

# WANYI HE

10 River Road, 17A, New York, NY 10044 | (646) 894-1184 | [dorahwy713@gmail.com](mailto:dorahwy713@gmail.com)  
Personal website: <http://wah4001.github.io> for more detailed information

## EDUCATION

### Cornell University – Weill Graduate School of Medical Sciences

Master of Science in Biostatistics and Data Science

New York, NY

GPA 3.85 | December 2019

### Purdue University

Bachelor of Science in Mathematics/Business

Bachelor of Science in Applied Statistics

• Minors in Management / Film and Video Study

**Honors:** *Dean's list* and *Semester Honors* for 4 years

West Lafayette, IN

GPA 3.64 | December 2017

## SPECIALIZED SKILLS

**Technical Skills:** R, SAS, Python, SQL, Microsoft Office (Access, Excel, PowerPoint, Word), REDCap

**Competencies:** Statistical Analysis and Modeling, Data Mining, Data Wrangling, Data Visualization and Interpretation

**Language:** Chinese (Mandarin)

## DATA SCIENCE EXPERIENCE

*Weill Cornell Medicine, New York, NY*

### ● Research Assistant

Jun. 2019 - Current

- Designed, implemented and maintained research database on Research Electronic Data Capture (REDCap) in the functionalities of relational database management system for the *EHE Retrospective Chart Review Study*
- Have designed 1,500+ variables in the form of survey questions
- Meet and communicate with the client monthly to update the detailed information of the database

### ● Biostatistician

May. 2019 – Aug. 2019

- Evaluated the racial and regional differences in Active Surveillance for low-risk prostate cancer in the USA
- Cleaned data from database that contains 200+ predictors and 310,000+ observations (363.4M)
- Collaborated with team to conduct multivariate logistic regression models and performed stratified analysis
- Performed sensitivity analysis to deal with 20,000+ unknown data in the outcome variable

### ● Statistical Programmer

- Collaborated with team to develop an **R package** for the NHANES database, aiming at efficiently cleansing, exploring and visualizing the data before analytics by designing 6 graphical and tabular data representation functions
- Individually developed a **Shiny App** by R for the Happy Valley Meat Company with the purpose of automatically generating an environmental impact report with the user-entered weight of meat consumptions
- Conducted data wrangling by **Pandas**, performed exploratory data analysis by **Matplotlib**, and implemented Hierarchical Agglomerative Clustering by **Scikit-Learn** in Python to identify the public school safety level in NYC

*Purdue University, West Lafayette, IN*

### ● Data Analyst | Dept. of Nutrition Science

Sept. 2017 – Dec. 2017

- Evaluated the effect of Supplemental Nutrition Assistance Program education on household food security status and dietary intake; Program sponsored by United States Department of Agriculture (USDA)
- Conducted high-dimensional nutritional data wrangling with 100+ variables by **SAS**
- Performed inferences with hypothesis testing and help with the statistical modeling
- Collected data by communicating with people about their characteristic living status and quantitative dietary intake

### ● Teaching Assistant | Dept. of Mathematics

Aug. 2015 – May. 2017

- Assist professors in grading linear algebra homework and quizzes for 160+ students
- Proctored math exams for all levels of math courses for 200+ hours

## FINANCIAL INDUSTRY EXPERIENCE

### Risk Management Internship | CredEx Fintech, LTD, Shenzhen, China

Jun. 2017 – Jul. 2017

- Participated in a start-up project for small enterprise loans
- Summarized best practices and collaborated with team to develop the score-card variable library (contained 700+ variables) for the subsequent risk rating and modeling

### Investment Internship | Shenzhen Capital Group CO., LTD, Shenzhen, China

Jun. 2016 – Aug. 2016

- Participated in the preliminary investment examination procedure by attending 30+ project-review sections
- Collaborated with cross-functional team to complete the meeting minutes and wrote summaries of investment projects with the purpose of providing references for subsequent follow-ups
- Understood Due Diligence Investigation and Feasibility analysis, including business analysis, equity structure, financial management, and business laws