Heet Kailas Gala

(315)-455-0053 | heetgala7@gmail.com | https://www.linkedin.com/in/heet-gala/

SUMMARY

Data Analytics Professional with 3 years of experience in Python, SQL, Tableau, and Power BI, delivering scalable data solutions. Skilled at translating business needs into pipelines, reports, and dashboards that drive efficiency and decision-making. Experienced in data quality, governance, ETL, and cross-functional collaboration.

EDUCATION

Master of Science, Applied Data Science, Syracuse University	**Aug 2023 - May 2025**
Bachelor of Engineering, Computer Engineering, Mumbai University	**Jul 2017 - Jun 2021**

EXPERIENCE

Research Assistant, Syracuse University - Syracuse, NY — **Feb 2025 - Present**

- Increased transparency in fundraising performance by building **5+ interactive Tableau dashboards** integrating **SQL and Python-driven datasets**, enabling directors to quickly identify high-performing strategies.
- Identified **$250K+ in donations** driven by specific alumni engagement campaigns by performing **cross departmental data analysis, data modeling, and KPI tracking**, then presenting actionable insights to leadership.
- Improved decision-making efficiency by **reducing manual reporting time by 40%**, through **automating ETL pipelines, data cleaning, and visualization workflows** using Python and Excel.

Analyst, Capgemini - Mumbai, IN — **Jul 2021 - Jun 2023**

- Designed **Power BI dashboards and SQL-based reports** tracking payment failures across 5 regions, reducing resolution time, ensuring SLA compliance, and boosting customer retention by 13%.
- Developed and implemented data driven strategies increasing client revenue by 17%.
- Engineered 3 ETL pipelines with **SQL, Salesforce CRM, ERP data, and AWS Redshift**, centralizing reporting and improving data quality validation.
- Enhanced PowerBI data models and **DAX queries**, cutting dashboard load time by **65%** and automating weekly reports.
- Automated **Excel/Smartsheet reports** (VBA, Pivot Tables, macros), saving **12-18 hours weekly** and improving forecasting.
- Partnered with **business stakeholders** to gather requirements, translate them into BI solutions, and improve **business acumen driven decision making**.

Data Scientist Intern, Pie-Infocomm - Mumbai, IN — **Jun 2020 - Jul 2020**

- Migrated 30+ static Excel reports into interactive Tableau dashboards, cutting manual effort by 50% and empowering leadership with real-time insights for performance tracking and strategic planning.
- Collected, cleansed, and transformed datasets from SQL, APIs, and Excel using Python Pandas, reducing data inconsistencies by 20% and improving the accuracy of business-critical ad hoc reporting.
- Built Automated data refresh and report generation with Power Automate, reducing turnaround time by 50% and ensuring stakeholders had timely access to updated dashboards.

TECHNICAL SKILLS

Programming: Python, R, SQL, SAS, SSRS, VBA
Tools: Power BI, Tableau, ArcGIS, Amazon QuickSight, Excel (Power Pivot, Power Query, Pivot Table), Google Analytics
Data: ETL, Data Cleaning, Data Wrangling, Data Validation, Data Modeling, Data Pipelines
Cloud & Databases: Snowflake, AWS (S3, Redshift, Glue, DynamoDB, EC2), Salesforce (CRM), Relational DBs, NoSQL
Reporting & BI Development: Dashboarding, Ad-hoc Reporting, Data Quality Validation, SOP Writing
Big Data / Workflow: Apache Airflow, Kafka, Spark
Libraries & Frameworks: Pandas, NumPy, Scikit-learn, Statsmodels, TensorFlow, Matplotlib, Seaborn, Plotly, PyTorch, PySpark

PROJECT

LLM & RAG Builder — **Jul 2025 - Dec 2024**

- Built Built a PDF ingestion pipeline using **LangChain**, **HuggingFace embeddings**, and semantic search to ensure data integrity.
- Integrated Chroma vector database with Ollama's Llama2 for retrieval-augmented generation (RAG), enabling accurate, context-aware responses and enabling faster insights from large document sets.
- Optimized data chunking, storage, and validation, cutting token usage by 20% while improving retrieval quality.

NYPD Call Analysis for Crime Hotspots Prediction — **Jan 2025 - Apr 2025**

- Built and tuned **ML models (Decision Tree, KNN, Random Forest)** on NYPD 911 + weather data, boosting response time prediction accuracy by **25%**.
- Created 4 Tableau dashboards to visualize crime hotspots and response delays, enabling stakeholders to spot high-risk zones.
- Integrated and transformed multi-source datasets using **Apache Airflow** pipelines and **Amazon S3** storage, delivering analysis-ready data that improved efficiency of response time and hotspot analysis.