

Education

Communication University of China

Data Journalism, 3.8/4.0

Beijing - China

September 2017 – June 2021

- Top 5% with excellent performance in Calculus, Linear Algebra, Probability, Statistics and various courses in programming.

University of California, Los Angeles

Summer Session, *Stats100C: Linear Models*, 4.0/4.0

Online

July 2020 – September 2020

- Top 20% in a proof-based course on the theory of regression analysis. Topics include Gauss Markov theorem, general aspects of maximum likelihood estimation, partial regression, constrained estimation, generalized least squares, and model selection.

Skills in Programming and Computation

My featured projects include a [case study](#) on the most up-to-date and comprehensive Titanic survival data, a [text classification model](#) differentiating Brontë Sisters, and another [case study](#) using Python and R in parallel for data tidying and visualization. I am writing [notes](#) in mathematics for data science and statistics. Please visit my [website](#) to see a full list of personal projects.

- [website](#) **R, Python and web development:** Experienced in data analysis with R's tidyverse, data.table and Python's pandas, ranging from data collection, cleaning, feature engineering, statistical modeling and data visualization. I have created my own miscellaneous R package available [here](#). I am also familiar with HTML, CSS and JavaScript (vanilla and Vue.js), having authored several static sites and Shiny apps.
- [templates](#) **Git, R Markdown and \LaTeX :** These tools are used extensively in all my projects for version control and reproducibility. I have taught them to new members in the research group for scientific publication.

Professional Experience (Internship)

Data Analyst

51Talk, Department of Big Data and Business Intelligence

September 2020 – Now

- At a leading K-12 online English education company, I am involved in establishing an integrated tutoring performance system. With R and Python, I conducted extensive exploratory analysis and research on the influence of hardware setup, speaking speed, facial expressions and time management on conversion rate and customer review. I am involved throughout the project in problem definition, data cleaning, feature engineering, regression modeling and the development of end-to-end web applications.
- Optimized adaptive testing system for placement, based on item-response theory and deep knowledge tracing.

Business Analyst

Analysys, Department of Data Analysis and General User Success

June 2020 – September 2020

- Innovated our web and app analytics service in aspects of data visualization tools and user interface.
- Provided supports for users across industries by designing tracking events and analysis of user behavioural data.
- Translated numbers into interactive reports and dashboards written in R Markdown and JavaScript, receiving immense positive feedback from client management.

Data Mining Engineer

DiDi Chuxing, International Division

March 2019 – August 2019

- Contributed to building fast, scalable and extensible AB testing environment with R and C++.
- Conducted overseas geo-marketing research targeting South America and Australia with techniques like density-based clustering, logistic regression, and correspondence analysis.

Academic Achievements and Awards

- 2020 **Changing Frontiers: 20 Years of China's Communication Study:** An manuscript in preparation probing China's trending research focus and academic network in Mass Communication study over the last 20 years. Investigated the abstract sections in graduate dissertations with dynamic topic modeling and citation network analysis.
- 1st Place in China International College "Internet +" Innovation and Entrepreneurship Competition:** A content analysis of misleading news coverage on COVID-19 and its effects on public health crisis. 1 out of 200 by faculty votes.
- Best Visualization and Creativity Award in China Data Journalism Competition:** An [interactive data-driven report](#) on COVID-19's impact on China's Film Industry.
- 2019 **1st Place Team in Nanfengchuang Social Research Contest:** A quantitative analysis of aquatic biomass and water pollution in two of China's largest freshwater lakes, to assess the effectiveness of the "River Chief" policy.