

# JAE YEON KIM

## Overview

- Behavioral Data Scientist studying how people think and behave using behavioral science, statistics and data science tools
- Research with sampling, survey design, survey experiments, natural experiments, factor analysis, cluster analysis, multivariate and hierarchical regression analysis, computational text analysis, machine learning
- Research areas: behavioral science, computational social science, experimental and survey methods, measurement, causal inference

## EDUCATION

Present	<b>University of California, Berkeley</b> PhD Candidate in Political Science	📍 Berkeley, California, USA
Summer 2019	<b>Russell Sage Summer Institute in Computational Social Science</b> Selected as one of 29 participants from a nationwide competition	📍 Princeton University, Princeton, USA
2016	<b>University of California, Berkeley</b> MA in Political Science	📍 Berkeley, California, USA
2011	<b>Korea University</b> BA in Political Science, Linguistics, and English	📍 Seoul, South Korea

## PROFESSIONAL EXPERIENCE

May 2019   Present	<b>Data Science Fellow and Consultant</b> <b>Data-intensive Social Sciences Lab</b>	📍 UC Berkeley
	<ul style="list-style-type: none"><li>• Consulted faculty, graduate students, and staff on R, Python, statistical modeling, data visualization, and computational text analysis</li><li>• Founded the working group on the bias and fairness in machine learning</li></ul>	
Spring 2020	<b>Data Science Education Program Fellow</b> <b>Data Science Education Program</b>	📍 UC Berkeley
	<ul style="list-style-type: none"><li>• Served as a research lead for the undergraduate students and project partners involved in 40+ data science discovery projects</li><li>• Designed workshops on project management, computational reproducibility, and bias and fairness in machine learning</li></ul>	

## TEACHING EXPERIENCE

Spring 2019	<b>An Introduction to Computational Tools and Techniques for Social Science Research</b> Lead Instructor	📍 UC Berkeley
	A graduate course on computational social science at UC Berkeley	

## CONTACT INFO

✉ [jaeyeonkim@berkeley.edu](mailto:jaeyeonkim@berkeley.edu)

🐙 [github.com/jaeyk](https://github.com/jaeyk)

in [linkedin.com/jae-yeon-kim](https://www.linkedin.com/in/jae-yeon-kim)

📞 +1 510-646-5183

For more information, please contact me via email.

## COURSEWORK

Statistical and Causal Inference, Experimental Design, Survey Methods, Game Theory, Computational Social Science

Passed [Political Behavior](#) (social and cognitive psychology, survey and experimental design) field exam with distinction

## SKILLS

Quantitative: Statistical and causal inference, Experimental and survey design

Computational: Computational text analysis, Machine learning, Web scraping, R (Advanced), Python (Intermediate), UNIX, Git, SQL (Familiar)

Qualitative: In-depth interviews, Archival research, Process tracing

Fall 2016

## Introduction to Empirical Analysis and Quantitative Methods

Teaching assistant

📍 UC Berkeley

An undergraduate course on statistical methods at UC Berkeley

Received [the best TA award](#) (given to less than 10% of Berkeley TAs)



## RESEARCH EXPERIENCE

Fall 2019

### Causal Inference and Machine Learning [\[GitHub\]](#)

PhD Candidate

📍 UC Berkeley

- Demonstrated how machine learning can help create critical data for causal inference by combining text classification and interrupted time series design
- Presented at [the 2019 Data Science Showcase](#) at UC Berkeley

2018

### Natural Language Processing and Machine Learning [\[GitHub\]](#)

PhD Candidate

📍 UC Berkeley

- Parsed unstructured historical newspaper articles (HTML files), turned them into a tidy dataset, and classified the text data using machine learning
- Selected to receive [the Best Paper Award in Asian Pacific American Politics](#) at the upcoming Western Political Science Association annual meeting

2016

|

2017

### Statistical Modeling of Time Series Data [\[GitHub\]](#)

PhD Candidate

📍 UC Berkeley

- Identified causal effects of the reduced government support on community organizing by collecting [original organizational data](#) and modeling time-series data
- Ensured the reliability of findings by doing robustness checks and running sensitivity analysis

2016

|

2017

### Experimental Research [\[GitHub\]](#)

Graduate student

📍 UC Berkeley

- Designed list experiments on sensitive political attitudes and embedded them in a nation-wide mobile survey experiment
- Reduced sampling bias by matching the online panel data with a preexisting probability sample before carrying out the survey

Summer

2018

### Survey Research [\[GitHub\]](#)

Graduate Student Researcher

📍 UC Berkeley

- Cleaned and wrangled a large-scale panel survey data, conducted multivariate and hierarchical regression analyses, and visualized results
- Imputed missing responses using multiple imputations and validated key survey constructs using factor analysis