

# NOAH JAFFE

## DATA SCIENTIST

✉ noah.h.jaffe@gmail.com  
☎ 415-686-6967  
📍 Portland, OR 97223  
in noah-h-jaffe/  
🌐 https://github.com/njaffe

## Skills

### PYTHON

Numpy  
Pandas  
SciPy  
Scikit-learn  
Matplotlib  
Seaborn

### MICROSOFT OFFICE

Word  
Excel  
Powerpoint

### MACHINE LEARNING

Linear Regression  
Logistic Regression  
KNN classification  
Bayesian classification

### SQL

PostgreSQL

### R

### TABLEAU

### NATURAL LANGUAGE PROCESSING

NLTK  
Spacy  
Principal Component Analysis

## Education

**San Francisco State University** Jan. 2020  
MS Marine Biology 2020

**Pitzer College** May 2016  
B.A. Organismal Biology 2016

## Experience

### Metis Data Science Bootcamp

Data Scientist

Live Online format  
Sept. 2020 to Current

- ACCET accredited 12 week immersive data science bootcamp focused on project-oriented learning
- Core curriculum centered around Python, statistics, supervised and unsupervised machine learning, exploratory data analysis, databases, and visualization techniques
- Completed five self-designed data science projects from conception to presentation; including data collection, data management, exploratory data analysis, modeling, and visualizations

### Compass Education Group

ACT/SAT Tutor

Larkspur, CA  
Fall 2018 to Summer 2019

- Tutored students in one-one-one ACT and SAT exam preparation

### San Francisco State University

Teaching Assistant, Marine Invertebrate Zoology Course

San Francisco, CA  
Aug. 2016 to Dec. 2016

- Co-organized student laboratory activities
- Co-led field collecting and surveying trips in intertidal habitats. Led laboratory dissections of all major invertebrate phyla
- Assisted students with lab and field data analysis exercises

Teaching Assistant, NSF REU program

San Francisco  
Summer 2016, Summer 2017, Summer 2018

- Guided original research in ecology, evolution and population genetics in National Science Foundation-funded Research Experience for Undergraduates program

Research Mentor, CSU STAR program

San Francisco, CA  
Summer 2018 to Fall 2018

- Mentored genetic research in California State University-funded Science Teacher and Researcher program
- Assisted with design, planning and implementation of genetic research project

### Richardson Bay Audubon Society

Teacher Naturalist

Tiburon, CA  
May 2014 to Aug. 2014

- Designed and taught scientific- and stewardship-focused summer curriculum
- Led tours and camp programming to promote stewardship at marine and bird sanctuary

## Projects

### Scientific publication: "Sea star wasting disease demography and etiology in the brooding sea star *Leptasterias* spp."

Fall 2016 to Winter 2019

- Published original ecological and biological survey research
- Worked directly and remotely with multiple collaborators

### Using Linear Regression to Predict MTA Subway Traffic

Sept. 2020 to Oct. 2020

- Project 1 of Metis Curriculum
- Collaborated with three other Metis students
- Optimized placement of non-profit volunteers through linear regression analysis on NYC MTA Turnstile data
- Provided recommendations to fictional company "Women Tech Women Yes" based on our results

### Using Linear Regression to Predict Magic: the Gathering Card Prices

Oct. 2020 to Oct. 2020

- Project 2 of Metis Curriculum
- Webscraped MTG card price data from CardKingdom.com
- Used linear regression techniques within to predict prices for Magic: the Gathering cards within \$1.50

### Comic character classification

Oct. 2020 to Nov. 2020

- Project 3 of Metis Curriculum
- Analyzed dataset of comic book character attributes from Kaggle.com
- Used classification models including logistic regression, K-nearest neighbors, Random forest, and XGBoost to categorize characters as "Good", "Bad", or "Neutral"
- Identified characteristics associated with moral alignment

### Using NLP to model topics in horror movies

Nov. 2020 to Nov. 2020

- Project 4 of Metis Curriculum
- Webscraped 125 horror movie scripts from IMSDB
- Performed topic modeling techniques including Singular Value Decomposition, Non-negative Matrix Factorization, and Latent Dirichlet Allocation
- Grouped movies into six subgenres using clustering and t-distributed stochastic neighbor embedding