Yi Luo

Mathematics, Economics & Data Science

Email: Luoy45802@outlook.com Phone: +01 484-542-2576 Location: Easton, PA

SUMMARY

Data-driven professional combining mathematics, economics, and data science to generate insights that shape research, policy, and decision-making. Skilled at building multi-year datasets, applying advanced statistical/ML models, and creating clear visualizations that uncover patterns in complex data. Strong technical foundation in R, Python, and Stata; effective communicator and collaborative problem-solver.

EXPERIENCE

Research Assistant 01/2025 – 05/2025 — Easton, PA
Lafayette College (Advisor: Sayorn Chin)

- Collected, cleaned, and merged multi-source datasets (state, race, sex, county) using Stata to build a harmonized panel for mortality.
- Designed and produced visualizations showing county-level heart disease death rate trends across time and demographic groups.
- Conducted descriptive and summary statistics in Stata (by age, sex, race) to reveal disparities in mortality patterns across cities and counties.

Research Assistant 12/2024 – 07/2025 — Remote
Columbia University (Advisor: J. Yang)

- Cleaned, imputed, and visualized datasets; engineered features for analysis.
- Trained models in Python (scikit-learn, TensorFlow); evaluated performance using accuracy, precision, and F1.
- Co-authored a research paper presenting empirical findings and methodology.

Financial Analyst Assistant Summer 2023 — Beijing, China
China Galaxy Securities Co., Ltd.

- Conducted ETF and financial product analysis to assess performance and risk.
- Coordinated data collection and reports to support investment decisions.

SELECTED PROJECTS

Decision Trees vs. Random Forests 01/2025–05/2025

- Implemented Decision Trees and Random Forests in **Python (scikit-learn)** to predict resort choice and household income.
- Compared models on accuracy and stability; Random Forest outperformed by reducing overfitting and improving predictive power.

Election Campaign Projections (Python) 09/2024–12/2024

- Processed state-level election data using **Pandas** and **NumPy**; built visualizations with **Matplotlib/Seaborn**.
- Identified swing states and temporal voting shifts through exploratory analysis.

Urban Heat Island Analysis (R) 09/2024–12/2024

- Used **R (sf, raster, ggplot2)** to map building density, land cover, and surface temperature.
- Quantified day–night temperature differences; demonstrated correlation between urban form and heat intensity.

Tumor Growth Modeling 01/2024–05/2024

- Applied **nonlinear regression in R** to fit Gompertz and Logistic models on tumor growth datasets.
- Determined Gompertz model was the better fit based on residual sum of squares and visualized growth curves.

Cancer Risk & Early Smoking 01/2024–05/2024

SKILLS

Data analysis · Statistical modeling · Econometrics · Machine learning · Forecasting · Data visualization · Regression analysis

CODING LANGUAGES

R	★★★★☆
Python	★★★★☆
Stata	★★★★☆
MATLAB	★★★☆☆
LaTeX	★★★★☆

EDUCATION

B.A. in Mathematics & Economics (Minor: Data Science) 09/2022 – 01/2026 — Easton, PA
Lafayette College

- Built **logistic regression models in Stata** to evaluate cancer risk factors (age at initiation, education, smoking intensity).
- Results emphasized higher risks for early initiation, highlighting the importance of early-intervention programs.

Lion & Hyena Population Dynamics (R) 01/2024–05/2024

- Developed **Lotka–Volterra models in R** to simulate predator–prey and interspecies competition.
- Analyzed coexistence, competition, and extinction dynamics within Maasai Mara Reserve.

Economic Indicator Analysis (Iowa) 09/2023–12/2023

- Conducted EDA and feature selection in **R**; imputed missing values and built linear regression models.
- Validated models with RMSE, R^2, and ANOVA F-test to evaluate predictive performance.