

Thomas S. Lee

U.S. Citizen | 737-224-4010 | leethomas0182@gmail.com | [linkedin.com/in/thomasleefn](https://www.linkedin.com/in/thomasleefn) | Austin, TX

SUMMARY

Data scientist transitioning from academia with a proven track record of delivering actionable insights from large and complex data sets using Python, SQL, and advanced statistical modeling. Skilled in A/B testing, causal inference, and data-centric presentations that influence strategic decision-making.

EDUCATION

The University of Texas at Austin

Ph.D. in Finance

Austin, TX
(Expected) Aug 2025

Korea University

B.B.A. | B.S. in Business Administration and Statistics

Seoul, South Korea
Feb 2018

EXPERIENCE

BGF Retail

Project Intern

Seoul, South Korea
Dec 2016 – Feb 2017

- Led end-to-end data science project to design a store classification system using customer sales data and machine learning (e.g., K-means clustering), helping improve customer experience through localized product optimization.
- Presented data-driven insights and recommendations to senior leadership, supporting strategic decision making and cross-functional initiatives.

The University of Texas at Austin

Graduate Researcher

Austin, TX
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.
- Built statistical models and forecast models based on large and complex data sets (e.g., 110GB+ CRSP, 40GB+ municipal bonds) using Python and SQL.
- Applied unstructured data analysis techniques to extract insights from 1M+ bond features and credit ratings, scaling knowledge and tools across research efforts.
- Communicated analytical findings through clear and compelling presentations at academic conferences (e.g., FIRS 2025), influencing financial and academic direction.
- Mentored peers on data best practices, including analytical processes, reproducibility, and forecasting techniques.

Texas Education Agency

Independent Researcher

Austin, TX
Feb 2022 – Present

- Developed statistical models including 2SLS IV and Difference-in-Differences (DiD) to evaluate revenue drivers and investment outcomes from state-led programs.
- Conducted large-scale data analysis on longitudinal education datasets (160M+ records), using Python and Stata to derive insights on investment outcomes.
- 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, Stata, SAS, R, Excel
- *Quantitative Methods:* Statistical modeling, Machine learning, A/B testing, Optimization models, Experimentation techniques, Forecasting, Predictive modeling, Causal inference, Unstructured data analysis