Heet Kailas Gala

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

SUMMARY

Innovative and detail-oriented Data Analyst with 3 years of experience in BI dashboards, ETL development, and data visualization at Capgemini and Syracuse University. Master's in Applied Data Science from Syracuse University with expertise in SQL, Python, R and Power BI. Skilled in streamlining reporting, improving data accuracy, and enabling clearer insights.

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.
- Partnered with finance teams to align fundraising data, ensuring **100% accuracy** in insights shared with senior leadership.

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**
- Consolidated 20,000+ records using **SQL, Python, and Excel**, improving **data readiness and KPI reporting** for executives.
- Developed **data cleaning and validation scripts**, reducing inconsistent records by 25% and enabling more reliable dashboards.
- Built an **Excel VBA automation tool** that reduced repetitive data processing by 8–10 hours weekly.
- Enhanced data processing workflows, reducing turnaround time by 42% and boosting reporting accuracy.
- Conducted **A/B testing and what-if analysis** to identify operational process improvements, resulting in a **15% efficiency gain**.

TECHNICAL SKILLS

Programming: Python, R, SQL, SAS, SSRS, VBA
Tools: Power BI, Tableau, ArcGIS, Amazon QuickSight, Excel (Power Pivot, Power Query, Pivot Table), Google Data Studio
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

Stock Price Analysis & Forecasting **Jul 2025 - Dec 2024**
- Built a document ingestion pipeline for PDFs files using LangChain, HuggingFace embeddings, and semantic search, ensuring data governance and 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.

MLB Player & Team Performance Dashboard **Jan 2025 - Apr 2025**
- Built **interactive Tableau dashboards** tracking 500+ players and 30 teams with metrics like batting average, RBIs, and strikeouts.
- Implemented predictive models (**Python, Numpy and XGboost)** achieving **85%** accuracy in forecasting player performance and game outcomes.
- Conducted **cohort and trend analysis**, identifying high-impact players and improving **data-driven decision** for team strategies.