

XINYAO WU

100 Haven Avenue, New York, NY 10032  
917-470-6513 [wuxinyao1996@gmail.com](mailto:wuxinyao1996@gmail.com)

## EDUCATION

---

**COLUMBIA UNIVERSITY, MAILMAN SCHOOL OF PUBLIC HEALTH**, New York, NY

**Master of Science in Biostatistics | GPA 3.79**

Aug 2018-May 2020

- **Courses:** Probability, Biostatistics Method I, Biostatistics Method II, SQL, Data Science, Epidemiology, Statistical Inference, Topics in Advanced Statistical Computing.

**CHINA AGRICULTURE UNIVERSITY**, Beijing, China

**Major in Agronomy and Biotechnology**

Sept 2014-June 2018

- **Courses:** Genetics, Biochemistry, Trail Design and Biostatistics, Molecular Biology, Epigenetics

## PROJECT EXPERIENCE

---

**Course Project about Markov chain Monte Carlo**, New York, NY

May 2019

- Randomly selected 80% data and develop an MCMC algorithm to estimate the posterior mean of the model parameters based on a given conditional probability model.
- Applied the model to track the remaining 20% hurricanes and evaluate the predictions.

**Course Project about bootstrapping**, New York, NY

May 2019

- Applied logistic-LASSO algorithm to select features for predicting.
- Compared the two approaches Lasso and bootstrap-smoothing in prediction.
- Explained the statistical justifications of my suggestion and implemented them in R.
- Carried out a CD optimization algorithm for a high-dimensional logistic model to predict breast cancer

**Course Project of Data Science**, New York, NY

Oct 2018-Dec 2018

- Filtered Design a data analysis study about the relationship between air qualities and asthma outcomes in NYC; Searched and collected different areas air pollution data and asthma data
- Visualized the relationship between different features and asthma rate with plots in R and made a website using shiny and flexdashboards.

## PROFESSIONAL EXPERIENCE

---

**Peking University, Research Assistant**, Beijing, China

Sep 2017-May 2018

- Build the construction of multiple regulation layers of *Arabidopsis thaliana* based on raw sequencing genetic data by Linux.
- Filtered and visualized gene-gene regulation network by weighted correlation network analysis with R.

**Peking University, Research Assistant**, Beijing, China

May 2017-Sep 2018

- Collected and sorted. The data and materials of international machine learning application
- Filtered Analyzed advantages and feasibility of artificial intelligence(AI) on future medical research

## SOCIETY EXPERIENCE

---

**Violinist**, Columbia University Medical Center Symphony Orchestra New York, NY Oct. 2018 –Present

**Volunteer, DatapiTHU (WeChat Subscription of Tsinghua University)**

Beijing, China, 2017

- Published the article Exclusive | *Teach you how to make a small program by scrapy(with the code attached)*

## SKILLS&CERTIFICATES

---

- **Computer Skills:** SQL, R, SAS, MS Office, Python
- **Certificates:** SAS Certified Advanced Programmer for SAS 9, SAS Certified Base Programmer for SAS 9
- **Coursera Course Certificate:** Machine Learning