Thomas S. Lee

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

SUMMARY

Data scientist with expertise in Python, SQL, and large-scale data analysis. Skilled in A/B testing in an industry setting, causal inference, data mining, statistical modeling, machine learning techniques, and experimentation techniques, with experience in descriptive and inferential statistics, clustering, regression, and presenting data to support strategic product decisions.

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), supporting localized product optimization and providing analytical services to client organizations.
- Presented data-driven insights and recommendations to senior leadership, aligning analytics solutions with decision-making and bottom-line contributions.

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 (including 2SLS IV), and causal inference models to evaluate policy impacts.
- Conducted data mining and quantitative analysis on large-scale data sets (e.g., 110GB+ CRSP, 40GB+ municipal bonds) using SQL and Python.
- Applied statistical and mathematical software including R, SAS, and Stata to build models and analyze data.
- Built deep understanding of risk data, reporting, and performance metrics for research purposes.
- Translated complex analytical findings into insights for non-technical stakeholders through academic conference presentations (e.g., FIRS 2025).

Texas Education Agency   Austin, TX
Independent Researcher   Feb 2022 – Present
- Designed and implemented econometric models (e.g., 2SLS IV, Difference-in-Differences) to evaluate state-led debt relief programs, applying advanced statistical techniques and machine learning models.
- Performed large-scale data processing and quantitative analysis on 160M+ longitudinal education records using Python and structured programming.
- Applied experimentation techniques to identify impact of $1,000/pupil capital investment on math score improvements (+0.12 SD), supporting stakeholder decision-making and understanding business value and process before focusing on technical solutions.

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, Unstructured data analysis