

# PHI NGUYEN

## DATA SCIENTIST

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

## EDUCATION

### Humboldt Universität zu Berlin

*Berlin, Germany | Oct 2016 – Nov 2019*

M.S. Statistics | 1.7 [equivalent to 3.6 GPA]

### UC Berkeley

*Berkeley, CA | Aug 2009 – May 2013*

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

## EXPERIENCE

### Researcher

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

- Developed causal inference models to estimate the deadweight loss of traffic congestion in Berlin.
- Used qGIS to analyze GPS data from over 32 million unique trips, the DBSCAN clustering algorithm to generate trip clusters, and R to build regression models.
- Proposed a time- and distance-based congestion charge that would yield welfare gains of 0.2 – 0.5% (approximately 1.2 – 3 billion EUR annually).

### Data Scientist, Working Student

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

- Developed and optimized several machine learning models to estimate app revenue, app stickiness, and record linkage for apps in both iOS and Google Play stores.
- Drove model changes that led to up to 17% improvements in MAPE (median absolute percentage error) for certain geographies and app categories.
- Used R, Python, and Airflow to automate loading of estimates into Google BigQuery.

### Senior Data Analyst

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

- Managed web analytics for Quickbooks.com using Adobe SiteCatalyst, SQL, and Tableau, including page flow analysis and funnel analysis.
- Led over 20+ AB tests ranging from pricing changes to product lineup presentation, using SiteSpect, Excel, R, and SQL for analysis of results.
- Collaborated cross-functionally to communicate key findings and insights to primary stakeholders.

## SKILLS

**Data Science & Statistics:** R, Python, SQL

**Visualization:** ggplot2, Tableau, R Shiny, qGIS

**Other:** Git, Markdown, command line, LaTeX, Excel