

# YI YIN

APT.4 239W 105th New York, NY 10025  
646-402-1589 [yin.yi@columbia.edu](mailto:yin.yi@columbia.edu)

## EDUCATION

---

**Columbia University**, New York, NY *Expected Dec 2019*  
M.A. in Quantitative Methods in the Social Sciences, Cumulative GPA: 4.0  
Coursework: Applied Machine Learning, Natural Language Processing, Social Network Analysis,  
Data Visualization, Modern Data Structure

**The Ocean University of China**, Tsingtao, China *July 2017*  
B.A. in Journalism and Communication, top 3%, Cumulative GPA: 3.7  
Honors: National Scholarship- 2014, Outstanding Thesis- 2017  
Study Abroad- 2015 (Columbia University)  
Coursework: Brand Communication, Sample Surveys Analysis, Experiment Design

## WORK EXPERIENCE

---

**Institute for Social and Economic Research and Policy, Columbia University** New York, NY  
*Research Assistant (Supervisor: Dr. Donald P. Green)* *Feb 2019 – Present*

- Implement randomization inference using Stata to obtain exact p-values, and re-evaluated over 20 published works in political science and economics that rely on approximations.
- Convert examples and solutions in the textbook *Field Experiments: Design, Analysis, and Interpretation* written in R code to Stata code, and produce validated output used for the updated edition of the textbook.

*Research Assistant (Supervisor: Dr. Amy E. Semet)* *Feb 2019 – Present*

- Analyze legal text using Python, extract titles, authors and decisions from judicial court cases, predict authorship via inverse document frequency (ti-idf), and visualize text analysis result using R.

**Kantar Media CIC** Shanghai, China  
*Summer Analyst* *May 2018 – Aug 2018*

- Compiled Chinese social networking market statistics, built behavior model to explain user preference transformation, and pinpointed the potential and risks of Artificial Intelligence in Social Media, which was published on the Kantar Media's annual report *China Social Media Landscape 2018*.
- Analyzed the Kantar Media global consumer survey involving over 5,000 respondents using SPSS, compared advertise preference between internet user, newspaper reader, magazine reader and radio listener, highlighted the distinct advertise block behavior among Chinese users, which was presented in the Roadshow *Dimension 2018*.
- Organized the current company internal sharing knowledge into a tree structure, and conducted research on enterprise knowledge management solutions over LAN network, FTP, enterprise social network services, which helped the manager decide to use Microsoft SharePoint as internal knowledge sharing and archiving tool.

**International Growth Center** Yangon, Myanmar  
*Research Assistant* *Oct 2017 – Apr 2018*

- Cleaned data, checked covariates balance, built difference-in-difference model, conducted sub-group analysis, and visualized results using R for the pilot survey experiment involving 952 subjects, which was presented in Quantitative China Studies Seminar (QCSS) in New York City.
- Analyzed full-scale experiment involving 1910 subjects, included in the International Growth Center policy brief, and produced illustrations, presented in the regional study filed research seminar in Shanghai.

## SKILLS

---

*Language:* Mandarin (Native), Japanese (Basic)  
*Programming:* Fluent in R, Python, Stata; Basic in SQL and SPSS

# 尹伊

yin.yi@columbia.edu +1 (646) 402-1589



## 教育背景

哥伦比亚大学

预计 2020 年 2 月毕业

硕士：社会科学定量方法（数据科学方向）GPA: 4.0

核心课程：机器学习应用，自然语言处理，社会网络分析，数据可视化，数据结构，定量研究方法

中国海洋大学

2013 年 - 2017 年

本科：新闻学（年级前三）优秀毕业论文：媒介使用与政治倾向的研究

核心课程：品牌传播，危机传播管理，公共关系学，新闻采访与写作，社会调查设计与分析，社会科学实验设计

## 实习经历

哥伦比亚大学 社会经济政策研究所 研究助理

2019 年 2 月至今

- 用蒙特卡洛法模拟实验随机分配过程，通过实验结果的概率分布计算精确 p 值，校准 20 余篇已发表的社会科学实验论文中通过回归模型估算的 p 值。将教材《实地实验：设计、分析与解释》中用 R 语言编写的案例和课后习题答案改写为 Stata 代码，为新版教材进行数据整理和代码编译、校对。
- 用 python 的自然语言处理模块提取法庭文献中的标题、法官、庭审决议等信息，通过文档词频预测文献作者。用 R 语言 ggplot2 模块将模型可视化。

Kantar Media CIC 战略智库研究中心 实习分析师

2018 年 5 月 - 2018 年 8 月

- 收集中国社交媒体行业数据，归纳用户使用社交媒体的偏好，预测品牌社交营销趋势，撰写年度白皮书《2018 年中国社会化媒体生态概览》。编辑月报 Trend Watch 科技前瞻专栏的人工智能主题文章。
- 用 SPSS 分析超过 5000 名受访者的全球消费者调查，对比新媒体用户和传统媒体用户的广告接受偏好，总结中国用户独有的隐私让渡和广告屏蔽习惯，为 2018 年 Dimension 路演上海场提供数据模型支持。
- 协助直线经理组织每两周一次的公司骨干培训会，为公司内部分享会制作 PPT，为战略智库研究中心和咨询部、大数据部、市场部等跨部门会议记录会议纪要。将公司内部分享的知识报告整理成树形结构。调查对比不同的企业知识分享工具（局域网、FTP、企业云社交服务等），帮助直线经理决定内部知识分享工具。

国际发展中心 研究助理

2017 年 10 月 - 2018 年 4 月

- 使用 R 语言对包含 952 名受访者的先导调查实验数据进行清理、建模分析和可视化呈现，分析模型在 2018 年纽约中国定量研究研讨会上展示。分析 1910 名缅甸公民参与的田野调查实验，为国际发展中心的缅甸地区政策简报提供数据支持，为课题在 2018 年“区域国别研究的田野调查案例”学术研讨会的展示提供可视化模型。

## 语言技能

英语：雅思 7.5，GRE 329（语文超过全球 90% 的考生），六级 589

编程：熟练运用 R 进行数据可视化；熟练运用 Python 建立机器学习、自然语言处理模型；掌握 SQL，Stata，SPSS

## 获奖情况

中国海洋大学 优秀毕业论文

2017 年

中国海洋大学 第一位前往哥伦比亚大学访学的本科生

2015 年 - 2016 年

教育部 国家奖学金

2014 年