

# Wei (Thomas) Wang

New York, NY

(917) 558-7720

✉ [thomas.wei.wang@gmail.com](mailto:thomas.wei.wang@gmail.com)

🌐 [weiwang.io](http://weiwang.io)

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

🌐 [linkedin.com/in/thomasweiwang](https://www.linkedin.com/in/thomasweiwang)

---

## Skills

- **Languages**, proficient in Python and R; familiar with C/C++.
- **Tools & Applications**, MySQL, ggplot2, D3.js, Git, Linux, Hadoop, Emacs.
- **PhD Research Areas**, Bayesian Statistics, Causal Inference, Survey Methodology, Gaussian Processes, Graphical Models.

---

## Education

2009–2015 **PhD, Statistics**, *Columbia University*, New York, NY.

2005–2009 **BS, Statistics**, *University of Science and Technology of China*, Hefei, China.

---

## Experience

Jan–Mar, 2015 **Fellow**, *Insight Data Science Fellows Program*, New York, NY.

- Built *Curatir* (<http://curatir.weiwang.io>), an art recommender for New York City.
- Scraped major NYC museums' online catalogs and wikipedia articles of painters; Used TFIDF and SVD on wikipedia articles to analyze similarity between artists and extracted recurring themes for each artist; developed a web application based on Flask and Bootstrap and deployed it on EC2.

May–Aug, 2013 **Research Intern**, *Microsoft Research*, New York, NY.

- Developed methods to handle highly biased on-line opt-in survey collected on Xbox gaming platform for election predictions and delivered predictive power on par with traditional poll aggregators that rely on expensive polling data.
- Story was picked up by media including the *New York Times* and *New Scientist*.

Jul–Aug, 2011 **Consultant**, *World Bank*, Washington, DC.

- Built the R-based back-end of World Bank *Comparative Living Standards Project* web interface for interactive statistical analysis and visualization of World Bank survey database.
- Implemented standard statistical analysis tools such as Linear Regression, Logistic Regression, Scatter Plot and Histogram on a web interface.

May–Jul, 2011 **Consultant**, *International Finance Corporation*, Washington, DC.

- Built model for performance of developmental projects financed by the International Finance Corporation. Identified important factors in project performance using cross-validation and improved over the base model by 2%.

2009 - 2015 **Graduate Researcher**, *Columbia University*, New York, NY.

- Published three research papers on peer-reviewed journals on Bayesian hierarchical modeling of structured data and causal inference.
- Used R and Stan extensively.

2009 - 2015 **Graduate Teaching Fellow**, *Columbia University*, New York, NY.

- Served as graduate instructor twice for advanced undergraduate statistics courses.
- Served as teaching assistant for courses including Bayesian Data Analysis, Data Mining, Applied Data Science, Generalized Linear Models, Exploratory Data Analysis and Visualization
- Received 2011 Department of Statistics *Minghui Yu* Teaching Assistant Award.