

VENKATA HARSHA

Data Analyst — Hyderabad, India — 2026 Graduate

+91-7013445796 — venkataharsha2809@gmail.com — [LinkedIn](#) — [Portfolio](#)

PROFESSIONAL SUMMARY

B.Com (Honors) 2026 graduate with hands-on experience analyzing multi-source datasets across government, health-care, and climate-economic domains. Proficient in SQL, Excel, and Power BI, with a commerce foundation that brings practical domain knowledge in financial and business data. Seeking entry-level Data Analyst roles involving data cleaning, reporting, and dashboard-driven decision support.

EDUCATION

Bachelor of Commerce (Honors)

Keshav Memorial Institute of Commerce and Sciences, Hyderabad

2023 – 2026

CGPA: 8.43 / 10

SKILLS

SQL & Databases	MySQL 8.0 – CTEs, Aggregate Functions, CASE, Joins, Data Cleaning & Validation
Spreadsheets	Microsoft Excel (Pivot Tables, VLOOKUP/XLOOKUP, Power Query), Google Sheets (cross-dataset cleaning, data standardization)
Visualization	Power BI (Dashboard Design, DAX, CALCULATE, ALL(), Data Validation), ECharts (Interactive Web Visualization)
Analytics	Data Cleaning & Validation, Exploratory Analysis, Pattern & Trend Analysis, Statistical Modeling, Reporting
Domain Knowledge	Financial Statement Fundamentals, Variance Analysis, MIS Reporting, Budgeting Basics, BFSI Data Concepts
Analytical Tools	RapidMiner Studio – Classification Models, 10-Fold Cross-Validation, Model Evaluation (AUC, Accuracy)

PROJECTS

Pacific at Risk – Pacific Dataviz Challenge 2026 — [Live GitHub Pages Site](#)

Jul 2026

Tools: Power BI (DAX) • ECharts • Google Sheets • GitHub Pages — Domain: Climate Economics / Data Visualization

- **Integrated and cleaned 7 heterogeneous datasets** (climate, economic, agricultural) covering Pacific Island nations using Google Sheets; resolved country name mismatches and unit inconsistencies to enable reliable cross-source comparison.
- **Built an interactive Power BI dashboard** with a composite vulnerability index; diagnosed scores exceeding 100 to a DAX normalization error involving CALCULATE and ALL(), corrected the logic, and validated outputs.
- **Published findings as an interactive ECharts site** on GitHub Pages, translating multi-source analytical results into accessible visual insights for a non-technical audience.

Heart Disease Risk Detection – Healthcare Analytics — [Read Case Study](#)

May 2026

Tools: RapidMiner Studio • Microsoft Excel • CDC BRFSS Dataset (50,000 records) — Domain: Healthcare Analytics

- **Cleaned and validated a 50,000-record CDC BRFSS dataset** across 19 features before evaluating Logistic Regression, Decision Tree, and Random Forest models using 10-fold cross-validation; Random Forest achieved **90.98% accuracy and 0.824 AUC**.
- **Translated model outputs into a risk-tiered dataset** (Low / Moderate / High / Critical) with tailored recommendations, demonstrating end-to-end ability to move from raw data to actionable analytical deliverables.

District Priority Index – UIDAI Data Hackathon 2026 — [View Hackathon Submission](#)

Jan 2026

Tools: Microsoft Excel — Domain: Government Data / Enrollment Analytics

- **Engineered a weighted District Priority Index** by collecting, cleaning, and aggregating 500,000+ Aadhaar enrollment and biometric records across district-level data sources.
- **Compared enrollment and update-rate patterns across districts** to identify priority areas and generate data-driven insights for a national-level hackathon submission.
- **Delivered a structured analytical report** containing visual summaries and district-level findings.

CERTIFICATIONS

- **Data Analysis Job Simulation – Deloitte (Forage):** Cleaned and analyzed client datasets, built visual dashboards, and communicated findings through structured analytical reports.
- **OTBI Workshop Certification – Digitalization of Villages:** Recognized for leading participant onboarding and technical facilitation at a state-level rural digitalization workshop.