

# MITUL SRIVASTAVA

India | +91 95824 80350 | mitul1999s@gmail.com | <https://www.linkedin.com/in/mitul-srivastava/> | https://mitul-srivastava-lqmvvxv.gamma.site/

## Summary

Data Analyst with proven experience in turning data into impactful business insights. Strong analytical acumen and attention to detail with a focus on accuracy, efficiency, and strategic decision support. Dedicated to continuous improvement through evidence-based analysis.

## Technical Skills

- **Core Skills:** Data analysis, data cleaning, data visualization, statistical analysis, predictive modeling, reporting, business insights
- **Databases:** PostgreSQL, MySQL
- **Languages:** Python, R, SQL
- **Libraries & Frameworks:** NumPy, Pandas, scikit-learn, TensorFlow
- **Visualization:** Power BI, Excel
- **Cloud:** Google Cloud Platform (GCP)
- **Tools:** Alteryx, MS Office

## Experience

**Advanced Associate – Assurance (Data Analytics)** 07/2023 to 08/2024  
**EY Global Delivery Services** Noida

- Improved audit data accuracy and reduced review time by validating, reconciling, and analyzing large financial datasets across global engagements.
- Delivered utilization and productivity insights that supported optimized resource allocation and improved chargeability tracking.
- Streamlined documentation and reporting workflows, cutting turnaround time by 10%.
- Resolved client data inconsistencies and delivered audit-ready analytical outputs for US engagements.
- Mentored new analysts to raise data validation quality and accelerate onboarding.

**Data Analyst – Assurance (Data Analytics)** 06/2022 to 02/2023  
**Alchemy Techsol (deputed to EY GDS)** Gurgaon

- Automated data preparation for financial audits, reducing manual effort by ~25% through standardized workflows.
- Built and maintained ETL workflows to extract, clean, and transform ERP financial data for audit testing.
- Strengthened reconciliation and exception testing processes to ensure accuracy of GL, sub-ledger, and payroll data.
- Delivered dashboards and reports that enabled faster anomaly detection and evidence-based audit conclusions.

## Education

**MSc: Data Science** 10/2025  
South East Technological University Ireland  
Recipient — Government of Ireland International Education Scholarship (GOI-IES)

**B.Tech: Computer Science** 08/2021  
Guru Gobind Singh Indraprastha University Greater Noida  
Recipient — Academic Excellence Award

## PROJECTS

## **Machine Learning Algorithms Repository**

- Evaluated 13 models (regression, classification, clustering, deep learning) to address varied prediction and segmentation problems.
- Achieved 5–10% performance gains via hyperparameter tuning, regularization, and feature selection.
- Deployed a logistic regression model on Hugging Face Spaces for real-time inference.

## **AI-Powered Chatbot for Mental Health Support**

- Fine-tuned a conversational model for empathetic, context-aware dialogue and implemented safety/preprocessing pipelines.
- Integrated sentiment detection, CBT-inspired response strategies, and a RAG knowledge layer for informed replies.
- Added voice I/O and validated responses using perplexity and semantic diversity metrics.

## **Big Data Architecture with Docker & Kafka**

- Built a containerized streaming pipeline (Docker + Kafka) simulating data every 2 seconds; automated ingestion via Python.
- Stored processed data in PostgreSQL and surfaced insights through Power BI dashboards for real-time monitoring.
- Implemented automated backups to ensure data reliability.

## **AGI Research Analysis**

- Analyzed hardware, data generation, investment, and research publication trends using R, revealing a 337% rise in GPU performance, 1200% surge in global data, and 1814% growth in AI investments over the decade.
- Web-scraped 500+ AGI-related papers from arXiv, processed and visualized submission trends showing a 1500% increase in AGI research output from 2013 to 2023.
- Combined multi-source datasets and created reproducible visualizations linking compute, data, and funding growth to AGI research acceleration.

## **Earthquake Data Analysis Dashboard**

- Processed and cleaned 5,900+ USGS earthquake records using Python (Pandas) to uncover trends in regional activity, magnitude distribution, and hemispheric frequency.
- Built an interactive Shiny dashboard with Plotly visualizations showcasing the top 10 active regions, maximum magnitudes by location, and station–error correlations.
- Delivered a data-driven analytical tool enabling stakeholders to explore geospatial earthquake patterns and make insight-backed interpretations.