

Shubh Mehta

United States | +1-812-803-4508 | shubh9mehta@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

Data Analyst who solves operational headaches with analytics. Automated inventory tracking systems that catch supply chain issues before they become costly problems. Uncovered hidden patterns in user behavior that redirected product strategy. Built predictive models that replaced guesswork with data-driven planning. Master's in Data Science from Indiana University with impact across hospitality tech, travel platforms, and academic research. Specialized in SQL, Python, Tableau/Power BI, modern cloud platforms, and making complex data ridiculously clear.

TECHNICAL SKILLS

- **Analytics:** Data Cleaning, Statistical Modeling, Forecasting, Feature Engineering, A/B Testing, ETL Pipelines
- **Programming:** Python (pandas, scikit-learn, NumPy), SQL (T-SQL, PostgreSQL), R, Stata
- **Databases & Tools:** SQL Server, MySQL, MongoDB, BigQuery, AWS S3/Redshift, Azure Data Factory, Google Cloud Platform, Git, Docker, REST APIs
- **BI/Visualization:** Tableau, Power BI, Looker Studio, Excel (PivotTables, VLOOKUP, Power Query), Google Analytics (GA4)
- **Business Skills:** Business Intelligence, Dashboard Development, Cross-Functional Collaboration, Impact Analysis, Stakeholder Reporting

EDUCATION

- Indiana University Bloomington

Master of Science, Data Science
- Aug 2023 - May 2025
- SRM Institute of Science & Technology

Bachelor of Technology, Computer Science & Engineering
- Jul 2019 - May 2023

WORK EXPERIENCE

- Macrotech Corporation

Data Analyst
- Aug 2025 - Present
- Irving, TX
- Reduced inventory mismatches by 24% by automating reconciliation across 400+ SKUs using **SQL** queries and **Excel** functions (**VLOOKUP**, **SUMIFS**, **PivotTables**), improving warehouse audit accuracy.
 - Improved demand forecasting accuracy by 21% building **Power BI** dashboards and weekly **SQL**-driven reports from **CRM** data, reducing stockouts and optimizing procurement planning.
 - Reduced **CRM** data entry errors by 18% by standardizing input protocols with warehouse and IT teams, improving downstream reporting reliability and audit accuracy.

- CruiseAlly

Growth Data Analyst
- Jun 2025 - Aug 2025
- Bloomington, IN
- Uncovered 37% checkout drop-off across 1,000+ user sessions, by analyzing **GA4** funnel data in **Looker Studio** to prioritize product roadmap fixes.
 - Decreased 0-result searches by 28%, improving failure rate(40% to 28.8%) by building a fallback logic in **Django/GraphQL** for alternative cruise recommendations.
 - Conducted comprehensive **SEO** audit analyzing 200+ URLs, removing broken pages and proposing **JSON-LD breadcrumb schema** implementation to strengthen search engine crawlability.
 - Supported product roadmap prioritization by delivering **GA4 funnel analysis** and user journey insights via Looker Studio dashboards used by product, marketing, and growth teams.

- Indiana University Research Data Commons

NLP Research Assistant
- Jan 2025 - May 2025
- Bloomington, IN
- Reduced information retrieval time by 90% (from 5 mins to <30 secs) by deploying a **Flask**-based **AI** assistant using **GPT-4o** and semantic search for 10,000+ records.
 - Automated the ingestion of 15+ disparate data sources (**Excel/Word**) to a structured **JSON** knowledge base, eliminating manual data compilation tasks for staff.
 - Optimized search accuracy by implementing **SentenceTransformers (all-MiniLM-L6-v2)** and **cosine similarity** to enable contextual, meaning-based queries across research archives.

- Kelley School of Business, Indiana University

Data Analyst Research Assistant
- May 2024 - Aug 2024
- Bloomington, IN
- Delivered 94% dataset match rate, successfully linking 21,000 of 22,383 IPO records to blockholder data, by applying **Python** fuzzy matching and **Stata** techniques for layoff analysis.
 - Revealed 17% reduced layoff probability in founder-led firms via **logistic regression** outputs, by incorporating FounderCEO status and blockholder variables into models.
 - Shortened processing runtime by 38%, from 2.5 hours to 1.55 hours, by optimizing **Stata** scripts with **vectorized operations** on large financial datasets.

- TeknoBiz Solutions Pvt Ltd.

Data Analyst Intern
- May 2022 - Sep 2022
- India
- Enhanced **sales forecasting** accuracy by 21%, improving **R²** from 0.65 to 0.78, by developing **Random Forest** and **Gradient Boosting** models in **SAS/Python** on 100,000+ records.
 - Streamlined manual data prep by 29%, reducing monthly reporting time from 21 to 15 hours, by implementing **ETL** pipelines with **T-SQL** for sales consolidation.
 - Lowered excess inventory by 14%, adjusting stock from 120% to 103% of demand, by building **Tableau** dashboards highlighting regional performance and **segmentation** insights.

ACADEMIC PROJECTS

- NYC Taxi Revenue & Operations Analytics (Cloud + BI)
- Built end-to-end **ETL** pipeline using **Mage-AI** on **GCP** processing 150,000+ trip records into **BigQuery**, creating automated **Looker Studio** dashboards tracking revenue patterns and operational metrics.
 - Designed **fact-dimension** data warehouse schema identifying peak revenue hours (3-7 AM) and high-value pickup zones, uncovering opportunities for optimized driver dispatch and airport route prioritization.
 - Automated dashboard refresh using **Google Cloud Scheduler** enabling real-time visibility into payment trends (82% credit card adoption) and \$4M+ transaction analysis.
- Sales Data Analysis & Visualization for Computer Hardware Company
- Analyzed multi-dimensional sales and financial data using **SQL/Excel**, building **Tableau** dashboards visualizing regional performance and post-sales profit margins to support strategic planning.
 - Created interactive dashboards with **filters** and **drill-downs** enabling stakeholders to explore revenue trends, identify underperforming regions, and assess profitability drivers.
 - Structured datasets for reusable reporting framework supporting ongoing **sales analysis** and executive decision-making.