

ELENA (YUXIN) LIANG

Phone: (267) 928-5015 | Email: issingliang@gmail.com | [LinkedIn](#) | [GitHub](#)

Address: 1701 Race Street, Philadelphia, PA, 19103

EDUCATION

University of Pennsylvania

MS., Computer & Information Technology

Ongoing

Philadelphia, PA

University of Pennsylvania

MS., Social Policy and Data Analysis, GPA:3.83

Sep. 2022 – May 2024

Philadelphia, PA

Fudan University

LLB., International Politics, GPA:3.41

Sep. 2017 – Aug. 2022

Shanghai, China

Waseda University

BA., Political Science, GPA:3.68

Sep. 2019 – Aug. 2020

Tokyo, Japan

RESEARCH & TECHNICAL EXPERIENCE

University of Pennsylvania

Data Scientist (full-time)

Dec. 2024 – Now

Philadelphia, PA

- **Designed, developed, and launched a webpage** to showcase postdoctoral fellows' research projects. Built the site architecture and layout using **HTML, CSS, and JavaScript**, incorporating **responsive design** and interactive features to ensure accessibility across devices. Utilized **ngrok** for seamless local development and testing, resulting in a robust, user-friendly platform that fosters engagement across the academic community.
- **Applied causal inference models (DiD, IV, PSM, RCT) using R** to evaluate international development programs funded by USAID, performed statistical analysis (e.g., regression, hypothesis testing, causal inference). Developed comprehensive analysis reports and produced well-designed, functional **maps, figures, and tables** to illustrate findings effectively.
- **Built and maintained end-to-end data pipelines** integrating civic datasets (311 service data, GIS layers, AirSage mobility data) using Python (e.g., sklearn, PySpark) to predict illegal dumping incidents in Philadelphia based on 10 years of historical data. **Adapted analytical modules** (bstpp) to align with the specific needs of four key parks identified by stakeholders. Built dashboards and data stories using visualization tools (plotly, seaborn, matplotlib) and guided stakeholders through data-informed decision-making and experimentation outcomes.
- **Supervised a student project leveraging the Gemini LLM API** to classify trash categories, using **PyTorch and TensorFlow for model fine-tuning** via prompt engineering, hyperparameter optimization, and dataset refinement. Evaluated performance using **precision, recall, F1-score, accuracy, and confusion matrix analysis**, applying insights to enhance classification robustness and LLM efficiency.

Centre for Guaranteed Income Research, University of Pennsylvania

Quantitative Research Coordinator (part-time)

Oct. 2022 – Now

Philadelphia, PA

- Executed ETL and managed data pipelines for ad hoc and diagnostic analysis using **Stata and R** (dplyr, lubridate), designing and analyzing **large-scale A/B experiments, cohort analysis, and regression modeling** to support policy evaluation. Drafted reports with R Markdown to communicate research findings and boost policymaking.
- Designed **quantitative survey** and collected data using **Qualtrics**; conducted **in-depth interviews and focus groups** for qualitative analysis. Ensured high-quality data monitoring for metric movement.

Ipsos

Market Analysis Intern

Mar. 2022 – June 2022

Shanghai, China

- Led **A/B testing, user surveys, focus groups, and interviews** to conduct **quantitative and qualitative analysis**.
- **Collaborated with cross-functional teams** to analyze market trends across real estate, beverage, and electronics industries, translating findings into **client reports, whitepapers, and presentations** to guide strategic decisions.

PROJECT EXPERIENCE

Broadway Musical Database and Geo-Analysis | Python & ArcGIS

- Web scraped the IBDB using **BeautifulSoup**, compiling a dataset of **10,000+ shows and 130,000+ theater professionals**, capturing details on show networks, tour stops, and artist relationships to analyze industry trends.
- Designed **ArcGIS StoryMap layouts and interactive maps**, integrating **Google Maps API** to geocode theater locations. Created a **visually engaging web-based geographic analysis** showcasing Broadway's long term expansion and industry patterns: [link](#).

In Progress

Colonial Legacy and Gender Equality in India | R

- Analyzed the impact of British colonial legacy on gender attitudes using USAID's **Demographic and Health Survey**. Applied a **two-stage least squares regression model with an instrumental variable approach**.
- Conducted **data cleaning in Stata**, and modeling in **R** using **dplyr, lfe, spdep, splm**, and **ggplot2** for visualization.

2023 Fall

Social Gender Analysis in U.S. Culture | Python

- Analyzed the evolution of social gender concepts using **HistWords** word embeddings and **applied NLP** techniques such as cosine similarity and vector arithmetic. Utilized **pandas, numpy, sklearn, scipy.spatial**, and **matplotlib** for preprocessing and visualization.

2022 Fall

SKILLS

Programming: Python, R, SQL, Stata, HTML/CSS/JavaScript

Data Science: PySpark, scikit-learn, TensorFlow, PyTorch, ArcGIS/QGIS, Tableau, Power BI

Tools: Git, LaTeX, Adobe Creative Suite, Figma, Microsoft Office

Languages: English (Fluent), Chinese (Native), Spanish(Basic)