Sean Mulherin

smulherin519@gmail.com
seanmulherin.github.io
Los Angeles, CA
336.681.0821

Research Interest

Time Series Analysis, Point Processes, Education, Social Statistics

Education

M.S. Applied Statistics and Data Science, University of California, LA 2023 - Present

Testing for Causal Clustering in K-12 Student Misconduct

Advisor: Dr. Rick Schoenberg

M.A.T Secondary Mathematics, University of North Carolina, Chapel Hill 2020

Advisor: Dr. Josh Corbat

B.S. Mathematics, North Carolina State University 2019

Professional Experience

Brentwood School 2024 - Present
- Physics Faculty
- Data Science Expository Course

Jackson Hole High School 2021 - 2023
- Math Faculty: Geometry, AP Prep Algebra II, Trigonometry/Precalculus
- Cross Country & Track Coach
- Chess Club Coach

Mountain Academy of Teton Science Schools 2020 - 2021
- Lead Math Faculty: Algebra, Geometry, IB Applications & Interpretations
- Health & Wellness Teacher
- Academic Advisor

Carrboro High School 2019 - 2020
- Student-Teacher: Geometry, AP Calculus AB, AP Calculus BC

North Carolina State University Tutorial Center 2016 - 2018
- Math Tutor: Calculus, Foundations of Mathematics, Differential Equations, Probability

Research Experience

Advanced Studies Institute in Mathematics of Data Science & Machine Learning 2024

- Sponsored by the National Science Foundation, I traveled to Uzbekistan to participate in a two-week-long workshop focusing on the mathematics of machine learning. Topics covered include model-based clustering, Hawkes point processes, benign overfitting, generalization, double descent, and mirror descent.

North Carolina State University, College of Design 2017 - 2018

- As a research assistant, I collected data pertaining to the efficacy of healthy diets on the social, emotional, and academic performance of elementary school students.

Projects

Master's Thesis, Testing for Causal Clustering in K-12 Student Discipline Data 2024

- The main objective is to test the degree to which causal triggering explains the event of student misconduct. To achieve this, a test introduced by Kresin (2023) and McGovern (2024) is applied, wherein likelihood-ratio tests are performed using information gain statistics to compare the fit of a Neyman-Scott model to that of a Hawkes model.

Financial Statistics Models 2024

- Programmed a web application that provides two tools operating at the confluence of statistics and finance. The first is a forecasting tool that uses the Prophet model to fit and predict the daily adjusted closing price of a specified asset. The second tool is one for portfolio optimization and valuation wherein the user inputs multiple assets and relevant performance metrics are displayed such as volatility, return, weights, and so on.

Forecasting Selling Price of Houses in the U.S. 2024

- Built an interface for users to input a city and forecast period to observe and explore the forecasted trends and prediction metrics computed by the Prophet model. This model is notorious for its accuracy and flexibility. Housing prices are atypically volatile so this model does particularly well at producing accurate home price forecasts.

An Artificial Neural Network Approach to Identifying Diabetes Risk Status 2023

- Programmed a neural network from scratch to classify a user's risk of developing type II diabetes after completing a 21 question survey. My model was trained using CDC data and achieves 84% accuracy in its validation-set predictions.

Tracking Carbonization 2023

- Engaged in a comprehensive statistical analysis of the current state of global carbon dioxide emissions. Data was obtained from the United National Development Program and analyzed using R.

A Classification Analysis on Breast Cancer Tumors 2023

- Built and compared models to classify breast cancer tumors as malignant or benign. Models compared include: linear discriminant analysis, quadratic discriminant analysis, support vector machines, logistic regression, random forests, Naive Bayes, and KNN. The most optimal model used linear discriminant analysis to predict with 97% accuracy on the validation set.

An Ethereum Regression Analysis 2023

- Build and compared multiple regression models to predict the USD-ETH price in the year 2030. Models include: linear, quadratic, cubic, exponential, and logarithmic regression.

Note: all projects can be found on my online portfolio linked in the header.

Skills & Appointments

National Institute of Statistical Sciences GSN Council Member	2023 - Present
UCLA Statistics Graduate Student Association, VP of External Affairs	2023 - Present
UCLA Math and Physical Sciences Council Member	2023 - Present
DataFest Conference Guest Speaker - Introduction to R	2024
DataFest Conference Guest Speaker - Data Cleaning and Wrangling in R	2024
R Programming Certification - DataCamp	2022
Python for Data Science Certification - DataCamp	2022
NCAA Division I Cross Country & Track Athlete	2015 - 2019