

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

June 2022

Bachelor of Engineering in Computer Engineering

Technical Skills

Languages: Python, SQL, R, JavaScript

Quantitative Modeling: Scikit-learn, TensorFlow, XGBoost, ARIMA, LSTM, Prophet

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

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

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

Other: Excel, Power Apps, GitHub Codespaces, Adobe Illustrator

Experience

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

July 2025 - Present

- Designed and executed descriptive statistics workflows on large-scale datasets using pure Python, Pandas, and Polars, enabling reproducible cross-library metric comparisons and reducing analysis time by 35%.
- Automated ETL pipelines to clean, transform, and aggregate structured datasets, cutting manual processing by 40% and improving accessibility for 5+ faculty research teams.
- Engineered and tested prompt-based queries for Large Language Models (LLMs) on structured sports datasets, improving model accuracy validation by 25% through statistical benchmarking.
- Created statistical reports and visualizations using Matplotlib, Seaborn, and Plotly, enhancing data storytelling and contributing to 3 published research presentations.

Data Scientist Intern, Bandhouse Music Group, Nashville, TN

May 2024 – August 2024

- Built and automated data pipelines in Python to collect and clean high-volume engagement data from Spotify, YouTube, and Instagram, improving data reliability and reducing manual effort by 20%.
- Performed EDA and statistical analysis using hypothesis testing and confidence intervals to uncover audience behavior trends, contributing to a 15% boost in user engagement.
- Designed interactive Tableau dashboards to visualize campaign performance, genre trends, and fan segmentation, enabling real-time strategic insights.
- Presented actionable insights and trend analyses to cross-functional teams, contributing to a 10% growth in streaming performance through data-backed recommendations.

Data Analyst Intern, Aromagasms Cafe, Mumbai, India

February 2023 – May 2023

- Automated order and inventory management processes, increasing operational efficiency by 30% and standardizing data collection methods.
- Identified customer satisfaction drivers via statistical modeling, leading to a 25% boost in service quality and improved retention metrics.
- Utilized SQL and Python scripts for data cleaning and automation, ensuring consistent and accurate data reporting.
- Delivered weekly marketing insights that facilitated data-driven decisions and contributed to enhanced campaign performance.

Capstone 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. Achieved $R^2 = 0.99$ and RMSE = 12.6, enabling revenue-focused dynamic pricing strategies (R).
- Technologies used:** PySpark, Python, Jupyter, MLlib, Pandas, Gradient Boosting, Random Forest

vBay Auction Data Warehouse & Business Intelligence, Syracuse University

January 2025 – May 2025

- Designed and implemented a Snowflake-based dimensional data warehouse and interactive Power BI dashboard for an eBay-like auction platform, enabling analysis of bidding trends, category performance, and regional engagement to inform pricing and auction strategies.
- Technologies used:** Snowflake, SQL, Power BI, Python, Excel