Thomas S. Lee

U.S. Citizen | 737-224-4010 | leethomas0182@gmail.com | linkedin.com/in/thomasleefin | Austin, TX

SUMMARY

Data scientist transitioning from academia with expertise in Python, SQL, and large-scale data analysis, leading end-to-end data-driven projects using A/B testing, causal inference, and experimental design. Strong foundation in data science analytics with a high agency mindset, applied econometrics expertise, and an interest in extending causal methods to marketing and retail contexts.

EDUCATION

The University of Texas at Austin — Austin, TX
Ph.D. in Finance — (Expected) Aug 2025

Korea University — Seoul, South Korea
B.B.A. | B.S. in Business Administration and Statistics — Feb 2018

EXPERIENCE

BGF Retail — Seoul, South Korea
Project Intern — Dec 2016 – Feb 2017

- Applied data science & analytics expertise by designing a store classification system using customer sales data and machine learning, helping managers quantify the impact of marketing investments through localized product optimization.
- Used unsupervised machine learning (e.g., K-means clustering) to uncover customer segments and support short- and long-term business objectives.
- Presented insights to senior leadership, supporting strategic decision making and cross-functional retail initiatives.

The University of Texas at Austin — Austin, TX
Graduate Researcher — Sep 2018 – Present

- Led multiple end-to-end data-driven projects using experimental design, A/B testing frameworks, and causal inference models (DiD, IV) to evaluate policy impacts.
- Ensured model accuracy and scalability by processing very large datasets (e.g., 110GB+ CRSP, 40GB+ municipal bonds) using Python and SQL.
- Automated data collection pipelines and conducted unstructured data analysis on 1M+ bond features and credit ratings.
- Wrote complex Python and SQL queries for rapid testing and reproducibility, supporting complex codes for quick prototyping.
- Presented research at academic conferences (e.g., FIRS 2025), communicating insights through visualization and demonstrating attention to detail and accuracy in analysis.

Texas Education Agency — Austin, TX
Independent Researcher — Feb 2022 – Present

- Designed and implemented data-driven projects using DiD-based 2SLS IV models to assess the effect of state-led debt relief programs on student outcomes.
- Applied experimental design, data science analytics expertise, and causal inference models to evaluate program efficacy and inform strategic decision making.
- Analyzed longitudinal data for over 160M+ students using Python and Stata, identifying a 0.12 standard deviation increase in math scores from a $1,000/pupil capital increase.

TECHNICAL SKILLS

- *Programming Languages:* Python, SQL, Stata, SAS, R, Excel
- *Quantitative Methods:* Causal inference, A/B testing, Statistics, Experimental design, Panel data analysis, Time-series modeling, Machine learning