizations** using **Power BI, Tableau, and Looker** to communicate insights and trends effectively to stakeholders.
- Advanced knowledge of **Python** and **R** for **data manipulation, automation**, and the application of **machine learning** algorithms in business and operational contexts.
- Hands-on experience in developing and implementing **predictive models** and **classification algorithms**, such as regression analysis, decision trees, and clustering.
- Strong foundation in **statistical analysis** and **hypothesis testing**, with practical experience in **A/B testing** and regression modeling to drive data-driven decision-making.
- Proficient in cloud platforms such as **AWS, Google Cloud Platform (GCP)**, and **Azure**, with a solid understanding of **ETL processes** and **data integration** pipelines.
- Experienced in **data engineering**, including the design, implementation, and optimization of **ETL workflows**, data pipelines, and data transformation processes.
- Collaborative team player with strong communication skills, adept at conveying technical findings to both technical and non-technical audiences, ensuring clarity and alignment across teams.

EXPERIENCE

The World Bank | [Juris AI](https://jurisai-1358123759.us-central1.run.app/) (<https://jurisai-1358123759.us-central1.run.app/>)
Data Analyst

As part of a **collaborative team**, I contributed to the development and deployment of **JurisAI**, an AI-powered legal chatbot aimed at enhancing legal access for citizens in **Ghana** and **Sierra Leone**. The project leveraged **multilingual capabilities** and **advanced AI and cloud technologies** to bridge the accessibility gap in legal services. My role involved applying **data analysis, business analysis**, and **statistical techniques** to improve the effectiveness and scalability of the tool.

Key Responsibilities:

- **Developed JurisAI**, a multilingual legal chatbot leveraging **LLaMA, Groq API, RAG architecture**, and advanced data analysis to enhance legal access in Ghana and Sierra Leone.
- **Automated data ingestion, OCR, and semantic search** with **OpenAI embeddings** and **Pinecone**, ensuring efficient, accurate, and reliable legal information retrieval.
- **Cleaned and structured large legal datasets** using statistical models and data wrangling, delivering high-quality inputs for AI-driven responses.
- **Deployed and scaled the application** via **Docker** and **Streamlit on GCP**, improving performance and handling high traffic volumes.
- **Collaborated with legal experts and AI specialists** to align development with user needs, presented results at the **World Bank**, and applied user feedback to optimize accuracy and user experience.

Environment: LLaMA, Groq API, RAG architecture, OpenAI embeddings, Pinecone vector database, Docker, Google Cloud Platform (GCP), Streamlit UI, OCR, Document Ingestion, Multilingual AI, Semantic Search

Washington, District of Columbia
January 2025 – May 2025

Vestino Technologies
Data Analyst (Intern)

As a **Data Analyst Intern**, I worked on developing advanced **data solutions** to improve the efficiency and accuracy of **waste monitoring** and **sorting systems**. By leveraging **statistical models** and **machine learning techniques**, I contributed to the optimization of **waste management processes**.

Key Responsibilities:

- Designed and implemented **machine learning algorithms** for **compost monitoring**, increasing **organic waste conversion** by **64%** through **predictive analysis** and **real-time data insights**.
- Automated **material sorting systems**, improving **sorting accuracy** by **76%**, significantly reducing the need for **manual labor** and optimizing **operational efficiency**.
- Applied **Gradient Discriminant models** to streamline **data processing workflows**, reducing **analysis time** by **hours** and boosting **processing efficiency** by **67%**.
- Engineered an automated **error detection system** for **waste audit data**, reducing **environmental impact** by **52%** through improved **data accuracy** and **error correction**.
- Collaborated with senior analysts to develop **dashboards** and **reports**, enabling **actionable insights** into **waste management operations** and driving informed **decision-making**.

Environment: Python (NumPy, pandas), Machine Learning (Gradient Discriminant Models), SQL, Data Analysis, Waste Management Optimization

Chennai, India
January 2023 – June 2023

CERTIFICATIONS | [LinkedIn](#)

Programming in Python – Meta
Mastered Python syntax, control flow, loops, functions, and data structures; gained expertise in procedural programming and logical problem-solving.

IBM Data Analyst Professional – IBM
Learned the data ecosystem, ETL processes, and big data fundamentals; developed strong skills in data gathering, identification, and cleaning for analysis.

Google Data Analytics – Google
Gained a comprehensive understanding of the data life cycle; applied diverse tools and applications to streamline analysis and improve efficiency and accuracy.

Associate Cloud Engineer – Google Cloud
Skilled in cloud computing with hands-on experience in Compute Engine, GKE, Cloud Storage, and Cloud SQL for scalable deployments and data management.

