

JANICE LEE

☎ 323-929-3371 | @ jl13491@nyu.edu | 📧 in/janicelee47 | 🌐 lee-janice | 📍 Brooklyn, NY

EDUCATION

New York University

New York, NY

M.S. in Applied Statistics for Social Science Research

Sep 2023 – Present (Exp. Dec 2026)

Relevant coursework: Probability, GLMs, Unsupervised/Supervised Machine Learning, Frequentist & Bayesian Inference, Large Databases, Messy Data & Machine Learning, Causal Inference, Data Science for Social Impact

Pomona College

Claremont, CA

B.A. in Computer Science

Sep 2017 – May 2021

Relevant coursework: Data Structures, Computational Statistics, Stat Theory, Linear Algebra, Operations Research
Honors: Pomona College Llewellyn Bixby Math Prize, Pomona College Scholar

WORK EXPERIENCE

The Research Alliance for NYC Schools at NYU Steinhardt

New York, NY

Data Production Specialist, Full-Time

Sep 2024 – Present

- Secure and de-identify millions of FERPA-protected records & implement ETL pipelines in SAS/SQL to maintain the Research Alliance's archive of 30+ years of data on NYC public schools, as well as project-specific data.
- Manage the Research Alliance's secure Linux server environment with Bash and other command line tools, including maintaining access for 100s of users, installing and updating research software, and managing resource use.
- Spearhead documentation of data, computing environments, security protocols, and programming standards, including leveraging GenAI tools to overcome staffing constraints and update stale information at scale.
- Foster relationships with NYCPs and other external affiliates to facilitate data sharing and research partnerships.

Data Analyst, Full-Time

Nov 2021 – Sep 2024

- Conducted statistical analyses in SAS and created data visualizations in R for an evaluation of New York City's [Computer Science for All \(CS4All\) initiative](#) and an [NYU Discovery Research Fund project](#) on school pushout.
- Assisted with field-based data collection efforts, including distributing thousands of surveys to teachers and students, coding qualitative data, and conducting focus groups, interviews, and classroom observations.
- Contributed to the preparation of reports, presentations, and briefings on findings for a diverse audience of stakeholders, such as policymakers, practitioners, students and parents, and education science researchers.

DoorDash, Geo-Intelligence Team

San Francisco, CA (Remote)

Software Engineering Intern, Part-Time Software Engineer

May 2020 – Jan 2021

- Proposed, architected, implemented, and documented a [discrete-event simulation system](#) written in Kotlin and a React-based web app to automate and scale testing of new location-based services.
- Assisted in the migration of geo-intelligence API endpoints from DoorDash's monolithic Python application to a microservices architecture, while ensuring the integrity of the data pipeline from the PostgreSQL database.

Pomona College, Department of Computer Science

Claremont, CA

Teaching Assistant, Part-Time

Sep 2018 – May 2021

- Mentored students in Introduction to CS in Python/AI (3 semesters), Intro to Languages & Theory (3 semesters).

Georgia Institute of Technology, Department of Mathematics

Atlanta, GA

Undergraduate Student Researcher

May 2019 – July 2019

SKILLS

Languages: Fluent in Python, R, SAS. Proficient in SQL, Java/TypeScript. Familiar with Stata, Java/Kotlin, C/C++.

Technologies: Proficient with Plotly/Dash, Flask, R Shiny, Git, Bash, L^AT_EX. Familiar with React, Neo4j, Docker.

Methodologies: Familiar with GLMs, Bayesian Inference, Causal Inference, Dimensionality Reduction, ML.

Other: Familiar with Qualtrics, Dedoose, Adobe Illustrator, Adobe Photoshop, and Adobe Premiere Pro.

PROJECTS

Interactive Learning Analytics Dashboard

Python, Plotly/Dash, Pandas, Bootstrap

Jan 2024 – Jun 2024

- Architected and developed a multi-page web application using Python, Dash, and Plotly to analyze and visualize student performance data for NYU CAS's Principles of Biology course.
- Engineered a responsive front-end interface leveraging `dash-bootstrap-components` and custom CSS themes to present multi-dimensional student data to non-technical educators.

WikiGraph | wikigraph.netlify.app

React/TypeScript, Python, Neo4j, neovis.js, Netlify

Nov 2022 – Mar 2023

- Processed and uploaded [Wikipedia clickstream data](#) from October 2022 into a Neo4j AuraDB instance.
- Built an interactive web app using React/TypeScript, neovis.js, and Netlify to provide users with a non-linear, graph-based approach to exploring crowd-sourced semantic links between various Wikipedia articles.
- Winner of the Best Dynamic Data Visualization for [NYU's Love Data Week 2023!](#)

NYC Schools Mapping Tool | leejanice.shinyapps.io/nyc-vis

R, RShiny, ggplot2, leaflet

Nov 2021

- Merged, cleaned, and utilized data from NYCPS and DOHMH to create an RShiny application to provide users with a mapping tool to explore the demographics of NYC schools and their geospatial contexts.

Automated Tariff Duty Recovery Pipeline

Python, Google Drive API, Pandas, Regex

Jan 2020 – June 2020

- Implemented a programmatic OCR pipeline using Python and the Google Drive API to ingest, authenticate, and convert thousands of unstructured tariff invoice PDFs into machine-readable plaintext files.
- Extracted fragmented invoice data points, including product descriptions, unit costs, and freight amounts.
- Engineered data transformation logic to flag invoice items excluded from tariffs, calculate exact order quantities, and automate 1% defective allowance modeling.
- Compiled multi-stage data into Excel workbooks to successfully secure thousands of dollars in duty recovery funds.

SELECTED PUBLICATIONS & PRESENTATIONS

Li, X., Jain, R., **Lee, J.** (2024). Computer science course taking and college major declaration: Longitudinal evidence from New York City [Paper presentation]. *American Education Research Association (AERA) 2024 Annual Meeting*.

Fancsali, C., & **Lee, J.** (2024). [Moving the Needle on Equity in Computer Science Education: Lessons from New York City](#). The Learning Curve at Urban Institute.

Li, X., Jain, R., **Lee, J.** (2023). From participation to outcomes: An investigation of the racial and gender disparities in K-12 computer science education in New York City [Paper presentation]. In *Association for Public Policy Analysis and Management (APPAM) 2023 Fall Research Conference*.

Lee, J., Fancsali, C., Clough, S. (2023). Reaching for “All”: Understanding the challenges and needs of schools lagging in CS for all efforts. *SIGCSE Technical Symposium 2023*. <https://doi.org/10.1145/3545945.3569783>

LEADERSHIP & VOLUNTEER EXPERIENCE

SIGCSE Technical Symposium

Reviewer (Panels Track)

New York, NY

Sep 2023, Sep 2024

Community Impact at Columbia University

GED Math and Science Instructor

New York, NY

Jun 2022 – Aug 2022

Pomona College, Asian American Mentorship Program

Student Mentor

Claremont, CA

Apr 2018 – May 2019

Asian Americans Advancing Justice, Health Access Program

Community Partner, as part of coursework on Asian Americans/Social Documentation

Los Angeles, CA

Jan 2018 – Jun 2018