

Christian Smith

714-261-7962 | smithc4751@gmail.com | christiansmith.netlify.app | github.com/CSmith47

EDUCATION

University of California, Los Angeles

September 2022 - December 2025

Bachelor of Science in Statistics and Data Science

Los Angeles, CA

- Honors: Summa Cum Laude (Cumulative GPA: 3.992/4.0)
- Relevant Coursework: Introduction to Mathematical Statistics, Introduction to Data Analysis and Regression, Introduction to Probability, Introduction to Computational Statistics with R, Introduction to Statistical Models and Data Mining

WORK EXPERIENCE/INTERNSHIPS

Data Analyst

June 2024 – Sept 2024

Edwards Life Sciences

Irvine, CA

- Collaborated with senior data analyst to perform feature selection and dimensionality reduction on 7 datasets from 2 clinical studies, merging and comparing data for analysis on a device used across both studies.
- Extensively utilized R for data manipulation, cleaning, and preparation of large clinical datasets, exporting them to SAS for submission to the FDA, ensuring compliance with regulatory standards.
- Validated the integrity and accuracy of clinical datasets by cross-checking with study documentation, identifying a critical discrepancy that had gone unnoticed for 3 years, improving data quality.

Techzone Technician

Aug 2023 – June 2024, Sept 2024 - Present

Ackerman Computer Store - ASUCLA

Los Angeles, CA

- Worked collaboratively in a team of 1-3 members in an Apple authorized repair store to provide exceptional customer service, frequently resolving technical issues for customers the same day they visit the store.
- Performed over 100 repairs as an Apple Certified Technician, addressing hardware and software issues for Apple devices as well as other brands, including Microsoft, Dell, and HP.
- Leveraged strong troubleshooting and diagnostic skills to efficiently address problems and present solutions to customers, fostering a reputation for reliability and competence in problem solving.

PROJECTS

The Role of Statistics in Evaluating College Quarterbacks | *Python, SHAP, Scikit-learn*

Feb 2025 – March 2025

- Poster link: [Link to PDF Image](#)
- Processed and transformed a dataset of 1,639 observations in Python to enhance model performance and extract insights.
- Utilized weighting and undersampling to balance prediction models and increase F1-Score performance by up to 12%.
- Engineered a weighted performance metric utilizing SHAP values for non-technical audience interpretability.
- Synthesized key findings into a poster presentation, effectively communicating insights to a non-technical audience.

Reimagining Reviews With Sentence Transformers | *Python, BERTopic, VADER, Scikit-learn*

Feb 2024 – March 2024

- Article link: <https://rb.gy/s5v10d>
- GitHub link: <https://github.com/CSmith47/Reimagining-Reviews-With-Sentence-Transformers>
- Trained a BERTopic model across 3 distinct types of business reviews aiming to enhance business performance through topic-driven insights.
- Generated 6 distinct visualizations using BERTopic to illustrate the relationships between topics, their variations based on star ratings, and the evolution of topic importance over time.
- Utilized a VADER sentiment analysis model to visualize and interpret review sentiment to explain the nuances and inconsistencies of natural language application using data-driven processes.
- Constructed a logistic regression model with Scikit-learn to demonstrate the power of sentence transformers in capturing complex language patterns for predictive modeling and analysis.
- Penned an analysis of my findings through an article published by the platform: [Python in Plain English](#)

The Overwatch League: Statistics For Success | *Python, Scikit-learn, SQL, GitHub*

May 2023 – July 2023

- Article link: <https://rb.gy/66oqj>
- GitHub link: https://github.com/CSmith47/OWL_Stats_For_Success
- Transformed and simplified one large, complex dataset with millions of observations into two, optimized datasets uploaded to GitHub, making the data more accessible for public use while preserving critical insights for analysis.
- Feature engineered a completely new, weighted statistic to analyze individual player performance and predict game wins and losses with high accuracy.
- Generated 20 visualizations to display key statistical insights among team performance and feature importances of a generated random tree classifier model from the Scikit-learn library to predict team wins from significant statistics.

TECHNICAL SKILLS

Languages: Python, R, SQL, C++, LaTeX

Libraries/Tools: Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, SHAP, BERTopic, GitHub, Visual Studio