

Xinge Hu

xh1051@nyu.edu • (617) 331-5722 • <https://www.linkedin.com/in/xinge-hu-5336b5155/>

EDUCATION

NEW YORK UNIVERSITY

MS in Data Science GPA: 3.76/4.00

- Relevant Coursework: Machine Learning, Probability & Statistics, Linear Algebra, Big Data, NLP

New York, NY, USA
Sep. 2021 – May 2023 (Expected)

BOSTON UNIVERSITY

BA in Statistics, Minor in Computer Science GPA: 3.79/4.00

Boston, MA, USA
Sep. 2017 - Jan. 2021

EXPERIENCE

Intellipro Group

Data Analyst Intern

New York, NY, USA

Jun. 2022 - Aug. 2022

- Performed descriptive analysis in **SQL** to analyze 1M+ students' information including major, gender, email, school location, etc, generated infographics using **Tableau**, and presented the data visualization results to company executives bi-weekly.
- Conducted data quality analysis and pre-processed modeling data by performing data cleaning and data transformation.
- Performed web scraping using **Python** Requests and Selenium packages, and expanded the company's database by 30,000+ records.

BETH ISRAEL DEACONESS MEDICAL CENTER

Data Science Research Intern

Boston, MA, USA

Jul. 2020 - Nov. 2020

- Collaborated with team's biologists, and performed statistical analysis to compare patients' protein data from different sources using **R**.
- Developed a Machine Learning classification model with Python package **scikit-learn** using **Logistic Regression** with L1 Regularization and Sequential Feature Selection with sigmoid SVM.
- Identified 20 most relevant proteins to surgery-caused delirium out of 1300 proteins, as well as signature proteins that distinguish different levels of delirium, and increased the overall **model prediction accuracy** from 70% to 89%.

BOSTON UNIVERSITY CENTER FOR CAREER DEVELOPMENT

Data Analyst Intern

Boston, MA, USA

Jan. 2020 – May 2020

- Conducted data cleaning and data transformations on survey results from 4000 college graduates.
- Developed a linear model that predicted college degree holders' salaries based on their demographics.
- Designed a User Interface to visualize salary prediction and distribution using **R Shiny** to provide better career consulting service.
- Presented data results to university executives and colleagues, and offered recommendations for data collection and survey design.

FINTECH GLOBAL

Research Analyst Intern

London, UK

Feb. 2019 - Apr. 2019

- Collected 12000+ global FinTech investments into MS **Excel**, doubling the size of the company's database.
- Analyzed investments and composed detailed analytical articles about FinTech subsectors on the company's research blog weekly.
- Collaborated with a 10-person team to organize a 3-day Global RegTech Summit attended by 500 industry professionals.

PROJECTS

Two-sided Market A/B Test Experiment Analysis (Metric Design, Experiment Evaluation)

Jun. 2022

- Performed **z-test** to evaluate key metric of Instacart shopper hiring funnel and validated the effectiveness of earlier background check.
- Analyzed the cost per acquisition(CPA) of multiple acquisition channels and provided marketing solutions for leads generation.
- Provided executable plans based on shopper hiring analysis to find bottlenecks and insights using **Python** and **Excel**.
- Recommended marketing strategy by analyzing the conversion rate and cost per acquisition of social media, referral, and job search.

Web User Activity Analysis (User Journey Analysis, Funnel Analysis)

May 2022

- Used **funnel analysis**, **cohort analysis** and **segmentation analysis** to acquire the reasons for decreased user email login rate.
- Analyzed Email campaign performance such as open rate and click through rate to diagnose user activity moving trend.
- Wrote **SQL queries** to impute retention rate and analyzed the moving pattern in the user engagement in **PostgreSQL**.
- Built **Tableau** dashboards to present the tendency in user engagement over time and breakdown by user type dynamically.
- Developed a methodology framework to provide practical recommendations as summarization.

Credit Card Users Segmentation (Unsupervised Learning, K-means)

Feb. 2022

- Conducted exploratory data analysis for geographic, demographic and transaction data from 10K+ credit card holders.
- Applied **PCA** and **normalization** for feature engineering, reducing 40+ features into 10+ features.
- Clustered users into five clusters with different models including **K-Means** and **RFM models**.
- Applied elbow method to identify the optimal number of clusters for K-Means models.
- Conducted profiling analysis for each cluster to interpret the characteristics and main driven features.

SKILLS

Data Analysis SQL | Python | Excel | R | Statistics

Data Visualization Tableau

Database PostgreSQL | MySQL

Machine Learning TensorFlow | PyTorch | Scikit-learn