

YEH CHAN YOO

Seattle, WA | yehchanwoo@gmail.com | [linkedin.com/in/yehchan-yoo](https://www.linkedin.com/in/yehchan-yoo) | [yehchanwoo.github.io](https://github.com/yehchanwoo)

SUMMARY

Recent Statistics Master's graduate from the University of Washington, possessing expertise in machine learning, statistical modeling, and data engineering, with attention to detail in building analytical models and robust data pipelines. Currently looking for a full-time job primarily in data analytics, data science, or machine learning/artificial intelligence.

SKILLS

- **Relevant Programming Languages:** Python, R, SQL, Kotlin, React/Node.js, HTML, CSS, JavaScript, C/C++, Git, Shell
- **Relevant Platforms:** Jupyter Notebook/JupyterLab, RStudio, Microsoft Office (including Excel), Microsoft Visual Studio, Google Analytics, Amazon Web Services, Google Cloud Platform, Microsoft Azure, GitHub, Android Studio, Tableau

EDUCATION

University of Washington - Seattle

March 2026

Master of Science in Statistics - Advanced Methods and Data Analysis

- **Relevant Coursework:** Advanced Machine Learning (Python), Advanced Regression Methods for Independent Data (R), Advanced Regression Methods for Dependent Data (R), Applied Regression (R), Applied Statistics Capstone (Python), Design and Analysis of Experiments (R), Sample Survey Techniques (R), Statistical Methods for Analysis with Missing Data (R), Statistical Computing (C/C++), Statistical Inference

University of California - Berkeley

December 2023

Bachelor of Arts in Statistics and Political Economy with a Minor in Data Science

- **Relevant Coursework:** Casual Inference (R); Concepts in Computing with Data (R, Git); Concepts of Probability; Concepts of Statistics (R); Data, Inference, and Decisions (Python); Introduction to Artificial Intelligence (AI) (Python); Modern Statistical Prediction and Machine Learning (R); Principles and Techniques of Data Science (Python, SQL)

PROFESSIONAL/RESEARCH EXPERIENCE

Allen Institute for Brain Science • Visiting Student (Machine Learning Engineer) • Seattle, WA

September 2025 - March 2026

- Developed, trained, and validated Bayesian models using Python to perform large-scale single-cell analysis on 500+ GB of data with modern deep learning techniques as well as distributed data processing and scalable analytics on the cloud
- Managed codebase integrity and documentation while executing experiments on AWS and other high-performance computing clusters to ensure scalable model training and evaluation
- Submitted a paper on one of the aforementioned models as the second author to the *Nature Methods* journal

University of Washington iSchool • HCI Researcher/Technical Lead • Seattle, WA

October 2024 - September 2025

- Led the software engineering effort to code and test an Android app product (Kotlin) and a web app product (React/Node.js) for generating emotional messages and stickers through API integration with OpenAI and Google's multimodal AI models
- Contributed to the technical portions of the paper by selecting the right AI models and ensuring that the AI models used would provide safe and applicable results for the research
- Part of a wider research project to test how children aged 6-10 can develop social-emotional skills in a digitalized world; submitted paper as the third author to the Association for Computing Machinery (ACM)

Mindful Conversion • Data Scientist • New York, NY

January 2024 - August 2024

- Served as the primary data scientist for the marketing data analytics agency and startup Mindful Conversion, leading and managing the agency's data science projects in a cross-functional team with both software engineers and marketers
- Developed and maintained a data engineering and ML pipeline for the agency's predictive SEO product Kixely using Python, Microsoft Azure, and Google Cloud -- performing data mining, data processing, and data wrangling on ~100,000 keywords and URLs while engineering and evaluating LLMs to develop predictive models for webpage ranking
- Analyzed and evaluated large and complex data sets using Python and SQL to identify and improve marketing data quality and content strategies for the agency and its clients
- Authored several data-driven reports and presentations that utilized data visualization, technical expertise, knowledge of Google's advertising ecosystem and its long-term product trends as well as strong written and verbal communication skills to present marketing business insights to the agency and its clients

UC Berkeley School of Education • Student (Research) Assistant • Berkeley, CA

June 2023 - December 2023

- Researched, coded, and experimented with deep learning-based natural language processing (NLP) models for automatic short answer grading of 200+ student answers to mathematics and statistics questions; achieved 75% test accuracy with a fine-tuned distilled Transformers-based LLM (RoBERTa) using PyTorch
- Assisted with the experimental design and writing of my lab's statistical assessments

UC Berkeley School of Education • Undergraduate Research Apprentice • Berkeley, CA

January 2023 - May 2023

- Graded and performed statistical analysis on written answers to experimental statistics assessment questions from more than 500 California high school and middle school students using Excel and R, helping researchers refine assessment methodology

SK Wyverns • Intern • Incheon, South Korea

June 2018 - August 2018

- Examined and visualized patterns in thousands of customer records from the South Korean baseball team SK Wyvern's database using Google Analytics, Python (e.g., Pandas, Matplotlib), SQL, and Tableau to identify user sentiment and analyze users' app and website usage patterns
- Built a sentiment analysis NLP model using Python to identify emotions in user feedback