Azuka Atum

azuka.at@gmail.com ·
Portfolio · GitHub · LinkedIn ·

SKILLS

Languages and Tools: R, SAS, SQL, basic command line interface (CLI), Microsoft Office Suite (Word, Excel, Powerpoint, Outlook, OneNote, Teams), Google G Suite, Git (github, gitlab, git bash), Database Management Systems (MySQL, Oracle), Unix/Unix-like systems (macOS terminal), Adobe Illustrator, QualtricsXM, ObsidianMD
Statistical Analysis: Report writing (Markdown, Quarto, LaTeX), math skills (statistics/statistical modeling/statistical analysis, Survival Analysis (Kaplan-Meier, Cox Proportional-Hazards, AFT Models), ANOVA, Regression Analysis (linear, logistic, categorical), data visualization
Communication: Presenting data insights to non-technical stakeholders, blogging, graphic design
Research Skills and Interests: Applying biostatistics methods to clinical research, survival models, machine learning with survival models, R Shiny, communicating statistical analysis through visualization, clinical trials, Database Research (NHANES, PubMed), Hypothesis testing

EDUCATION

California State University, East Bay **Hayward, CA**
Master of Science: Biostatistics **May 2023**

- **Center for Student Research (CSR) Scholarship** (2022-2023): Scholarship funding to engage in professional development sessions designed to develop broader academic and professional skills.
- **Selected Coursework**: Probability and Statistical Theory, Regression analysis (Linear/Logistic Regression in SAS and R), Clinical Trials in the Biomedical/biotech Industry, Survival Analysis, Categorical Data Analysis, Applied Deep Learning (tensorflow in R), R for Data Science, Applied Research (AI/artificial intelligence, machine learning)

San Francisco State University **San Francisco, CA**
Bachelor of Science: Microbiology **May 2019**
Capstone: The Effect of Sugary Drinks on the Growth Rate of Cutaneous Staphylococcus sp.

- **National Institutes of Health: SFBUILD** (2018-2019): Participated in competitive research internship emphasizing health disparities, community engagement, and the intersection of public health policy and biomedical sciences.
- **College of Science and Engineering (CoSE) Student Project Showcase** (2018): presented capstone project at university symposium.
- **Selected Coursework**: Microbiology lab (genetic engineering, bacterial isolation), Designing Clinical Research (graduate course), Computational Statistics, Computational Biology

PROFESSIONAL EXPERIENCE

KAISER PERMANENTE **Oakland, CA**
Division of Research
Data Reporting and Analytics Consultant II **Feb 2024 - present**

- Utilized SAS and Oracle SQL within the KPRB Data and Analytic Platform (KDAP) to manage and analyze 50+ million rows of survey, genomic, environmental, clinical (EHR), and biospecimen data from the Kaiser Permanente Biobank, creating de-identified datasets that drove epidemiological research and actionable insights for stakeholders.
- Reduced long-running queries from 8+ hours to 1–4 hours by adopting a modular programming approach for data pulls, carefully monitoring error logs to swiftly identify and resolve issues. Maintained version control of SAS scripts and documentation via internal GitLab, enhancing collaboration, reproducibility, and data integrity across the team.
- Leveraged advanced computing infrastructure and DBMS to streamline data retrieval and ensure operational efficiency through detailed methodology documentation and workflows. Enhanced research capabilities by developing and maintaining essential, compliant research resources, promoting cross-functional collaboration and supporting a wide range of health and disease studies.
- Ensured HIPAA compliance and adhered to institutional review board (IRB) guidelines when handling patient-level EHR data, strictly controlling access to protected health information (PHI) and facilitating IRB-approved de-identification procedures.

University of California, San Francisco **San Francisco, CA**
Department of Epidemiology and Biostatistics
Research Data Analyst II / Biostatistician **Mar 2023 - Feb 2024**

- Analyzing xenobiotics, pharmaceutical and genomics data for impact of pharmaceutical metabolites on pre-term infants with broncho-pulmonary dysplasia.
- Creating detailed summary reports and analyses in R for respiratory samples collected from infants and children with broncho-pulmonary dysplasia.
- Assisting in research project linking gut microbiota to metabolites found in blood using MiME database.

- Effectively communicated findings and ongoing activities during lab meetings, fostering collaboration and informed decision-making within the research team.
- Staying abreast of new research in the field of genomics, biostatistics, and microbiology research.

Administrative Assistant **Feb 2020 - Aug 2022**

- Synthesized annual cohort data for 20 scholars using Qualtrics survey software, translating participant feedback into actionable insights that helped refine an informational program aimed at encouraging future clinical research coordinators.
- Managed scheduling and correspondence for 20 scholars and program administrators using tools like WhenIsGood and Outlook, ensuring a smooth training environment and effective communication processes.
- Supported program outreach by collaborating with senior administrators to improve data collection methods and amplify visibility through channels like Instagram and LinkedIn, including the design of promotional materials (logo, flyers) to attract potential clinical research coordinators.
- Created and maintained video tutorials from online classes and developed a detailed manual for software troubleshooting to support operational and educational needs for scholars.
- Engaged with scholars as a guest lecturer for the Study Design course, delivering presentations that clarified complex research concepts and encouraged academic participation.

Research Data Analyst I **Dec 2019 - Aug 2022**

- Modeled using SAS to analyze NHANES (national health and nutrition examination survey) and BRFSS (behavioral risk factor surveillance system) data to project the incidence and preventive impact of increased fruit and vegetable intake on cardiovascular health, estimating a prevention of 98,000 cardiovascular diseases and 16,000 deaths over a decade among U.S. adults aged 35-94, offering valuable insights that could influence dietary and health policies.
- Co-authored abstract in Circulation and presented a poster at the AHA EPI|Lifestyle Scientific Sessions 2022, disseminating key findings on dietary impact on cardiovascular health, highlighting potential substantial health benefits particularly among hypertensive populations.

Project Policy Analyst I **Aug 2019 - Dec 2019**

- Collected input data for nutrition as a risk factor for hypertension and cardiovascular disease for the CVD (cardiovascular disease) policy model.
- Worked with other researchers, analysts, and post-doctoral scholars on projecting incidence of heart disease in various groups in the U.S. using common heart disease risk factors integrated into the CVD policy model.
- Acquired training in statistical programming languages such as SAS, R.

Undergraduate Research Scholar **Jan 2019 - May 2019**

- Collected input data for cardiovascular disease policy model using nutrition as a risk factor.

San Francisco State University **San Francisco, CA**
College of Science and Engineering (CoSE)
Undergraduate Research Scholar **Jun 2018 - May 2019**

- Co-designed experimental methodology that studied the effects of the environment on the stress levels of 14 SFBUILD undergraduate scholars teaching at the Latino Cultural Center in Richmond, CA.
- Received faculty mentoring and training in biomedical research using SAS at UCSF on research project modeling the effect of diet on cardiovascular disease by reducing hypertension.

TEACHING EXPERIENCE

University of California, San Francisco **San Francisco, CA**
Department of Epidemiology and Biostatistics
Guest Lecturer **Annually, 2022 - Present**
Study Design | *Clinical Research Coordinators: Learners for Equity (CIRCLE) Program*

- Delivered an annual lecture on **study design** as part of the Day 2 session of the CIRCLE program.
- Focused on the basics of research, sampling schemes, **randomization**, **bias reduction**, and **statistics** tailored for prospective clinical research coordinators.

LEADERSHIP, VOLUNTEERING, AND COMMUNITY SERVICE

Los Medanos College **Pittsburg, CA**
American Medical Student Association

Vice President **Aug 2015 - May 2016**

- Coordinated medical professional guest speakers for career development for members.
- Helped to raise over $1,000 in fundraisers on campus geared towards Leukemia and Lymphoma awareness.
- Worked with the EMT department to provide free CPR AED classes to AMSA members.
- Won Outstanding Club of the Year award for 2015-2016 Fall/Spring semester.

KAISER PERMANENTE **Antioch, CA**

Volunteer **Jul 2015 - Feb 2017**

- Providing assistance to care providers such as nurses or other personnel by discharging members from any area in the hospital using wheelchairs.
- Helped Kaiser members navigate hospital campuses with ease using appointment lookups with medical record numbers.
- Respected protected health information by maintaining HIPAA compliance.
- Maintained hygienic standards with sanitizing policies consistent with reducing spread of nosocomial infections in the hospital such as stocking medical-grade wipes, medical- grade disinfectant, and proper glove disposal.
- Amassed over 200 hours of volunteer work.

PUBLISHED WORKS

Miguel Guardado, Dara Torgerson, Cheryl Chapin, **Azuka Atum**, Ryan D. Hernandez, B Ronald, Rebecca Simmons, Samuel Parry, Philip L. Ballard. *Urinary acetaminophen metabolites and clinical outcomes in premature infants.*

Atum, A. BS, Hennessy, S. PhD, Bsc (Hons), et. al. *Projected Reduction in Incident Cardiovascular Death from Increasing Fruit And Vegetable Intake in the United States.* 2022.

PRESENTATIONS AND POSTERS

Atum, A. BS, Hennessy, S. PhD, Bsc (Hons), et. al. *Projected Reduction in Incident Cardiovascular Death from Increasing Fruit And Vegetable Intake in the United States.* 2022. Poster presented at: AHA EPI|Lifestyle Scientific Sessions; March 1-4, 2022; Chicago, IL.

Gaitan, Leeza, Jackson, Eric, **Atum, Azuka**, Marquez-Magaña, Leticia, Ph.D. 2018. "The Effect of Environmental Change on Salivary Cortisol Concentration as an Indicator of Stress in SFBUILD Scholars". Poster Presented at: SFBUILD Student Showcase.

Atum, A., Dennis, S. 2018. "The Effect of Sugary Drinks on the Growth Rate of Cutaneous Staphylococcus sp." Poster presenteed at: 20th Annual SFSU College of Science and Engineering (CoSE) Student Project Showcase.

PORTFOLIO PROJECTS

Fixed-Effect Analysis of Baseline Serum Creatinine across Sex and Blood Pressure **Sept 2023 - Oct 2023**

- Performed an ANOVA to investigate baseline serum creatinine differences across sex and blood pressure categories, using electronic medical records from 491 patients at Tawam Hospital in Abu Dhabi (2008). Imported and formatted the dataset in SAS, categorized blood pressure levels, and employed PROC GLM to model main effects and interactions. Developed proficiency in analyzing health data, conducting post-hoc comparisons, and interpreting findings to understand kidney function variability.
- Corresponding SAS code and presentation.

Survival Analysis: Unraveling Factors in Post-Transplant Disease **Apr 2023 - May 2023**

- Performed a detailed survival analysis on chronic graft-versus-host disease (CGVHD) in leukemia patients using specific SAS procedures. Utilized the PROC LIFETEST for Kaplan-Meier estimates and PROC LIFEREG to fit generalized gamma Accelerated Time Failure (AFT) models, examining demographic and clinical covariates such as age, sex, CMV status, and waiting time to transplant.
- Corresponding SAS code.

Forecasting Wins and Losses with Machine Learning **Sep 2022 - Dec 2022**

- Crafted a predictive model utilizing Linear Discriminant Analysis to estimate the win-loss record of the Lakers during the COVID-19-impacted season. Gathered and refined game data from Basketball Reference, implementing R for data preprocessing and analysis. Executed modeling through the lda() function from the MASS package and visualized data interactions using ggplot2. Assessed model effectiveness with a split-sample validation technique to confirm its predictive accuracy.
- Corresponding R-code and presentation.

Tall People and Risk of Cardiovascular Disease **Apr 2023 - May 2023**

- Conducted a comprehensive analysis using SAS to explore the relationship between height and cardiovascular disease risk. Managed data importation, cleaning, and transformation tasks with SAS procedures like PROC IMPORT and PROC FREQ. Utilized PROC LOGISTIC for model development, focusing on variable selection and model optimization to identify

key risk predictors. This project enhanced my proficiency in SAS programming and statistical modeling, providing valuable insights into cardiovascular health metrics.
- Corresponding SAS code and presentation.

Regression Analysis of the Lakers' Regular Season **Mar 2022 - May 2022**

- Conducted a comprehensive statistical analysis to evaluate the predictive power of various basketball performance metrics, such as effective field goal percentage, free throw percentage, and games played, on points per game during the Lakers' 2021-2022 season. Utilized data from NBA.com and basketball-reference.com, addressing challenges related to data discrepancies and updating issues. Implemented variable elimination techniques to manage high collinearity among predictors, refining the model to identify key indicators of player performance. The insights derived from this analysis provide valuable contributions to the field of sports analytics.
- Corresponding R code and presentation.

Improving Diet to Prevent Cardiovascular Disease **Jan 2019 - Jan 2022**

- Utilized simulation modeling to assess the impact of dietary changes on cardiovascular health outcomes, focusing on statistical programming and data manipulation in SAS.
- Corresponding abstract and poster.

The Effect of Sugary Drinks on the Growth Rate of Cutaneous Staphylococcus sp. **Jan 2018 - May 2018**

- Orchestrated the design and execution of an investigation to explore how sugary drinks influence bacterial growth using BHI media and Staphylococcus cultures. Utilized turbidity measurements for monitoring bacterial proliferation, effectively managing limited laboratory resources to ensure accurate data collection. This endeavor enhanced my proficiency in experimental design, data analysis, and problem-solving in a team-oriented laboratory setting.
- Corresponding poster.

PROFESSIONAL MEMBERSHIPS AND CONFERENCES

American Statistical Association **August 2021 - Present**

- Conference on Statistical Practice (Feb 2023)
- Joint Statistical Meetings (Aug 2023)

American Heart Association **August 2020 - Present**

- Presented accepted abstract/poster at AHA EPI|Lifestyle Scientific Sessions March 2022.

College of Science and Engineering (CoSE) **May 2018**

- Presented accepted poster at SFSU student showcase.

R for Data Science Book Club **May 2023 - Present**

International Statistical Engineering Association **Nov 2023 - Present**

Society for Clinical Trials **Nov 2023 - Present**