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. Skilled in A/B testing, causal inference, and statistical modeling to solve business problems and drive impact. Adept at synthesizing insights for strategic decision-making and communicating results to non-technical stakeholders.

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), helping managers improve customer experience through localized product optimization.
- Presented data-driven insights and recommendations to senior leadership, supporting strategic decision-making and cross-functional initiatives across business analytics and product management.
- Presented insights to senior leadership, including the Executive Director, supporting product and inventory strategies.

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 to evaluate policy impacts.
- Analyzed large and complex data sets (e.g., 110GB+ CRSP, 40GB+ municipal bonds) using Python and SQL, ensuring model accuracy and scalability.
- Automated data collection via web scraping and text parsing for over 1M+ bond features and 1.2M+ credit ratings.
- Translated complex analytical findings for non-technical stakeholders through presentations at academic conferences (e.g., FIRS 2025).

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, contributing to strategic thinking and public investment decisions.
- Conducted large-scale data analysis on longitudinal education datasets (160M+ records), using Python and Stata to derive insights on investment outcomes and education policy.
- Applied experimentation techniques to identify impact of $1,000/pupil capital investment on math score improvements (+0.12 SD).

TECHNICAL SKILLS

- *Programming Languages:* Python, SQL, SAS, Stata, R, Excel
- *Quantitative Methods:* Causal Inference (DiD, RDD, IV, Synthetic Controls, Propensity Score Matching), Statistical modeling, Machine learning, A/B testing, Optimization models, Experimentation techniques, Causal inference, Unstructured data analysis