

Tejas Mistry

tejas2000930@gmail.com | (315) 849-8584 | linkedin.com/in/tamistry/ | github.com/tamistry30

Education

Syracuse University, School of Information Studies

May 2025

Master of Science in Applied Data Science

University of Mumbai, Lokmanya Tilak College of Engineering

May 2022

Bachelor of Engineering in Computer Engineering

Technical Skills

Languages: Python, SQL, R, JavaScript

Quantitative Modeling: Scikit-learn, TensorFlow, XGBoost, ARIMA, LSTM, A/B testing, Logistic Regression, K-Means Clustering

Big Data & Platforms: PySpark, Apache Spark, Databricks, Redshift, Hadoop, dbt

Cloud & Tools: AWS, Azure, GCP, Snowflake, Docker, GitHub, JupyterLab

Visualization: Tableau, Power BI, Looker Studio, Google Analytics

Other: Excel(Pivot Tables, VLOOKUP/XLOOKUP, Conditional Formatting), Power Apps, GitHub Codespaces, Adobe Illustrator

Experience

Research Analyst – Data & Behavioral Insights, Syracuse University, Syracuse, NY

July 2025 - Present

- Engineered ETL workflows with robust schema design that cut manual processing by 40% and improved data consistency for 5+ faculty teams, **reducing publication data errors and speeding research turnaround.**
- Built Matplotlib/Seaborn/Plotly visuals used in **publication figures, grant proposals, and curriculum insights**, improving faculty decision-making.
- Standardized descriptive statistics workflows in Python (Pandas/Polars) to ensure **reproducible cross-library metrics**, enabling **faster experiment iterations and quicker validation cycles**, resulting in approximately 35% faster analysis.
- Optimized LLM prompt workflows on structured sports data using **OpenAI/Hugging Face + LangChain**, yielding **+25% validation accuracy** and **more reliable faculty analytics use cases.**

Data Scientist Intern, Bandhouse Music Group Co, Nashville, TN

May 2024 – August 2024

- Developed interactive Tableau dashboards with Spotify and YouTube API integrations, providing real-time insights that improved stakeholder decision-making and **drove a 10% lift in streaming performance.**
- Built **Python and SQL pipelines (Pandas, API clients)** to ingest and clean Spotify, YouTube, and Instagram engagement data, **improving data reliability, reducing manual effort by 20%**, and orchestrating workflows with **Airflow and dbt.**
- Led exploratory data analysis and statistical testing (including **A/B testing, hypothesis testing, and confidence intervals**) to identify audience behavior trends, achieving a **15% uplift in user engagement.**
- Implemented audience segmentation models using Logistic Regression, K-Means, **validated with ROC-AUC, accuracy, and silhouette score**, enabling targeted marketing campaigns that boosted engagement.

Data Analyst, Aromagasm Cafe Pvt. Ltd, Mumbai, India

May 2022 – July 2023

- Automated order and inventory management workflows using Python and SQL, enhancing efficiency by 30%, **reducing stockouts, limiting over-ordering, and lowering costs** while standardizing data capture across the cafe.
- Streamlined bi-weekly Excel reports (pivots/charts; VLOOKUP/XLOOKUP; conditional formatting) to reconcile sales vs. inventory for **accurate, repeatable business reporting.**
- Modeled customer satisfaction drivers to inform operations, yielding **+25% service-quality improvement** and higher retention metrics.
- Partnered with marketing and operations teams to deliver actionable insights that **optimized campaign performance, strengthened customer engagement.**

Projects

Airfare Prediction and Optimization, Syracuse University

September 2024 – December 2024

- Collaborated in a team of 3 to analyze airfare pricing patterns using big data from flight routes and seasons (S/T). Built an end-to-end ML pipeline in PySpark with feature engineering and models (Gradient Boosting, Random Forest) to predict optimal pricing. Delivered **$R^2 = 0.99$ accuracy with RMSE = 12.6**, enabling precise revenue-optimized pricing.
- Technologies used:** PySpark, Python, Jupyter, AWS S3, MLlib, Pandas, Gradient Boosting, Random Forest

vBay Auction Data Warehouse & Business Intelligence, Syracuse University

January 2025 – May 2025

- Architected and deployed a Snowflake star-schema data warehouse and interactive Power BI dashboards for an auction platform, **cutting query response times and unlocking real-time insights into bidding trends, category performance, and regional engagement** to optimize pricing strategies.
- Technologies used:** Snowflake, SQL, Power BI, Python, Excel