# 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 a passion for using data to solve complex business problems. Adept at translating ambiguous business challenges into scalable analytical solutions through clear communication and collaboration. Committed to delivering actionable insights and building sustainable, high-impact analytics products.

## 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.

**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 (e.g., DiD, IV) to evaluate policy impacts.
- Analyzed large and complex data sets (e.g., 110GB+ CRSP, 40GB+ municipal bonds) using Python and SQL, identified and extracted suitable data sources, performed extensive data quality checks, and ensured datasets were fit for advanced modeling and decision-use.
- Applied data governance, management, and quality standards to ensure data accessibility, scalability, and relevance for decision-making.
- Built and evaluated machine learning models including classification and regression techniques (e.g., logistic regression,)
- Automated data collection via web scraping and text parsing for over 1M+ bond features and 1.2M+ credit ratings.
- Provided actionable insights and communicated analytical findings effectively to technical and non-technical audiences at 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:* Classification and Regression, Forecasting, Causal Inference (DiD, RDD, IV, Synthetic Controls, PSM), A/B Testing, Experimental Design, Data Governance & Management, Data Acquisition & Quality Checks, Unstructured Data Analysis, Web Scraping, Model Evaluation (ROC, RMSE, Chi-square)