Diana Li

Boston, MA | (302) 803-9517 | li.daij@northeastern.edu | https://dianali495.github.io

I have 5 years of experience in statistical modeling, machine learning, and working with large-scale datasets. I've also managed projects and collaborated with teams to solve complex problems. I am a U.S. citizen; I do not require sponsorship.

SKILLS & CERTIFICATIONS

Technical Skills: Python | SQL | ArcGIS | R | STATA | Tableau | Power BI | JAVA | C++
Analytical Techniques: Data Visualization | Machine Learning | Statistics | Demand Estimation | Causal Inference
Languages: English (Advanced), Mandarin (Native), Français (Beginner A1-A2)

PROFESSIONAL EXPERIENCE

Research Assistant **September 2019 – December 2023**
Northeastern University, Boston, MA

- Processed survey data for development research using Geographic Information System (GIS), producing preliminary results that contribute to a forthcoming publication
- Researched and summarized 50+ consummated merger case details for an antitrust study, contributing to a conference presentation and preparation for future publication

Doctoral Research **September 2019 – December 2024**
Northeastern University, Boston, MA

- Conducted data cleaning and preparation for welfare analysis of the infant formula market, including filtering, labeling, and aggregation of over 25,000,000 observations using Python Pandas and NumPy
- Estimated consumer demand for infant formula across 800 households over a 52-week period, providing behavior insights using STATA and SciPy
- Proposed and evaluated policy changes to improve welfare by 10% through three counterfactual exercises
- Conducted a priceanalysis of airline mergers using 16 quarters of ticket sample data, applying a quasi-experimental (difference-in-difference) design to identify causal effects
- Investigated price impacts across three competition categories, finding a 8% increase on routes with high pre-merger market shares between merging airlines

PROJECTS

Semester Project **February 2019- June 2019**
Math Techniques in Data Science, University of Delaware, Newark, DE

- Performed regression, clustering, random forest, XGBoost, and neural network algorithms using TensorFlow and scikit-learn to make predictions on graduate school admission rate and relevant characteristics
- Achieved 96% prediction accuracy using a dataset of 500 observations

Capstone Project **July 2022 – August 2022**
Data Science for All

- Led a team of 6 members, created a detailed 6-week project plan and assigned deliverables
- Conducted exploratory data analysis using SQL and Python. Translated findings into visualizations using Tableau
- Used causal inference models and predictive techniques to analyze travel patterns and the spray of infectious disease
- Presented findings to a diverse audience of 50+ stakeholders at the Conclusion Symposium

EDUCATION

Northeastern University **December 2024**
Ph.D. in Applied Economics
Fields: Industrial Organization, Labor

University of Delaware **June 2019**
B.S. in Mathematics and Economics
Minors in Computer Science
Coursework: Object Oriented Programming, Data Structure, Numerical Methods, Probability Theory and Statistical Methods, Partial Differential Equation, Fundamentals of Optimization