# PHI NGUYEN

42266 Live Oak Circle, Fremont, CA 94538
phi.nguyen@outlook.com | 510-850-6733 | www.phinguyen.me

## EXPERIENCE

### Researcher

*Mercator Research Institute | Berlin, Germany | Nov 2018 – Nov 2019*

- Developed statistical and machine learning models in R to estimate the welfare loss of traffic congestion in Berlin, and proposed a congestion charge that would reduce traffic congestion by 0.2 – 0.5% (approximately 1.2 – 3 billion EUR revenue gain annually).
- Solved identification challenges in supply and demand framework by using an instrumental variable approach and by generating counterfactual observations.
- Used DBSCAN clustering and over 32 million geospatial data points to model traveler heterogeneity.

### 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.
- Used Fellegi-Sunter probabilistic record linkage approach with Jaro-Winkler distance to match records from different sources in order to generate reports for clients.

### Senior Data Analyst

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

- Managed web 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 – Nov 2019*

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

### UC Berkeley

*Berkeley, CA | Aug 2009 – May 2013*

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

## SKILLS

**Data Science & Statistics**: Python, R, SQL
**Visualization**: ggplot2, Tableau, R Shiny, qGIS, d3.js
**Other**: Git, Markdown, command line, LaTeX, Excel