Thomas S. Lee

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

SUMMARY

Data scientist and Ph.D. researcher experienced in designing and optimizing data science models, A/B testing, and causal inference using large and complex data sets. Proficient in Python, SQL, and statistical modeling, with a strong record of translating business requirements into ML solutions. Adept at driving data-driven decisions, with a problem-solving mindset and the ability to navigate ambiguity.

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 and analytics expertise to design a store classification system using customer sales data and machine learning (e.g., K-means clustering) to uncover customer segments and support strategic decisions.
- 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 data-driven research projects using experimental design, A/B testing, and causal inference models to evaluate policy impacts.
- Engineered and optimized data pipelines to process 110GB+ of CRSP Mutual Fund and 40GB+ of municipal bond data using Python and SQL.
- Automated data collection and deployed unstructured data analysis techniques to extract insights from 1M+ bond features and credit ratings.
- Presented research findings at academic conferences (e.g., FIRS 2025), emphasizing strong communication skills and clarity in conveying technical content.

Texas Education Agency   Austin, TX
Independent Researcher   Feb 2022 – Present
- Designed advanced econometric models including 2SLS IV and Difference-in-Differences (DiD) to evaluate state-led debt relief programs.
- Conducted large-scale data analysis on longitudinal education datasets (160M+ records), using Python and Stata to derive insights on investment outcomes.
- Applied advanced statistical and machine learning methods to identify impact of \$1,000/pupil capital investment on math score improvements (+0.12 SD).

TECHNICAL SKILLS

- *Programming Languages:* Python, SQL, Stata, SAS, R, Excel
- *Quantitative Methods:* Machine learning, Data science, Optimization models, A/B testing, Causal inference, Statistics, Unstructured data analysis