

PHI NGUYEN

888 Douglass Street, San Francisco, CA 94114

phi.nguyen@outlook.com | 510-850-6733 | www.phinguyen.me

WORK EXPERIENCE

Data Scientist | *Gusto* | *San Francisco, CA* | *August 2021 – Present*

Data Scientist | *Cisco Meraki* | *San Francisco, CA* | *May 2020 – August 2021*

- Created models in Python to improve end-to-end customer support experience, such as topic modeling to identify common issues and binary classification to understand drivers of case creation.
- Spearheaded AB testing efforts organization-wide, including an experiment that reduced support case close time by 16%.

Data Scientist | *Priori Data* | *Berlin, Germany* | *Nov 2017 – Oct 2018*

- Developed several machine learning models (GLM with regularization) to estimate app revenue & app stickiness, and used Python, Airflow, and Google BigQuery to schedule model deployment.
- Improved model performance metric MAPE (median absolute percentage error) by 17% in certain geographies and categories by creating benchmark observations and changing model formulation.

Senior Data Analyst | *Intuit* | *Mountain View, CA* | *Aug 2013 – Feb 2016*

- Led analytics and AB testing for Quickbooks.com using SiteCatalyst, SQL, R, and Tableau.
- Evaluated performance (conversion rate, retention, revenue, mix shift) for high-profile AB tests that led to introduction of new product line on the website (Quickbooks Self-Employed).

EDUCATION

Humboldt Universität zu Berlin | *Berlin, Germany* | *Oct 2016 – Oct 2019*

M.S. Statistics | 3.6 GPA [1,7 in German grading system]

Master's Thesis: Measuring the Cost of Congestion in Berlin using DBSCAN-Identified Travelers

UC Berkeley | *Berkeley, CA* | *Aug 2009 – May 2013*

B.A. Operations Research and Management Science | 3.86 GPA [Honors with Distinction]

ADDITIONAL EXPERIENCE

Data Scientist Volunteer | *LifeMoves* | *Menlo Park, CA* | *February 2021 – Present*

- Devised a data science strategy and roadmap to aid fundraising efforts.

SKILLS

Tools: Python (pandas, numpy, sklearn, statsmodels), R (tidyverse, mlr, R Shiny), SQL, git, Markdown, AWS + GCP (limited working knowledge), Docker

Techniques: regression, classification, clustering, machine learning, statistics, econometrics, time series, topic modeling, survival analysis, hypothesis testing (frequentist + Bayesian), visualization