

Daniel Ryan Furman

1935 Jefferson Street • San Francisco, CA 94123 • daniel_furman@berkeley.edu

[Portfolio](#) • [LinkedIn](#) • [Google Scholar](#) • [GitHub](#)

EDUCATION

University of California, Berkeley, School of Information – Berkeley, CA

Anticipated May 2023

Master of Information Management and Systems

Specialization: Applied Data Science and Machine Learning | Graduate Certificate in Applied Data Science

University of Pennsylvania, College of Arts and Sciences – Philadelphia, PA

Graduated May 2020

Bachelor of Arts with Distinction in Earth Science

Minor: Mathematics • **Cumulative GPA:** 3.79/4.0

Honors: CURF Rose Research Award, NSF Award #1757952, CURF Climate Action Grant, EES Hayden Scholar

Activities: NCAA D1 Golf Team • Experimental Geophysics Lab

Harvard-Westlake School – Los Angeles, CA

Graduated June 2016

Cumulative GPA: 4.2/4.0 (Weighted) • **ACT:** 34 (Composite Score) • **Activities:** Varsity Golf Team Co-Captain

RELEVANT EXPERIENCE

Data Science Consultant – Contract & Freelance

March 2019 - Present

- Refined geospatial software for **Comon Solutions** working cross-functionally as an intern on the Engineering team by developing automated image processing and vision modeling pipelines in Python & AWS, products which drive real-time natural resource strategy for land stakeholders via big geo-data insights (files ~ 100's GBs). [[Link](#)]
- Strategized recommendations for **De Castro, West Inc.** to a real estate client restructuring 11 unevenly shared assets into 3 majority ownerships by building a search algorithm that minimized stake change between partners (value + debt), finding an optimal state with only ~\$2.9mm in stake change against ~\$159mm in total assets. [[Blog](#)]
- Extracted product insights for the **Official World Golf Rankings Ltd.** with two decades of proprietary data (16 feats. x 3400 tournaments across 21 tours). Developed PCA biplot visualizations that helped drive stake-holders to ship ranking system updates in Q3 '22, which more accurately capture relative performance and strength of field.

Data Mining Teaching Assistant – Berkeley Haas School of Business

September 2021 - Present

- Collaborated on "Descriptive & Predictive Data Mining" with Prof. Luyi Yang by sourcing data sets for assignments and answering questions about R & exploratory data analysis for non-technical MBA students. [[Syllabus](#)]

Data Science REU – Harvey Mudd College

May 2019 - August 2019

- Constructed GIFs of climate impacts for four marquee CA species from bioclimate features and species presence records (19 feats. x 6800 obs.), resulting in a rich ML-powered visualization of CA eco-risks. [[DataViz](#)] [[Poster](#)]
- Developed a Python class with Catboost & XGBoost geo-classification through an advanced experimental design with multi-seed/multi-sample frameworks (AUCs>0.95, domain SOA). Built R functions for ETL & post-processing.

PROJECTS & INTERESTS

SurfCrowds.ai – Python (NumPy, Pandas, Scikit-learn), AWS (Lambda, S3), Dash, Heroku. [[Web-App](#)]

- Built a crowd size predictor for Malibu, CA with Random Forest & Gradient Boosting ensemble regression running as an AWS Lambda function, equipping my hometown surf community with new types of AI-driven forecasts.

Instagram Style Clustering with NLP – Python (NLTK, SpaCy, HuggingFace), R (shiny, tidyverse). [[DataViz](#)]

- Explored author attribution on Instagram with captions from health food influencers with highly related topical content, using KMeans on BERT embeddings (capturing stylistic context) to cluster three profiles (NMI = 0.6).

Kaggle Competitions – Python (PyTorch, TensorFlow, Keras), Neptune.ai. [[BirdCLEF21 Blog](#)]

- BirdCLEF21: Won solo bronze out of 816 teams for Sound Event Detection with multi-label CNNs (ResNest18) and metadata classifiers (397 classes, ~63,000 obs.), overcoming a large train/test shift and weak labeling (F1=0.61).
- SETI Search: Placed top 17% out of 768 teams for radio anomaly detection by bagging ViTs and CNNs (AUC=0.77).