

WILLIAM ZHU

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

Professional Summary

Data Analytics Consultant with 2 years of experience in food tech e-commerce and survey research. Skilled at extracting and communicating business insights from complex consumer data.

Tools: Python (pandas, numpy, statsmodel, sklearn, matplotlib, seaborn, networkx, nltk, spacy, gensim)

SQL, R, PySpark, Dask, Google Data Studio, Google Analytics, Git/Github, HTML/CSS

AWS (Sagemaker, EC2, EMR, Lambda, Redshift, DynamoDB, S3, Kinesis)

Machine Learning: Linear and Logistic Regressions, Random Forest, Boosting, KNN, Model Selection, PCA, Clustering

Analytical Methods: Hypothesis Testing, Cohort Analysis, Network Analysis, Time-series Forecasting, Text Analysis

Education

The University of Chicago (Chicago, IL)

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

June 2022 (Expected)

Courses in Python Programming, Machine Learning, Deep Learning, Algorithms, Databases, Natural Language Processing, Social Network Analysis, Large Scale Computing

Marketing Concentration at Chicago Booth School of Business

Courses in Competitive Strategy, Marketing Strategy, Pricing Strategy, Consumer Behavior, Experimental Marketing

Haverford College (Haverford, PA)

BA Sociology, Minor in Statistics, GPA: 3.73/4.0, GRE: Verbal: 165/170, Quantitative: 167/170, Writing: 5/6

May 2019

Work Experience

RealEats (Geneva, NY)

Data Analytics Consultant

September 2021 – present

- Drove analytics insight that led to \$16 million Series A funding by automating customer retention reports for CFO, investors, and board of directors (Python and SQL on AWS Sagemaker and Redshift)
- Saving 6 hours of CMO and Director of Marketing's time every week by building dashboards that track promo code acquisition and retention performance (SQL and Google Data Studio)

Data Science Intern

June 2021 – September 2021

- Improved 2-month customer lifetime value (LTV) by 7% by identifying and implementing promo codes associated with high Average Order Value (AOV) and low churn rate (Python and SQL)
- Solidified brand positioning strategy by identifying top customers attributes using logistic regressions and decision trees; compiled a customer dataset from sources including sales records, Google Analytics API, U.S. Census, gender-guesser Python package, and Zillow home value index (Python and SQL)
- Delivered 8 presentations in front of the CEO, CFO, CMO, and VP of Operations; attended weekly strategic meetings

Westat (Rockville, MD)

Research Assistant

July 2019 – June 2020

- Received the highest performance rating for all 8 evaluation projects by performing survey data analysis and report writings for clients including the U.S. Department of Education, Verizon, and National Science Foundation (R, Excel)
- Co-authored an [evaluation report](#) that led to \$260K in grant funding for Racial Equity Action Leadership Program by showing that the program is effective at guiding 38 participants to craft company-wide racial equity plan through 8 participant observations, 2 focus groups, and 10 surveys
- Co-authored a peer-reviewed [journal article](#) titled "Using state data sets and meta-analysis of low-powered studies to evaluate a school-based dropout prevention program for students with disabilities." in *Studies in Educational Evaluation*

Projects

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

March 2021 – June 2021

- Discovered associations between Chicago's public bike stations' usage volume and crime rates, socio-economic status, network effect and demography by merging and analyzing 300K Divvy bike trip records, 210K Chicago crime records, 11K bus stops records, U.S. Census, and Zillow home value index (Python pandas, networkx, statsmodels)

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

March 2021 – June 2021

- Showed that firms with an agile culture are more likely to receive M&A and investment offers using keyword frequencies of 4 Million Glassdoor employee text reviews from 6K firms in the U.S. (Python pandas, statsmodel, nltk, spacy, fuzzywuzzy)