

Ganraj Puyad

+91 7249557447 | ganraj.puyad.work@gmail.com

LinkedIn: [linkedin.com/in/ganraj-puyad-data](https://www.linkedin.com/in/ganraj-puyad-data) | GitHub: github.com/g-puyad-patil

PROFESSIONAL SUMMARY

Highly motivated Data & AI Trainee candidate with a strong foundation in Python, SQL, and Machine Learning. Proven ability to engineer Medallion Architecture pipelines and perform complex SQL-based data cleaning on interconnected datasets. Committed to high learning velocity and eager to contribute to Xebia's Data & AI CoE by building scalable, validated data solutions and prototypes.

TECHNICAL SKILLS & COMPETENCIES

- **Programming:** Python (Pandas, NumPy), SQL (PostgreSQL, MySQL), PySpark, Bash/Shell Scripting.
- **Data Engineering:** ETL/ELT Pipelines, Medallion Architecture (Bronze/Silver/Gold), PII Masking, Data Validation.
- **CS Fundamentals:** Data Structures & Algorithms, Operating Systems, Networking basics.
- **AI & Analytics:** Statistics & Probability (Distribution, Correlation), Supervised Learning, GenAI/LLM workflows.
- **Tools & Cloud:** Docker, Git/GitHub, Apache Spark, Tableau, AWS exposure.
- **Soft Skills:** High Learning Velocity, Ownership Mindset, Analytical Problem Solving, and Collaborative Technical Documentation.

DATA & AI PROJECTS

Secure Medallion Data Pipeline (Spark & Python)

- Engineered a multi-stage Medallion Architecture (Bronze, Silver, Gold) to process 10,000+ synthetic records using Python and Apache Spark.
- Implemented automated PII masking using custom PySpark UDFs and SHA-256 hashing to ensure 100% data privacy compliance (GDPR/CCPA).
- Developed robust error handling and logging mechanisms to maintain high data quality and system reliability.

TechMart: Global Retail Analytics & Data Validation (SQL)

- Resolved 100% of data quality issues across 4 interconnected tables by standardizing "nan" entries and non-numeric representations into validated data for reporting.
- Engineered 5+ high-impact analytical views using Joins, CTEs, and Window Functions to track global sales trends and employee productivity.
- Optimized query logic by leveraging Generative AI tools to draft query skeletons and refine complex multi-table joins.

EDUCATION

Vellore Institute of Technology | *B.Tech in Computer Science and Engineering* | 2026

- **Relevant Coursework:** Database Management Systems (DBMS), Data Structures & Algorithms, Applied Statistics & Probability, Machine Learning Basics.
- **Honors:** CGPA: 8.03

CERTIFICATIONS & LEARNING

- **Data Engineering:** Data Engineering Specialization – IBM (Coursera).
- **Advanced Analytics:** Google Advanced Data Analytics & IBM Data Science Professional Certificates.
- **Business Intelligence:** Tableau Business Intelligence Analyst Certification.
- **Active Participant:** HackerRank (SQL) challenges.