Tyler Katz

tylerkatz110@gmail.com | (908) 731-2663 | linkedin.com/in/tylerkatz1 | Portfolio | GitHub

EDUCATION

Syracuse University, School of Information Studies   **Syracuse, NY**
Bachelor of Science in Applied Data Analytics   *May 2027*
- **GPA:** 3.9/4.0; Dean's List (Fall 2023, Spring 2024, Fall 2024, Spring 2025)

WORK & LEADERSHIP EXPERIENCE

Cuse AI   **Syracuse, NY**
Engineer Lead   *September 2025 - Present*
- Scope AI/ML projects with student teams, translate ideas into milestones, unblock executions, and keep deliverables on track
- Act as the technical go-to, answer questions on Python, model training/deployment, and review designs to raise code quality
- Maintain the Cuse AI GitHub organization, structure repos, manage CI/CD, and enforce engineering best practices

Syracuse University   **Syracuse, NY**
Undergraduate Data Analytics Research Assistant   *January 2025 - Present*
- Engineer clean, integrated datasets from raw census files, geospatial shapefiles, and web-scraped sources using R, streamlining data preparation workflows and improving analysis efficiency for the project
- Design and implement data visualizations, geospatial maps, and statistical tables that were directly incorporated into multiple chapters of *Machine [Gun] Politics*, enhancing the clarity and persuasiveness of research findings
- Translate complex, multi-source datasets into actionable insights through advanced data wrangling, exploratory analysis, and visualization techniques, driving progress toward book completion and supporting evidence-based conclusions

Nexis Student Technology Lab   **Syracuse, NY**
Data Science Research Team Member   *September 2024 - Present*
- Collaborate with a team of 4 researchers to analyze multiple datasets, identifying which provide the most valuable insights for further exploration and analysis
- Leverage R and Python to clean, analyze, and visualize data, creating compelling narratives from the findings
- Develop and present data-driven insights to peers and faculty, showcasing results and potential implications for future research

PROJECTS

Restaurant SQL Database | *SQLite, DB Browser* | GitHub Repository   *August 2025*
- Designed a normalized 10-table SQL database from raw CSVs using conceptual and physical ERDs to define schema, keys, and constraints
- Created a Python script to convert CSV data into SQL insert statements for SQLite
- Wrote advanced SQL queries with joins, CTEs, and views to analyze revenue, customer behavior, and performance

Sales Forecasting Predictions Using XGBoost | *R, Shiny* | GitHub Repository   *July - August 2025*
- Engineered 28+ time-series features, including lag variables, rolling averages, improving RMSE from 1433.87 to 663.47 and MAE from 1081.46 to 438.57
- Applied time-series cross-validation and hyperparameter tuning to optimize XGBoost predictions
- Built an interactive Shiny dashboard for store-level sales forecasts over 7–90 days to visualize actual vs. predicted sales

Fake News Detection Using NLP | *Python, Streamlit* | GitHub Repository   *May - July 2025*
- Built a meta-classifier combining Logistic Regression and Random Forest, achieving 95.5% accuracy on 72K+ news articles
- Engineered a text preprocessing and sentiment analysis pipeline in Python using NLTK, TF-IDF vectorization, and parallelized processing (Swifter) to optimize performance on large-scale text data
- Developed an interactive Streamlit web app enabling real-time article classification and confidence scoring

Customer Segmentation Using K-Means Clustering | *Python, PySpark* | GitHub Repository   *March - May 2025*
- Applied K-Means clustering in PySpark to segment 2,240 customers into 3 groups (silhouette score 0.3136)
- Developed three recommendation systems: FP-Growth association rule mining, hybrid cluster–personalized recommendations, and targeted deal suggestions, to inform marketing and inventory strategies
- Engineered key demographic and behavioral features, improving cluster interpretability via PCA

TECHNICAL SKILLS & CERTIFICATIONS

Programming Languages: Python | R | SQL
Software: Microsoft Office | MS SQL Server Management Studio | AWS | Minitab | Git | Tableau | Power BI
Certifications: Microsoft Office Specialist (MOS): Excel Associate | Google Analytics | Advanced Google Analytics