Michael K. Gureghian

Los Angeles | mguregh@ucla.edu | 310 560 7243 | https://www.linkedin.com/in/mgureghian/

Education

University of California, Los Angeles, BS in Statistics & Data Science Sept 2023 – June 2025

- **Coursework:** Data Analysis and Regression, Experimental Design, Mathematical Statistics, Monte Carlo Methods, Linear Algebra, Statistical Models

Pasadena City College June 2022 – May 2023

Experience

Undergraduate Laboratory Assistant, David Geffen School of Medicine at UCLA – Los Angeles, CA July 2024 – Present

- Processed and made visualizations for DNA concentration data in decellurization treatment study.
- Designed and administered comparative analysis of protein intensity data collected in experimental study.
- Wrote IRB proposal for surgical evaluation of a biomedical device.

Data Engineering Intern, Tala – Santa Monica, CA Aug 2023 – Sept 2023

- Created a back-end tool to automate large-scale Snowflake administrative tasks.

Research Assistant, Seattle Children's Hospital – Seattle, WA July 2021 – Aug 2023

- Developed RedCap databases and data collection instruments for clinical data.
- Sorted, formatted, and transferred appropriate clinical data from Epic EHR into databases for analysis by senior researchers.

Projects

Predicting Student Alcohol Consumption 2024

- Studied the influence of 28 factors in predicting secondary school students' level of alcohol consumption given survey data with 1044 observations.
- Designed a logistic regression model with 10 predictors selected with random forest and stepwise feature selection.

Kaggle Competition: Predictive Analysis of Obesity Diagnostic 2024

- Performed EDA and MICE imputation on training dataset with 29 predictors and 32014 observations.
- Built random forest model with most important features selected by cross-referencing VIF, random forest variable importance, and stepwise AIC.
- Achieved testing accuracy of 99.06%.

Technologies

Languages: R, Python, SQL, C++

Technologies: Snowflake, Microsoft SQL Server, docker, pandas, TensorFlow, Epic EHR, RedCap, Excel