

# DISHA GUPTA

Lucknow | +91 6393738728 | dishaa28118@gmail.com | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

Data Analyst with hands-on experience in data visualization, statistical analysis, and dashboard development using Python, SQL, and Power BI. Passionate about turning data into actionable insights through clean, interactive, and impactful visualizations.

## EXPERIENCE

**Virtual Internship Trainee** – Tata Data Visualisation: Empowering Business with Effective Insights (Forage)

Duration: July 2024 – April 2025 **Key Responsibilities & Achievements:**

- Completed practical business tasks focused on framing scenarios, selecting appropriate visuals, and communicating insights.
- Gained hands-on exposure to data visualization techniques aligned with business goals.
- Developed professional-level deliverables using Power BI and data storytelling strategies.
- Strengthened ability to present insights clearly for strategic decision-making.

## PROJECTS

### Business Health Dashboard

[GitHub](#)

- Engineered an AI-powered **Streamlit** dashboard to automate data analysis, featuring modules for data cleaning, anomaly detection, and custom visualizations using **Plotly**.
- Implemented a PDF report generation feature using **FPDF**, reducing manual reporting time by an estimated **90%**.
- Impact:** Enabled non-technical users to perform complex data analysis, leading to faster and more informed business decisions.

### DaaS\_API (Data Cleaning as a Service)

[GitHub](#)

- Developed a robust **microservice** using **FastAPI** and **Pydantic** to provide automated data cleaning functionalities, including deduplication, missing value imputation (mean, median, mode), and schema validation.
- Impact:** Streamlined data preprocessing workflows, ensuring a **100%** clean and validated dataset for analytics teams and improving the reliability of subsequent models.

### House Price Prediction

[GitHub](#)

- Built and evaluated regression models to predict housing prices based on location, area, and amenities.
- Performed comprehensive **Exploratory data analysis (EDA)** with **Matplotlib** and **Seaborn** to uncover key correlations between features and housing prices.
- Engineered **10+ new features** from the raw dataset, including location-based and amenity-based variables, which improved **model accuracy** by **15%**.
- Impact:** Created a predictive tool that provides accurate housing price estimates, demonstrating end-to-end machine learning project lifecycle management.

## EDUCATION

Bachelor of Technology in Computer Science & Engineering

Shri Ramswaroop Memorial University, Lucknow — Aug 2022-Present CGPA:

7.54 / 10 (Current)

## TECHNICAL SKILLS

**Languages:** Python, SQL, HTML/CSS

**Tools:** Power BI, Excel, Tableau, Streamlit, Jupyter Notebook, Docker, VS Code, Git, GitHub

**Competencies:** Data Cleaning, Data Visualization, Statistical Analysis, Prompt Engineering

**Soft Skills:** Critical Thinking, Communication, Team Collaboration, Attention to Detail, ProblemSolving, Adaptability

## CERTIFICATIONS

- Power BI – Microsoft (Jul 2024)
- Data Science and Business Analytics Workshop – IIT Mandi (March 2025)
- Data Science with Python – Digipodium (May–Jul 2024)
- Deloitte Data Analytics Job Simulation (Forage - April 2025)

## PUBLICATIONS

1. “AI-Driven Early Disease Detection” (*Upcoming*) To be published in **Cognitive IoT in Biomedical Devices: The Next Evolution in Smart Healthcare**, 2025

- Focuses on integrating AI with IoT-enabled biomedical devices for real-time diagnosis and predictive healthcare analytics.
- Chapter highlights models for early detection using medical sensor data, cloud integration, and deep learning pipelines.