

Ziqi Xu

+1 (217)979-0662 | ziqi12@illinois.edu | www.linkedin.com/in/ziqi12 | https://github.com/ziqixu22

EDUCATION

- University of Illinois Urbana-Champaign** Urbana, IL
Bachelor of Science - Statistics and Mathematics; Minor - Computer Science; GPA: 3.93 May 2026
Core Courses: Intro to Computer Science, Discrete Structures, Statistical Modeling, Time Series Analysis, Data Science

SKILLS SUMMARY

- Languages:** Python, C/C++, Java, JavaScript, R, SQL, CSS/HTML
- Tools:** Tableau, TensorFlow, JAX, Docker, Git, Power BI, AWS, Excel, Powerpoint, GeoPandas
- Soft Skills:** Leadership, Teamwork, Communication, Event & Time Management, Writing, Creativity, Problem-solving

WORK EXPERIENCE

- National Center for Supercomputing Applications** Urbana, IL
Undergraduate Researcher June 2024 - Present
 - Leveraged advanced statistical and machine learning techniques (e.g., **Lasso Regression**, **Random Forests**) to evaluate feature importance and optimize predictive accuracy using metrics such as Mean Squared Error (MSE) and R-squared
 - Applied spatial analysis tools, including **GeoPandas** and spatial regression models, to investigate distribution patterns
- MyEdMaster LLC** Remote
Data Scientist Intern January 2024 - May 2024
 - Conducted literature reviews by examining 10+ diabetes-relevant research and 2+ public databases to understand important features for determining the profitability of the health industry
 - Trained **Random Forest model** to predict individual health score over 10+ features (dietary supplements frequency, types), achieving RMSE of 0.98, utilized by Product and Market team to improve customer purchase frequency by 4%
- ILLINOIS ATLAS Teaching and Learning with Technology** Urbana, IL
Data Analyst August 2023 - December 2023
 - Designed a cloud dashboard with 40+ visualizations (geospatial maps, line charts), by **Tableau** to analyze ethnic diversity in 5 states and 6 universities, utilized by UIUC across ATLAS data services team to monitor the demographic trends
 - Built data **ETL** with **Python** (Pandas, Numpy) and **Excel** (VLOOKUP, INDEX MATCH) to gather data from 5+ sources, cleaned the data by standardizing, and removed duplicates, reducing manual effort by 30%
- Illinois Mathematics Lab** Urbana, IL
Undergraduate Researcher August 2023 - December 2023
 - Developed a **Python** function to construct phylogenetic trees from a NumPy faculty distance matrix, utilizing hierarchical clustering algorithms, and generated intuitive dendrogram visualizations for clear, impactful analysis
 - Drafted reports conducting a comprehensive evaluation of the resulting phylogenetic trees and associated clusterings to ensure alignment with the established departmental structure, ensuring accuracy and organizational coherence
 - Found the barycenter by applying the Geodesic Treetop Problem (GTP) algorithm and the Python package PATHTREES to identify geodesics between three trees, optimizing branch lengths through an iterative circular looping method

PROJECTS

- ZRX-USD Exchange Rate Analysis and Prediction | Team Leader**
 - Examined ZRX-USD Exchange Rate using an **ARIMA** time-series model on past returns and utilized VAR to assess risk
 - Leveraged sentiment analysis by **NLTK** on relevant news to improve Python algorithm accuracy by 1.3%
- Illinois Statistics Datathon | Team Leader**
 - Cleaned 5 million + data by removing top 5% outliers and standardized data with **Python (pandas)**
 - Selected 7 out of 10 features, based on correlation coefficients and trained a **multiple linear regression** to predict credit card charge-offs for 12 months, achieving RMSE 4.89
 - Utilized **K-Means** model to subgroup users from historical user data, studied the influencing factors and calculated macro-economics coefficients with least squares matrix to predict the charge-offs
- Elevation Path Analysis and Visualization | Individual Project**
 - Utilized **C++** to manage file I/O operations, dynamically allocating arrays for data handling and visualization in grayscale
 - Implemented **greedy algorithms** to interpret elevation data, generating images to visualize least-elevation-change paths
 - Wrote unit tests to validate all functionalities, ensuring the application while practicing test-driven development (TDD)