Drake Watson

253-341-2882 watsond1@uci.edu https://www.linkedin.com/in/drake-watson/ github.com/datadraco

Highly motivated graduate student with a strong foundation in mathematics and diverse data science methods. Experienced statistical researcher with effective communication and collaborative skills seeking opportunities to contribute to a progressive team working on innovative projects.

Skills

Python, R, SQL, Data Wrangling + Exploratory Analysis, Visualizations, Dashboards (Tableau) — Machine Learning: Linear Regression, Neural Networks, Decision Trees, KNN, etc. — Database Management: Querying, Efficiency, Structures (Relational, Json, Graph, etc.), Applications (Postgres, AWS, Cassandra, Neo4j, Spark)

Education

University of California, Irvine — MS in Data Science **2023 – December 2025**

- Statistics, Database Management, Big Data Management, Algorithms, Artificial Intelligence, Machine Learning, Data Structures, etc.

University of Washington, Seattle — BA in Mathematics **2020 – 2022**

- Advanced + Abstract Linear Algebra, Real + Numerical Analysis, Data Visualization, Data Programming, Probability
- Thomas P. Bleakney Endowed Scholarship in Mathematics – 2022
- Academic Excellence Scholarship – 2022
- Dean's List Autumn 2020 through Winter 2022

Tacoma Community College **2017 – 2020**

- Associates in Mathematics / Computer Science
- Calculus, Linear Algebra, Differential Equations, Mechanical + Electromagnetic Physics, Java Programming, etc.

Experience

Statistical Research Programmer **September 2023 – October 2024**

Chapman University *Irvine, CA*

- Collaborated with researchers within the School of Pharmacy at Chapman and various other universities across the country on end-to-end manuscript development.
- Provided critical analysis of previously developed research in the process of editing for resubmission for publication.
- Built out an ETL pipeline for locally reading and annually updating dataset of over 300 million observations.
- Created framework for performing statistical analysis in support of diverse research manuscripts focused on drug prescriptions for the state of California from 2010-2023.
- Applied techniques such as data processing, geospatial analysis, interrupted time series, difference of means testing, data visualization, and manuscript editing.

Publications

Measurement Bias in ML-Enhanced Opioid Risk Scoring Systems **Nature, October 2024**

- Analyzed California's PDMP data and ZIP code-based socio-demographic characteristics to examine correlations between the common predictive features used in clinical decision support tools and privacy protected patient attributes.
- Findings indicated that predictive features might not favor females, minority populations, older patients, neighborhoods with high percentages of disability and unemployment, and Medicare patients.
- End to end development of data collection, cleaning, collaborative interpretation with the Chapman School of Pharmacy, statistical analysis, visualizations, and manuscript development.

Examining Bias in the Narxcare Score **Value in Health, June 2024**

- Scrutinized key predictive features used by a prominent AI/ML-based clinical decision support system known as NarxCare.
- Focused on unveiling likely discriminatory patterns that negatively affected equitable opioid prescribing.
- Formulated system of statistical analysis that could be expanded for larger scale research regarding clinical decision support systems.

Changes in Opioid Prescribing Before and After COVID-19 in CA **researchsquare, November 2024**

- Constructed a Poisson regression model to perform interrupted time series analysis interrogating COVID-19's affect on drug prescription behaviors in California.
- Discovered statistically significant evidence that multiple problematic drug prescribing patterns had meaningfully increased following the onset of COVID-19.
- Collaborated with the Chapman School of Pharmacy on data interpretation, visualizations, and manuscript development.

Buprenorphine Access in CA **American Journal of Pharmaceutical Education, September 2024**

- Aimed to compare the availability of active buprenorphine-prescribing clinicians in California to the Substance Abuse and Mental Health Services Administration (SAMSHA)-listed DATA waived prescribers under each 5-digit ZIP Code
- Developed choropleth visualization of California ZIP Codes in order to evaluate the relationship between the demand and availability of buprenorphine.
- Collaborated with multiple universities and researchers on final manuscript and visualization.

Projects

NFL Quarterback Performance Projection | *python, R, scikit-learn, tensorflow* **July 2024**

- Engineered several diverse NFL quarterback datasets into a curated database posted publically on Kaggle.
- Defined and calculated statistically meaningful tiers of quarterback success in a way that reflected real-world consensus.
- Trained a neural network using keras and tensorflow with the purpose of predicting quarterback success levels and found the strongest possible OLS linear regression model.
- Wrote a detailed article walking through creation and decision making process of the project and published to Medium.

Video Game Sales Analysis | *python, scikit-learn, plotly* **Spring 2021**

- Created a dataset encapsulating global video game performance from several unique sources, performed EDA to find meaningful correlations, and developed an interactive Tableau dashboard for presentation.
- Built several decision tree models using scikit-learn in an attempt to train a model for sales projection performance.