Joshua Tollefson

San Francisco, CA **|** 612-387-3467 **|** josh.tollefson9@gmail.com
linkedin.com/in/joshua-tollefson **|** github.com/josh-tollefson

SKILLS

Languages:	Python, SQL, Matlab
Libraries:	NumPy, SciPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Keras / TensorFlow, NLTK
Tools:	Jupyter, Linux, Bash**,** Github
Methods:	NLP, ARIMA, Hypothesis Testing, Markov Chain Monte Carlo (MCMC), Image Processing, Regression (Linear Least-Squares & Logistic), Naive Bayes, Random Forest

EXPERIENCE

Insight Data Science Fellow　　　　　　　　San Francisco　　　　　　　　May 2020 - Present
- Completed data science training for academics transitioning into industry, including building a product in four weeks that provides insights for board game developers seeking feedback on their prototypes.
- Predicted sentiments and categorized contents of >100k board game text reviews by writing a Python NLP pipeline with binary and multi-label classification methods, achieving accuracies of 80-90%.
- Built a web application with Streamlit and hosted on AWS to provide a report predicting the general sentiment and common features from user-uploaded board game text reviews.

Postdoctoral Scholar　　　　　　　　University of California Berkeley　　　Sep 2019 - Present
- Led team of 5 collaborators across two observatories to image Neptune in order to provide scientists context on how the outer solar system formed and benchmarks for future space missions to Neptune.
- Processed 20 GB of this data with Python-based deconvolution methods, boosting image resolution by a factor of 5 over prior work.
- Retrieved atmospheric properties of Neptune by developing MCMC and chi-squared testing routines for a pre-existing Python package, improving confidence on parameters by 200%.
- Mentored two undergraduate students through Python-based image processing and photometric projects, seeing one student get a first-author publication.

Graduate Student Researcher　　　　　　University of California Berkeley　　　Aug 2013 - Aug 2019
- Tracked atmospheric variations on 200 Jupiter images by creating an image cleaning and auto-correlation pipeline in Matlab and Python, providing up-to-date profiles for NASA's *Juno* mission.
- Determined the location and velocity of hundreds of Neptunian clouds by devising a Python and SQL program using linear least-squares regression on photometric data from over 100 images.
- Designed Python and Jupyter tutorial to help students integrate programming into their end-of-term projects by teaching data manipulation and visualization techniques with Numpy and Matplotlib. Recognized as an Outstanding Graduate Student Instructor for having 90% favorable ratings.

PROJECTS

- Built a Python application forecasting air quality with ARIMA and VAR in collaboration with Insight fellows on Github. Scraped and cleaned five years of daily California sensor data from the EPA.
- Constructed a Jupyter notebook giving a step-by-step tutorial of ARIMA modeling in Python.
- Built a tic-tac-toe board game and trained an AI to win using reinforcement Q-learning in Python.

EDUCATION

Ph.D.	Earth & Planetary Science	University of California Berkeley	2019
B.S.	Applied Mathematics & Planetary Science	California Institute of Technology	2013