

# Jooho Kim

✉ jooho991122 (at) gmail (dot) com    🏠 [Personal Website](#)    🐙 [GitHub](#)

## Education

|   |   |
|---|---|
| <b>Seoul National University</b><br><i>MS in Statistics</i>                                     | Seoul, South Korea<br>Mar. 2024 – Feb. 2026   |
| <b>Korea University</b><br><i>BE in Food and Resource Economics, Double major in Statistics</i> | Seoul, South Korea<br>Mar. 2018 – Feb. 2024   |
| <b>The University of Texas at Austin</b><br><i>Exchange Program, Economics</i>                  | Texas, United States<br>Aug. 2022 – Dec. 2022 |

## Research Interests

Missing Data, Survival Analysis, Causal Inference, Statistical Uncertainty Quantification in AI

## Preprints

**Kim, J.**, Saegusa, T., and Shin, Y. E. (2025). “Scalable and Robust Multiple Imputation for Case-Cohort Studies via Influence Function-Based Supersampling.” *Preprint*. [\[arXiv:xxxxx\]](#) [🔗](#)

## Research Experience

|  |  |
|--|--|
| <b>Prediction Model Lab, Seoul National University</b><br><i>Graduate Researcher (Advisor: Dr. Yei Eun Shin)</i>   | Seoul, South Korea<br>Jun. 2024 – Present      |
| <ul style="list-style-type: none"> <li>◦ Led as the primary graduate researcher on a project funded by the National Research Foundation of Korea: “Multiple Imputation for Missing Covariates due to Epidemiological Cohort Sampling Designs”.</li> <li>◦ Proposed an influence function-based supersampling approach to impute only a subset (e.g., 3%) of the missing covariates while preserving efficiency and unbiasedness.</li> <li>◦ Devised weight calibration equations that reconcile heterogeneous sampling weights for a unified Cox proportional hazards regression analysis.</li> <li>◦ Applied the proposed method to NIH-AARP Diet and Health Study, reducing the imputation time by approximately 95% in the SMC-FCS method.</li> </ul> |  |
| <b>Urban Informatics Lab, The University of Texas at Austin</b><br><i>Undergraduate Research Assistant (Connected through Dr. Arya Farahi)</i>   | Austin, United States<br>Oct. 2022 – Dec. 2022 |
| <ul style="list-style-type: none"> <li>◦ Aggregated and cleaned geotagged electric vehicle tweets using regular expressions to handle misspellings and variant notations of the U.S. states.</li> <li>◦ Conducted hotspot analysis across the U.S. to identify regions with significant EV-related public sentiment.</li> </ul>  |  |

## Presentation

|   |           |
|---|-----------|
| <b>Scalable and Robust Multiple Imputation for Case-Cohort Studies via Influence Function-Based Supersampling</b><br>Korean Statistical Society, Seoul, Korea (Oral Presentation).                  | Dec. 2025 |
| <b>Multiple Imputation for Incomplete Survival Data with Missing Covariates: Toward Valid Causal Inference</b><br>The 2nd Symposium on Causal Inference, Seoul, Korea (Oral Presentation, English). | Jun. 2025 |

## Honors and Awards

|   |             |
|---|-------------|
| <b>Next Generation Scholarship for Fundamental Research</b><br><i>Awarded by Seoul National University for academic excellence and research potential</i>             | 2024, 2025  |
| <b>Graduate Research Fellowship in Science and Engineering</b><br><i>Awarded by the National Research Foundation of Korea through a competitive selection process</i> | 2024 – 2025 |

|   |  |
|---|--|
| <b>Special Scholarship</b><br><i>Awarded by Korea University for academic excellence</i>  | Fall 2022, Spring 2023                 |
| <b>Semester High Honors</b><br><i>Recognized by Korea University for academic excellence</i>  | Fall 2018, Spring 2022,<br>Spring 2023 |
| <b>Agricultural Economics Alumni Scholarship</b><br><i>Awarded by Korea University, Department of Food and Resource Economics for academic excellence</i> | Spring 2022                            |



## Teaching Assistantship

---

|  |             |
|--|-------------|
| <b>Survival Data Analysis and Lab</b><br><i>Advanced Undergraduate Course</i>                              | Fall 2025   |
| ◦ Led hands-on lab sessions on survival analysis and graded assignments and exams.                         |             |
| <b>Selected Topics Seminar</b><br><i>Introductory Undergraduate Course</i>                                 | Spring 2025 |
| ◦ Organized weekly discussion sessions on economics and statistics, and advised on data analysis projects. |             |
| <b>Mathematical Statistics 2</b><br><i>Core Undergraduate Course</i>                                       | Fall 2024   |
| ◦ Held office hours, graded assignments and exams, and prepared solution sets.                             |             |
| <b>Statistics Lab</b><br><i>Introductory Undergraduate Course</i>  | Spring 2024 |
| ◦ Evaluated Python programming coursework, and held office hours.  |             |

## Projects

---

|  |                       |
|--|-----------------------|
| <b>Mediation Analysis of Mental-Health Effects on Dietary Outcomes</b>   | Nov. 2025             |
| ◦ Advised on estimating direct and indirect effects of mental-health indicators on dietary outcomes using KNHANES data with SPSS PROCESS and R <code>mediation</code> package. |                       |
| <b>Modeling Risk Factors for Mortality and Hospitalization</b>   | Sep. 2025             |
| ◦ Analyzed risk factors for mortality and hospital stay using GLMM with multiple imputation, addressing repeated events and missing data.                                      |                       |
| <b>Weight Design Project for the Longitudinal Survey Panel</b>   | Oct. 2024             |
| ◦ Constructed stratified sampling weights using the R <code>survey</code> package and advised on missing data handling in the SNU student survey.                              |                       |
| <b>Bitcoin Chart Pattern Image Recognition and Price Prediction Project</b>  | May 2022 – Jul. 2022  |
| <a href="#">Github Repository</a>   |                       |
| ◦ Implemented Monte Carlo Dropout in the N-BEATS time-series neural network to quantify predictive uncertainty and visualize prediction intervals.                             |                       |
| ◦ Augmented chart image data using probability distributions resulting in 10% increase in accuracy.  |                       |
| <b>Optimizing Pricing Strategies for a Low-Demand Food Product</b>   | May 2022 – Jun. 2022  |
| ◦ Designed an online survey and conducted a conjoint analysis to identify consumer preference.   |                       |
| ◦ Developed an algorithm in Python to estimate the profit-maximizing bundle price for the products.  |                       |
| <b>Data Visualization of Job Openings in Korea</b>   | Nov. 2021 – Jan. 2022 |
| <a href="#">Github Repository</a>  (In Korean)  |                       |
| ◦ Extracted and preprocessed 36,000 job postings and 11,000 resumes by identifying html patterns.  |                       |

## Skills & Languages

---

**Software** R, Python, LaTeX, SAS, ArcGIS, Stata, SPSS  
**Languages** Fluent both in English and Korean