

Akash Pandey

akashpandey98717@gmail.com | +91 8368706164 | [LinkedIn](#) | [GitHub](#)

Summary

Final-year B.Tech. student in Computer Science (AI) with hands-on experience building production-grade generative AI systems, ML pipelines, and BI/analytics solutions. Independently built Tensor Veil, a CTGAN-based synthetic data engine, from research through deployment, and an end-to-end SQL + Power BI analytics project on PostgreSQL. Completed the DeepLearning.AI Machine Learning Specialization (Andrew Ng) and NPTEL Data Analytics with Python (Elite + Silver, 79%), and earned the Oracle OCI AI Foundation Associate and Google Cloud AI certifications. Seeking Machine Learning Engineer, Data Scientist, or Data Analyst roles / internships.

Education

KCC Institute of Technology and Management

B.Tech. in Computer Science (AI) — CGPA: 7.1/10

Sept 2023 – 2027

New Delhi, India

Projects

Tensor Veil – Synthetic Data Engine | [GitHub](#)

Jan 2026 – Present

- Built a generative AI system using CTGAN (GAN-based deep learning) to synthesize privacy-safe training data while preserving full statistical fidelity of source datasets.
- Engineered a scalable, production-grade ML pipeline for data ingestion, automated cleaning, missing-value imputation, and hyperparameter optimization — built to run in production rather than as notebook-level scripts.
- Implemented a validation layer (KS test, Total Variation Distance, ML efficacy via train-synthetic-test-real) comparing real vs. synthetic distributions, functioning as an auditable QA system for generated data.
- Set up a Supabase database with row-level security to track experiments and results.
- Deployed as a full-stack Streamlit web application with CI/CD via GitHub Actions and unit tests, and open-sourced on GitHub.

Olist E-Commerce Sales & Delivery Analysis – SQL + Power BI Dashboard | [GitHub](#)

- Built an end-to-end analytics project on ~99,000 orders from the Olist Brazilian e-commerce dataset, using PostgreSQL for data storage/analysis and Power BI for dashboarding and DAX.
- Modeled 7 relational tables (orders, customers, payments, reviews, products, sellers, geolocation, etc.) into a star-schema-style structure in PostgreSQL, including a custom Date dimension table.
- Wrote SQL to answer core business questions on sales performance, customer retention, delivery reliability, and payment behavior, then designed a 4-page interactive Power BI dashboard (Sales Overview, Customer Insights, Delivery & Reviews, Payments) with KPIs, geo maps, and DAX measures.
- Surfaced quantified business insights — e.g., late deliveries drop average review score by nearly 2 points ($4.3 \rightarrow 2.57$), and only ~3% of customers make a repeat purchase, pointing to a clear retention gap.

Top Rated TMDb Movies – Cleaned Dataset | [Kaggle](#)

- Performed exploratory data analysis (EDA) on raw TMDb movie data to identify quality issues, inconsistencies, and missing patterns.
- Cleaned and preprocessed the dataset, handling missing values, removing duplicate records, and standardizing fields for downstream analysis.
- Published the cleaned dataset publicly on Kaggle for use by other practitioners.

Technical Skills

Languages:	Python (primary), C++, Java
ML/DL Frameworks:	Scikit-Learn, TensorFlow, CTGAN
Data Engineering & Analysis:	NumPy, Pandas, Data Cleaning, Exploratory Data Analysis (EDA), Feature Engineering
Statistics & Validation:	Model Evaluation
Databases:	PostgreSQL, Supabase (PostgreSQL-based)
BI & Visualization:	Power BI (DAX, dashboarding), Streamlit
Generative AI & Automation:	Synthetic Data Generation, Prompt Engineering, Claude Code
Tools & Platforms:	Google Cloud Platform (GCP), Git/GitHub
Concepts:	Supervised & Unsupervised Learning, Statistical Modeling, Data Structures & Algorithms

Certifications & Achievements

- Machine Learning Specialization – DeepLearning.AI / Andrew Ng (Coursera): Supervised learning, Unsupervised learning, Advanced algorithms, Neural networks
- Data Analytics with Python – NPTEL (Elite + Silver, Score: 79%)
- Oracle OCI AI Foundation Associate (2025, Score: 85%)
- AI-Debate AI Case Competition by BC of IIT Delhi – Rank #67
- HackerRank – Introduction to Programming Using Python, Certificate of Achievement