V Balamurugan

Aspiring Data Engineer | B.Tech CSE (Data Science – AI)

Perambur, Chennai | 9080217416 | bm20061027@gmail.com | linkedin.com/in/bala-murugan-a41007374 | github.com/balamurugangethub

SUMMARY

Third-year CSE (Data Science – AI) student with strong algorithmic thinking and SQL/Python fundamentals, hands-on experience with IBM Cloud and Cognos Analytics, and two deployed projects including a data pipeline built with Airflow and dbt. Seeking a data engineering internship to build reliable data pipelines and grow into a full-time data engineer.

EDUCATION

B.Tech, Computer Science & Engineering (Data Science – AI) 2024 – 2028

Dr. M.G.R. Educational and Research Institute, Maduravoyal, Chennai | CGPA: 8.17 (after Semester 3)

- Industry-integrated programme with IBM: a dedicated IBM subject every semester, taught by IBM trainers, with hands-on labs and a project each semester.
- Worked with IBM Cloud and IBM Cognos Analytics in labs and projects.

PROJECTS

Retail Sales Data Platform – Data engineering pipeline and dashboard 2026

Python, Airflow, dbt, DuckDB, Flask, Docker, GitHub Actions | github.com/balamurugangethub/retail-sales-dashboard

- Built a daily pipeline on 1M+ retail sales rows, orchestrated with Apache Airflow and GitHub Actions, loading raw, staging and marts layers in DuckDB/MotherDuck.
- Modelled star schemas with dbt (incremental fact table, 51 dbt tests); added data-quality validation with quarantine of bad files and Discord/Slack alerts.
- Built a Flask dashboard with KPIs, a SQL workbench and pipeline monitoring; 175+ automated tests, deployed on Vercel.

Sakhi Saathi – Voice-first web app for rural women 2026

JavaScript, Node.js, Gemini API, Web Speech API, PWA, Docker, Cloud Run, GitHub Actions | github.com/balamurugangethub/-sakhi-saathi

- Built for the PromptWars × HackArena hackathon (finished 9th): a voice-first, offline-capable PWA that helps first-time rural women use government services and welfare schemes by voice, in 11 languages (4 hand-written, 7 via Gemini translation).
- Kept the API key server-side on a zero-dependency Node.js server; profile data stays on-device; deployed on Google Cloud Run with CI and 33 automated tests (security, accessibility, eligibility rules).

SKILLS

Strengths	Algorithms and problem-solving, Data Structures & Algorithms (DSA)
Languages (working knowledge)	Python, Java, C, C++, SQL
Data engineering	Apache Airflow, dbt, DuckDB, star-schema modelling, data quality, Docker, GitHub Actions
Building	AI-assisted app development, deploying applications, VS Code and developer tooling
Data & IBM	IBM Cloud, IBM Cognos Analytics, IBM Watson
Design & Office	Canva, Photoshop, DaVinci Resolve, Microsoft Excel, PowerPoint, Word
Spoken	English (fluent), Tamil (native), Spanish (basic)

CERTIFICATIONS

Python for Data Science – NPTEL, IIT Madras (4-week course, proctored exam) Jul – Aug 2026

Technical Support Fundamentals – Google IT Support Professional Certificate, Coursera Feb 2026