

EISHAAN KHATRI

☎ +91-8171380034 ✉ eishaankhatri@gmail.com 🔗 LinkedIn/Eishaan Khatri 🐙 Github/Eishaan Khatri

Professional Summary

Final-year Mathematics & Computing student with a strong quantitative foundation in statistical modeling, operations research, and predictive analytics. Experienced in building end-to-end analytical solutions — from SQL-based data extraction and R/Python modeling to stakeholder-ready Tableau and Power BI dashboards — across financial, industrial, and business domains. Published researcher with a demonstrated ability to synthesize complex findings into structured, client-facing deliverables. Thrives in collaborative, fast-paced environments and brings rigorous problem-solving and clear communication to every engagement.

Education

Rajiv Gandhi Institute of Petroleum Technology, Amethi May 2026 (Expected)
B.Tech in Mathematics & Computing | Minor in Business Consulting CPI: 8.60 | **1st Rank, Mathematics Dept.**

- Relevant Coursework: Operations Research, Optimization Methods, Statistical Inference & Hypothesis Testing, Econometrics, Database Systems & Data Warehousing, Business Analytics, Numerical Methods, Linear Algebra

Delhi Public School, Hapur Apr 2021 - Jun 2022
Class 12 90.67%

Experience

Indian Institute of Management Shillong May 2025 - Present
Research Analyst Intern | Mentor: Dr. Shravan Kumar Trivedi Remote

- Built and optimized **regression and classification models** in Python and R for financial decision systems, improving model accuracy by **15%** through systematic feature engineering and cross-validation
- Applied **LIME and SHAP** explainability techniques to decompose model predictions into key business drivers, enabling non-technical stakeholders to act on model outputs with confidence
- Designed **Tableau dashboards** to communicate analytical findings to senior stakeholders, reducing reporting turnaround and enabling faster data-driven decisions
- Performed **SQL-based data extraction and aggregation** to prepare structured datasets for modeling pipelines; currently authoring a research paper synthesizing findings into a client-ready analytical framework

Polka Labs Pte. Ltd. (Polkasassembly, Singapore) May 2025 - November 2025
AI Developer Intern Remote

- Engineered a **multi-platform data pipeline** with standardized schemas, reducing integration time from 1 week to under 5 hours — a **5× efficiency gain** directly improving operational throughput
- Designed **modular, reproducible data workflows** handling structured outputs across multiple services; applied consistent data formats to ensure downstream analytical reliability
- Collaborated across time zones in a fast-paced product environment, delivering analytical solutions iteratively against evolving client requirements

Projects

Stock Price Forecasting — Reliance Industries (NIFTY 50) 🐙 [Code](#) | Python, R, Excel 2024–2025

- Developed a **gradient boosting regression model** on 1,305 days of daily trading data, achieving **R² of 0.9998** and MAE of 3.48 on a held-out 241-day test set
- Engineered **10+ analytical features** (SMAs, RSI, log-returns, rolling volatility); applied **correlation and multicollinearity analysis** in R to select optimal predictors and eliminate redundant signals
- Performed **residual diagnostics, bias analysis, and time-series cross-validation** to prevent data leakage and simulate real-world deployment conditions
- Delivered a **structured analytical report** with executive summary, methodology, results, limitations, and recommendations — mirroring a consulting client deliverable format

Industrial Sensor Analytics & Anomaly Detection 🐙 [Code](#) | Python, SQL, Power BI, Tableau 2024

- Extracted, cleaned, and transformed **378,000+ rows** of industrial sensor data using **SQL queries and Python pipelines**; resolved error strings, nulls, and temporal inconsistencies for reliable analysis

- Engineered **28 domain-informed features** (pressure drop, heat transfer efficiency, rolling volatility, rate-of-change) improving anomaly detection signal quality by **40% over baseline**
- Detected **20+ significant operational anomaly periods** using Isolation Forest with temporal smoothing; performed **Z-score root-cause analysis** to rank contributing sensor deviations by severity
- Built **Power BI and Tableau dashboards** translating technical findings into prioritized operational recommendations for engineering teams

Pharma Commercial Analytics — Forecasting, Segmentation & A/B Testing [Code](#) | Python, R, SQL, Excel, Tableau, I

- Queried and aggregated **600,000+ pharma sales transactions** spanning 6 years across **57 drug categories and 8 ATC classifications** using **SQL joins, window functions, and GROUP BY aggregations** to build a structured analytical dataset
- Built **linear and ridge regression models in R** to forecast monthly drug sales by category; segmented products into high/mid/low performance tiers using **K-means clustering**, validated with **ANOVA** confirming statistically significant differences across tiers at **$p < 0.05$**
- Conducted **A/B test analysis** on seasonal campaign periods using **two-sample t-tests and chi-square tests**, identifying a **statistically significant sales uplift** at **95% confidence** — synthesized into a go/no-go recommendation for campaign rollout
- Delivered findings via **Tableau dashboard and Power BI report** with forecast vs. actuals and segment profiles; summarized in an **Excel pivot model with VLOOKUP-driven scenario tables** enabling category managers to act on underperforming drug lines

Publication

Target Based Offensive Language Detection in Hinglish [Link](#) | *Expert Systems with Applications, Vol. 308, May 2026*

- Peer-reviewed publication in a **top-tier Elsevier journal**; self-annotated a dataset of **6,800+ instances** and developed a novel token-level annotation framework achieving high accuracy across pre-trained models

Achievements

- **1st Rank** in the Mathematics Department at RGIPT
- **Finalist** — National Policy Hackathon 2024 by **IIM Calcutta & IIFT Kolkata**
- **Finalist** — FloodAI Hackathon by **IIEC IIT Gandhinagar**
- **JEE Advanced 2022: Qualified** — Top 1.5 percentile among **1.2 million candidates**

Technical Skills

- **Analytics & Modeling:** Regression, Time-Series Forecasting, Classification, Clustering, Hypothesis Testing (t-test, chi-square, ANOVA), A/B Testing, Anomaly Detection, Operations Research, Explainable AI (LIME, SHAP)
- **Programming & Querying:** Python, R, SQL (joins, window functions, aggregations), C++
- **Data & Visualization:** Excel (Pivot Tables, VLOOKUP, Financial Modeling), Tableau, Power BI, Snowflake, Data Warehousing, Apache Spark, AWS
- **Libraries & Tools:** Scikit-learn, PyTorch, TensorFlow, HuggingFace, Pandas, NumPy, Git

Leadership & Service

- **Overall Management Head** — OWASP, RGIPT: Led cross-functional team initiatives and guided **500+ freshmen** through structured orientation programs
- **Academic Tutor:** Delivered structured tutoring in **Numerical Methods** and Biology to junior students; recognized for clarity of explanation and student outcomes