

WILLIAM ZHU

wzhu4@outlook.com | Chicago, IL 60637 | 610-517-3106 | [LinkedIn](#) | [GitHub Website](#)

Education

The University of Chicago (Chicago, IL)

MA Computational Social Science (STEM), Quadrangle Scholar, GPA: 3.8/4.0

June 2022 (Expected)

Relevant Courses: Python Programming, Machine Learning, Deep Learning, Text Analysis, Algorithms, Database, Large-Scale Computing, Competitive Strategy (Booth), Consumer Behavior (Booth)

Haverford College (Haverford, PA)

BA Sociology, Minor in Statistics, GPA: 3.73/4.0

May 2019

Work Experience

RealEats (Geneva, NY)

Data Analytics Consultant (part-time, remote)

September 2021 – present

- Extracting insights in consumer behavior for a weekly meal subscription service by compiling a customer database on AWS Redshift, building dashboards on Google Data Studio, and leading survey research design

Data Science Intern (remote)

June 2021 – September 2021

- Improved the 2-months customer lifetime value (LTV) by 4% by identifying promo codes associated with high average order value (AOV) and customer retention rates
- Saved 20 hours of CFO's time per month by automating Python scripts on AWS Sagemaker and building customized cohort analysis dashboards on Google Data Studio
- Provided strategic support to CMO by identifying top customers attributes using logistic regressions and decision trees; compiled a customer dataset using Python and SQL from sources including sales records, Google Analytics API, U.S. Census, gender-guesser Python package, and Zillow home value index
- Presented findings at 6 meetings to the CFO, CMO, VP of Operations, and Director of Growth

Westat (Rockville, MD)

Research Assistant (full-time)

July 2019 – June 2020

- Received the highest performance rating for all 8 evaluation projects by performing data analysis in R and report writings for clients including U.S. Department of Education, Verizon, and National Science Foundation
- Co-authored a [journal article](#) titled "Using state datasets and meta-analysis of low-powered studies to evaluate a school-based dropout prevention program for students with disabilities" in *Studies in Educational Evaluation*
- Co-authored an [evaluation report](#) to show that Racial Equity Action Leadership (REAL) Program is effective at guiding 30+ participants to craft company-wide racial equity plan via 8 participant observations and 10 surveys

Projects

Exploring Variations in Divvy Bike Station Usage Volume [\[5-min presentation, blog post\]](#)

March 2021 – June 2021

- Demonstrated that public bike share stations' usage rate are strongly linked to crime rates, socio-economic status, network effect and demography using linear regressions with log transformations; compiled a station-level dataset from 7 databases including 300K Divvy bike trip records, 210K Chicago crime records, 11K bus stops records, U.S. Census, and Zillow home value index

Uncovering Links between Cultural Values and Company Directions [\[5-min presentation\]](#)

March 2021 – June 2021

- Demonstrated that firms with an agile culture are more likely to initiate share buyback, attract private placement, and receive M&A offers within one year by measuring keyword frequencies of 4 million Glassdoor employee reviews from 6K firms in the U.S.; applied logistic regressions to control for firm size, location, and average ratings
- Matched 20K company names between S&P Capital IQ and Glassdoor employee review databases by constructing a customized string similarity measure based on the fuzzywuzzy Python package

Skills

- Python, R, SQL, AWS (Redshift, Sagemaker), Google Analytics, Data Studio, Git/Github, Latex, HTML/CSS
- Linear and Logistic Regressions, KNN, Random Forest, PCA, clustering, Cross-Validation, Model selection
- Hypothesis Testing, Cohort Analysis, Customer Segmentation, Network Analysis, Text Analysis