

# Jingfei Cai, Ph.D.

I am a data scientist who can quickly adapt to various challenges and is motivated to acquire new skills. I am eager to combine data science techniques with my experience in scientific research and communication to catalyze data-driven decision-making and deliver high-quality work.

## SKILLS

**Data Science:** Python (including libraries such as numpy, pandas, sklearn, statsmodels, matplotlib, seaborn, requests, nltk), SQL, Tableau, R

**Development:** Git, Unix, Linux

**Web Development:** JavaScript, HTML, CSS, jQuery

**Cloud Computing:** Amazon Web Services, Google Cloud Platform

**Languages:** Fluent in English, Cantonese, and Mandarin (both spoken and written)

## EXPERIENCE

**General Assembly Data Science Immersive** | *Student* | Washington, DC | June 2017 – September 2017

Participated in a 12-week, full-time, project-based training program that focused on statistics, data analysis and visualization, and machine learning. Projects include:

- **Capstone: Topic Evolution in Life Sciences Research** (extracting data from XML files with lxml; Natural Language Processing; dynamic topic modeling)
- **Liquor Sales + Linear Regression** (linear models to predict sales; cross-validation to evaluate model performance; regularization tactics to compare model performance; executive summary)
- **Web Scraping & Classification** (web scraping; BeautifulSoup; classification models including random forest, logistic regression, k-nearest neighbors, and support vector machine; executive summary)
- **Predicting Terrorism/Bayesian Inference** (Markov Chain Monte Carlo; Bayesian inference; filling in missing data)

**The Journal of Visualized Experiments** | *Scientific Video Editor* | Cambridge, MA | June 2013 – April 2017

- Edited manuscripts and accompanying videos (over 1000) from various scientific disciplines (*e.g.* medicine, biology, neuroscience, immunology, behavior, bioengineering, engineering, chemistry, and environment)
- Addressed comments from authors and enhanced presentation of research methods and results
- Performed quality control on manuscripts/videos as a final step before publication

**LanguageLine Solutions** | *Over-the-phone Interpreter (contractor)* | August 2015 – June 2017

Interpreted conversations between Cantonese/Mandarin and English speakers (~500 hours), assisting clients in diverse fields including health care, emergency, government, law enforcement, court, and general business.

**Boston College, Department of Chemistry** | *Research Assistant* | Chestnut Hill, MA | July 2006 – April 2013

- Designed and implemented biochemical/biophysical research studies with focus on protein-membrane interactions
- Analyzed data collected from experiments and summarized research progress in forms of slides/poster presentations, written reports, abstracts to scientific conferences and manuscripts to journals

## EDUCATION & TRAINING

**General Assembly** | Washington, DC

*Data Science Immersive* | June – September 2017

**Boston College, Graduate School of Arts & Sciences** | Chestnut Hill, MA

*Ph.D. in Chemistry* | May 2013

Thesis Title: “Investigations of Lipid-specific Interactions of Lung Surfactant Proteins and EF-hand Domain of Mammalian Phosphatidylinositol-specific Phospholipase C”

**Peking University, College of Chemistry and Molecular Engineering** | Beijing, China

*Bachelor of Science in Chemistry* | July 2005