

Zhiyu (Audrey) Wei

206-581-8684 | zw2975@cumc.columbia.edu | New York, NY | www.linkedin.com/in/audreywei

EDUCATION

Columbia University Mailman School of Public Health

M.P.H. Candidate at Department of Environmental Health Sciences
Concentration in Applied Biostatistics (GPA 3.9)

New York, NY

Expected Graduation: May 2025

University of Washington, Seattle

B.S. in Biochemistry, Dean's List for three semesters

Seattle, WA

June 2023

SKILLS

Data analysis: R (data wrangling, statistical analyses & modeling, data visualization), Shiny, SAS (biostatistics), SQL, QGIS
Microsoft Office, quantitative and qualitative analysis, Problem solving, Communication, Time management, Team collaboration.
Language: English, Chinese, Japanese

DATA ANALYSIS & RESEARCH EXPERIENCE

NYC Department of Health and Mental Hygiene: H RTP Intern

New York, NY

HIV and Housing Research Data Analyst

Nov. 2024 – Present

- Analyzed large NYC datasets using R and SAS, applying quantitative and qualitative methods to evaluate disparities in housing for people with HIV, and employing geospatial mapping to reveal patterns in neighborhood housing quality and client distribution based on factors like housing violations, income levels, rent burden, and gun violence.
- Conducted literature reviews to investigate potential indicators affecting housing quality for clients' housing placements, contributing to further improvement in project's data-driven insights and policy recommendations.

SAS Statistical Analysis: Cognitive Performance and Metabolic Syndrome

New York, NY

Statistical Programmer

Dec. 2024

- Conducted statistical analysis using SAS, PROC SQL, and Macro language to explore the relationship between cognitive performance (SRT scores) and metabolic syndrome determinants in a dataset of 267 participants.
- Applied ANOVA, Pearson's correlation, and multiple linear regression to identify significant predictors, including systolic blood pressure and triglycerides, and presented findings in tables, boxplots, and scatterplots, highlighting significant differences in cognitive scores across metabolic syndrome and between sexes.
- Designed and refined predictive models, explaining 9.35% of the variance in SRT scores (adjusted $R^2 = 0.0935$, $p < 0.0001$), while controlling for sex as a confounding variable, incorporating SAS Macros to streamline repetitive analysis processes.

Data Science Webpage Project: Healthy Food Accessibility in NYC

New York, NY

Project Contributor

Nov. 2024 – Dec. 2024

- Conducted a comprehensive data analysis project on food justice in NYC, structuring data cleaning strategies for 8 datasets aligned with the research focus and employing statistical modeling, geospatial analysis, and interactive visualization to derive key insights.
- Investigated relationships between food accessibility, proximity, income levels, and housing prices, delivering a report summarizing key findings, visualizations, and recommendations to address food access disparities and inform urban planning.
- Structured interactive webpage using Shiny to present findings through dynamic maps, charts, and dashboards, improving the accessibility and impact of the analysis for diverse audiences.

Geospatial Analysis Project

New York, NY

Food Scarcity and Health Outcomes

Oct. 2024 – Dec. 2024

- Conducted a geospatial mapping analysis on food scarcity and marginalization, examining their impact on obesity and diabetes rates across 62 New York counties, and presented findings in a poster on geospatial patterns of food scarcity and health equity.
- Analyzed datasets on population income levels, grocery access rates, and health outcomes using regression analysis and Moran's I to evaluate relationships and geographical clustering.
- Created bivariate maps and visualizations to highlight correlations between limited grocery access, income disparities, and higher prevalence of obesity and diabetes.

P30 HEAT SEEK Project: A Pilot Study on Heat Exposure in Northern Manhattan

New York, NY

Research Assistant

Nov. 2023 – Jun. 2024

- Spearheaded data collection efforts for a pioneering study on heat exposure within marginalized communities, demonstrating effective management of research protocols, field visits, participant data through Qualtrics, and follow-up efficacy tracking.
- Led the recruitment process, enrolling 70+ participants in the initial month and maintaining a follow-up rate above 50%.
- Proactively collaborate with the team to use generative AI to brainstorm and implement comprehensive outreach strategies in partnership with local community organizations, significantly broadening the diversity of participant demographics.

PFAS Substance Levels and Total Cholesterol: Environmental Health Analysis

New York, NY

Core Member

Feb. 2024 – May 2024

- Conducted comprehensive statistical analyses on extensive datasets of 5000+ NHANES entries using R, focusing on key variables from publication reviews to eliminate 80% of irrelevant data, enhancing dataset relevance and analysis efficiency.

- Formulated 2 innovative statistical hypotheses for the research study, focusing on three key predictive variables. Utilized combinations of training and testing models and applied statistical methods to identify patterns in epidemiological effect of PFAS substances on human health.

Health Apps' Impact on Fibromyalgia Patients: A Biomedical Engineering Research

Seattle, WA

Research Assistant

Sept. 2022 – Jan. 2023

- Led the validation and testing of statistical results for a major clinical research paper using R, optimizing coding schemes for 30+ medical characteristics and analyzing 600+ survey responses.
- Executed in-depth qualitative and quantitative analyses on the psychological conditions and challenges faced by fibromyalgia patients on social media platforms, yielding significant insights into patient experiences and needs.

Using R to Analyze Video Games on Steam

Seattle, WA

Core Member

Feb. 2022

- Performed sophisticated statistical analysis using R to scrutinize a dataset comprising 50,000+ data sourced from Kaggle, enhancing data wrangling efficiency by 80% through process optimization.
- Directed the creation of a research report on 6 predictive models across three diverse proposals, employing rigorous hypothesis testing to verify model accuracy and reliability, leveraged applied statistical methods to convey data-driven decision making.

PROFESSIONAL & LEADERSHIP EXPERIENCE

Vasco Knight (Evidence Medical Service)

New York, NY

Event Coordinator | Translator

Oct. 2023 – Present

- Facilitated by generative AI, translated, compiled, and presented more than 5 cardiovascular medical journals in Chinese and English, utilizing infographics to campaign for cardiovascular health on social media.
- Coordinated with hybrid VEITH cardiovascular conferences, introduced in-person participants to innovative surgical equipment.

39 Health Net (Medical Information Website)

Seattle, WA

Health Consultant

July 2021- Sept. 2021

- Delivered tailored health recommendations and consulting service to over 60 clients, utilizing specialized knowledge in chronic disease management and preventive care, achieving a 90% positive feedback rate.
- Enhanced medical expertise through rigorous review of 50+ accredited publications, significantly improving consultation efficacy.