Tengyu Song

(516) 644-3959 • New York, NY • Email • LinkedIn • Portfolio

Education

George Mason University **Sep 2024 –**
PhD in Statistics Fairfax, VA

- Honors: Presidential Scholarship

Columbia University **Sep 2022 – Dec 2023**
Master of Arts in Statistics, GPA: 4.3/4.0 New York, NY

- Honors: Chair's List of Academic Achievement, JSM Award, Winner of 2023 Columbia Data Science Hackathon
- Coursework: Linear Models, Advanced Machine Learning, Information Theory, Bayesian Statistics

Shanghai University of Finance & Economics **Sep 2018 – Jun 2022**
Bachelor of Science in Statistics, GPA: 3.4/4.0 Shanghai, China

- Honors: People's Scholarship (Top 10%)
- Coursework: Probability, Inference, Multivariate Statistics, Programming, Database, Data Mining, NLP

Work Experience

Columbia University Engineering Graduate Student Council **Mar 2023 – Dec 2023**
Data Analyst New York, NY

- Spearheaded an overhaul of the ETL process by integrating Python (pandas, matplotlib) with Tableau to streamline data loading, extraction, and visualization. Enabled easy adaptation for future datasets and improved data reliability.
- Effectively communicated with stakeholders to propose diverse perspectives from data findings. Led to a well-attended town hall and the creation of a publicly acclaimed report.

JD.com **Sep 2021 – Nov 2021**
Data Engineer Intern Shanghai, China

- Developed an AI-driven system that synthesizes user behavior and coupon data to dispense personalized coupons. Spurred a 10% uplift in the company's 7-day Gross Merchandise Volume (GMV).
- Constructed a SQL and PowerBI-based dashboard, presenting key performance indicators in charts, graphs, and tables. Enhanced operational efficiency and empowered stakeholders with the ability to perform incisive trend analysis.
- Led the anomaly detection initiative on R&D process flow data related to deviations from expected behaviors. Utilized ML algorithms such as Isolation Forest to detect anomalies. The model achieved 80% detection accuracy and resulted in a 25% reduction in the average bug response time.

Yum China Holdings **Apr 2021 – Aug 2021**
Data Analyst Intern Shanghai, China

- Orchestrated a project "Cross Analysis of User Acquisition and Retention" for Pizza Hut. Extracted engagement data via SQL and employed statistical analysis (ANOVA, Causal Inference) to identify key factors affecting user behavior. Utilized insights to refine SMS campaigns, boosting retention user numbers by 20%.
- Merged web-scrapped demographic metrics with internal same-store sales figures to forecast new store performance with regression analysis. Guided market expansion decisions and optimizing Return on Investment (ROI) for the franchise.

Skills

Programming Languages: Python (Proficient), R (Proficient), SQL (Proficient), SAS (Proficient), C/C++ (Proficient)

Tools & Frameworks: Pandas, Numpy, Scikit-learn, PyTorch, TensorFlow, PySpark, Tableau, MATLAB, Git, Excel
Statistical Analysis & Data Science: R, SAS, SPSS, Linear Regression, A/B Testing, Causal Inference
Web Development & Scripting: HTML/CSS, JavaScript, VB.NET

Projects and Research

Relationship between Mental Imagery and Perception – Python, R, Statistical Analysis **Sep 2023 – Aug 2024**

- Engineered robust data pipelines for processing and analyzing extensive datasets, including advanced eye-tracking and EEG data using Python and R. Conducted in-depth statistical analysis to uncover correlations between perception and imagination patterns in different tasks, utilizing statistical models (e.g., Logistic Regression, Log-normal Model) and machine learning techniques (e.g., Fuzzy-C Clustering, KAMILA).

Dynamic Pricing for MTA Subway System – Python, Scikit-learn, Machine Learning **Oct 2023**

- Innovated a hybrid dynamic pricing model created from historical turnstile and metro card usage data. Merged Random Forest and ARIMA to predict the future ridership at each station. This strategy can effectively mitigate over-crowdedness and bolstering fare revenue from the simulation conducted.

Coronary Heart Disease (CHD) Risk Prediction Model – SAS, Statistical Modeling **Jan 2022 – May 2022**

- Developed a logistic regression model using SAS to predict 10-year risk of CHD, leveraging a dataset from the Framingham Heart Study including over 4,200 records and 16 health indicators. Applied advanced analytical techniques such as SMOTE sampling for class imbalance and cost-sensitive learning in logistic regression, resulting in improved model accuracy (75.1%) and precision (74.5%). Conducted thorough statistical analyses to identify key risk factors influencing CHD, presenting findings with actionable public health recommendations.

Automatic Defect Detection of PV Cell Panels – Python, OpenCV, PyTorch **Aug 2021 – Dec 2021**

- Performed defect detection related to cracks, discolorations, or irregularities on photographed images of PV cell appearance using a finetuned Mask-RCNN model and an innovating image cropping algorithm, achieving accuracy rate of 98%, surpassing industry benchmarks.

Leadership

Introduction to Statistics **Jan 2023 – May 2023**
Teaching Assistant New York, NY

- Hosted bi-weekly office hours in small group sessions or one-on-one consultations, providing tailored assistance for students' homework and final project, boosted students' performance in the final exam by 5%. Developed detailed lab materials on R data analysis and conducted monthly lab sessions for 80 students, leading to a marked 20% increase in student engagement.

Publications and Presentations

Razzak, E., Mousseau, T., **Song, T.**, & Spagna, A. (2024). The Synergy of Visual Mental Imagery and Working Memory in Problem-Solving. ASSC27, Tokyo.

Liu, Z., **Song, T.**, Chuaqui, K., Tomar, S. S., Chang, K., Baer, N., Frank, R., Liu, J., Bartolomeo, P., & Spagna, A. (2024). Visual mental imagery: An English-language assessment battery for different perceptual and imagery domains with clustered results. Cognitive Neuroscience Society Annual Meeting 2024. Cognitive Neuroscience Society.

Heidenry, Z., **Song, T.**, Holton, T., Chuaqui, K., Chuko, S.-Y., Tien, L., Huang, X., & Spagna, A. (2024). The impact of mental rotation on bistable motion perception lasts longer than you might imagine! Cognitive Neuroscience Society Annual Meeting 2024. Cognitive Neuroscience Society.