

Howard Baik

(206)-591-4758 | howard.baik@yale.edu | [Blog](#) | [GitHub](#) | [LinkedIn](#)

Education

Oregon State University Postbaccalaureate Degree in Computer Science	Online Sept. 2023-Present
University of Washington Master of Science in Biostatistics	Seattle, WA March, 2023
University of Washington Bachelor of Science in Statistics (Minor: Mathematics)	Seattle, WA June, 2021

Work Experience

Yale School of Public Health Data Scientist I	New Haven, CT Sept. 2024-Present
<ul style="list-style-type: none">◆ Develop and execute statistical models for analyzing EHR and biobank datasets.◆ Create and maintain software packages, as well as offer user support and training in coding and software tools.	
Fred Hutch Data Science Lab Software Development Engineer I	Seattle, WA March 2023-Present
<ul style="list-style-type: none">◆ Built an open-source Shiny app that enables the creation of automated videos from Google Slides or Microsoft PowerPoint.◆ Integrated an open-source text-to-speech generation library, Coqui TTS, into an existing text-to-speech R package for enhanced text-to-speech generation capabilities and improved accessibility.◆ Developed and submitted conrad, an R package for converting text into synthesized speech, to CRAN.	
Merck Contractor	Remote Feb. 2022-March 2023
<ul style="list-style-type: none">◆ Designed internal R packages and a Shiny app for analysis and reporting in a regulatory environment.◆ Produced internal Shiny training material for biostatisticians and statistical programmers.◆ Consulted stakeholders with differing technical backgrounds to meet their requirements for R packages and Shiny apps.	
Merck Biostatistics Intern	Upper Gwynedd, PA June 2022-Aug. 2022
<ul style="list-style-type: none">◆ Created an open-source R package for algorithmically querying ClinicalTrials.gov and building an analysis dataset for drug discovery and development within the broader scientific community.◆ Programmed an interactive Shiny App for Merck users to extract analysis datasets by searching for treatment or condition.◆ Tested two open-source R packages developed at Merck for standard adverse events analysis. (metalite & metalite.ae)	
University of Washington School of Medicine NIH Research Assistant	Remote Feb. 2022-June 2022
<ul style="list-style-type: none">◆ Developed a Shiny Dashboard that allows patients and clinicians in addiction treatment to monitor patients' progress and goals over time.◆ Prototyped a shinyMobile application of the dashboard for optimal viewing on a smartphone.	
NOAA (National Oceanic & Atmospheric Administration) Machine Learning Intern	Remote Jan. 2018-Aug. 2018
<ul style="list-style-type: none">◆ Created an algorithm to automatically detect upwelling, an oceanographic process that has a significant impact on global fishery production.◆ Applied principal component analysis, k-means clustering, and hierarchical clustering to real-world image data to group patterns of upwelling.◆ Documented findings in a blog post that elaborates on the detection algorithm and the significance of visualizations.	
George Mason University Dept. of Computer Science NSF REU Educational Data Mining Research Intern	Fairfax, VA June 2018-Aug.2018
<ul style="list-style-type: none">◆ Analyzed real world datasets from Stanford Lagunita's course, Statistics in Medicine, by creating exploratory visualizations in ggplot2 and employing k-means clustering.◆ Built an interactive web application (shinydashboard) that presents my analysis and visualizations in R.◆ Presented a demo of the dashboard at the 9th International Learning Analytics and Knowledge (LAK) Conference.	