

Christian Smith

714-261-7962 | christian6251@gmail.com | [linkedin.com/in/smithc47](https://www.linkedin.com/in/smithc47) | github.com/CSmith47

EDUCATION

University of California, Los Angeles

Los Angeles, CA

Bachelor of Science in Statistics and Data Science, Minor in Data Science Engineering

Expected to Graduate: March 2026

- Cumulative GPA: 3.989/4.0
- 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

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 in-depth analysis of my findings through an article posted on the Medium platform that was subsequently 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
- Uploaded and managed a dataset containing over one million rows into a SQL database for efficient querying and data exploration, optimizing performance for large-scale data analysis.
- Transformed and simplified one large, complex dataset 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.
- Constructed a random tree classifier model using Scikit-learn to predict match outcomes based on team and player statistics.
- Generated 20 visualizations, 18 of which display statistical variations among teams' performances and 2 of which show the most important features of the predictive model.
- Penned an in-depth analysis of my findings through an article posted on the Medium platform.

TECHNICAL SKILLS

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

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