SYLVIA LI

CONTACT

+1 617 335 1065

sylvialileyan@gmail.com

https://www.linkedin.com/in/sylvialeyanli/

https://github.com/sylvia010402/Sylvia-Data-Portfolio/tree/main

EDUCATION

Master of Education - Education Policy Analysis

Harvard University (2024 - 2025)

Bachelor of Arts - Economics

University of Manchester (2020 - 2023)

CERTIFICATION

MIT Applied data science program

MIT Professional Education
July 2025

SKILL

Programming Language / Software

R, Python, SQL, Tableau, Excel

Statistical Methods

AB Testing, ANOVA, Data Wrangling, Exploratory Data Analysis, Machine Learning, Regression Analysis, Survey Design, Time Series Analysis

LANGUAGE

English (proficient)
Chinese (proficient)

WORK EXPERIENCE

Program Manager – MetaAI EdTech Evaluation

Sustainable Living Lab (SL2) | Singapore (Remote) | 01/2025 – 05/2025

- Performed data collection from 200+ students and teachers to inform program design, co-led evaluations of an AI-powered education platform across Singapore, India, and Indonesia
- Designed and deployed Qualtrics-based user feedback instruments, cleaned survey data in SQL, for student engagement and learning outcomes to uncover key usage trends and inform future program design
- Visualized quantitative results in Excel and presented to international product and curriculum teams, shaping the platform's expansion roadmap and ensuring culturally responsive design

Policy Research Assistant

Hebei Provincial Department of Urban Development | China | 09/2023 – 06/2024

- Cleaned local fiscal data using SQL and R, identifying investment patterns that increased regional visitor engagement by 15%
- Writing R script for reproducible data pipeline to standardize and visualize municipal funding and revenue data, enabling policy teams to model ¥50 million in budget allocations
- Translated quantitative analytics into briefings and dashboards, presenting for government leaders for urban investment strategy

PORTFOLIO PROJECT

Online Learning Retention Analysis

HarvardX-MITx MOOC Learner Certification Study

- Utilized Python libraries such as Pandas, Numpy and Matplotlib to clean 260,000+ data, built interactive dashboards for behavioral patterns of learners
- Built multiple machine learning models using Scikit-Learn to identify key predictors for course certification, with high model performance (highest F1 Score: 0.73), summarized in clear visuals for the platform teams to guide course design and improve user experience.
- Discovered that learner engagement (e.g., login frequency) outperformed demographic factors in predicting success, shifting platform focus toward modifiable user behaviors.

Education Funding Equity Analysis

Multilevel Modeling with NCES & SEDA Dataset

- Conducted multilevel modeling in R using lme4, analyzing across 10,000+ U.S. school districts to examine how SES and per-pupil spending intersect to shape student achievement.
- Identified interaction effects between district wealth and funding efficacy, guiding more nuanced interpretations of investment equity in K–12 policy.
- Summarized findings in a final report to support grant strategy, district decision-making, and academic publication.