Quinn Keck

614-596-2292 | keck.quinn@gmail.com | San Francisco, CA
github.com/keck343 | linkedin.com/in/quinn-keck/

EDUCATION

University of San Francisco **June 2019 (Expected)**
M.S. in Data Science
- Courses: Statistics, Linear Regression, Relational Databases (SQL), Data Acquisition, Time Series Analysis, Distributed Computing (Spark), Machine Learning, and Design of Experiments (A/B testing).

Hampshire College **May 2012**
B.A. in Liberal Arts

RELEVANT EXPERIENCE

Bay Area Metropolitan Transportation Commission
Data Science Consultant **2018-Present**
- Developed machine learning (Python, sklearn) and times series models (R) to predict daily Clipper Card usage within 4%.
- Analyzed Clipper and FasTrack data, tracked key performance indicators (Jupyter Notebooks, SQL) and created interactive dashboards (Tableau, Power BI, Google Data Studio).
- Centralized reporting infrastructure (AWS, S3, Python) to remove silos so multifunctional teams could upload and access data.
- Enabled leadership to make data-driven decisions by building more efficient, timely, and consistent reporting.

ADDITIONAL EXPERIENCE

Special Education Assistant Teacher
Hilltown Cooperative Public Charter School **2017-2018**
- Engaged students in grades 2-6 with individualized education plans in lessons and projects and led after school activities.
- Designed learning materials and curriculum for students with learning disabilities in coordination with classroom teachers.

Western Mass Recovery Learning Community
Peer Advocate **2014-2016**
- Led public workshops and community support groups. Presented at a national rights protection and advocacy conference.
- Organized biweekly meetings to update and revise community norms with input from community members and staff.

DIAL/SELF Youth Services
Youthserve AmeriCorps Member **2012-2013**
- Organized helmet safety campaign and youth civic engagement events with the Gill-Montague Community School Partnership.
- Screened and recruited volunteers to mentor elementary students, led mentorship activities, and coordinated fundraising efforts.

DATA PROJECTS

Open Payments
- Detected physician payment anomalies by predicting if a physician's average payment from drug manufacturers would be in the top decile, which may be indicative of potentially unethical behavior, to serve as an audit list for an investigating agency.
- Using PySpark, developed data pipelines, engineered features and fit a Random Forest Model to 29 GB of data to predict if mean payments of 900,000 individual physicians were in the top decile, resulting in an accuracy of 0.92 and F-score of 0.91.
- Co-author of *Predicting Unethical Physician Behavior At Scale: A Distributed Computing Framework*, SmartWorld 2019. github.com/unethical-physician-predictions/open-payments

Aligned Yoga App
- Developed an end to end web app that could tell users if their yoga pose was properly aligned and safe (EC2, Python, Flask).
- Processed videos users uploaded with their webcam using Python and Deep Learning on an AWS EC2 instance.
- Wrote the code for the core multi-label model that took inputs from a deep learning model, classified user's poses, and produced direct and accurate feedback on what a user should change so their pose is safe and correct. youtu.be/t8HMLYR1-FE

Predicting In-App User Purchases
- Predicted if users of an app would make a purchase in the next 7 or 14 days. Analyzed 45 GB of data spanning over 3 months.
- Using Python and pandas, engineered features based on user behavior, location, and device type. Fit Random Forest and XGBoost models; the final model achieved an AUC of 0.965.

Yelp Reviewer Toughness Classification
- Using Python and sklearn, fit Logistic Regression and Support Vector Machine Classification models to identify which users give harsher reviews (whose ratings are one standard deviation below a restaurant's average review.)
- Model achieved a 93.8% accuracy and F1 score of 0.936. github.com/keck343/yelp_user_toughness

TECHNICAL SKILLS

Python, R, SQL, PySpark, MongoDB, AWS, Github, Data Visualization (ggplot, plotly), Machine Learning.