

Zeran Ni (Simon)

📍 10 Waterside Plz ✉ zn2133@nyu.edu 📞 332 276 9551 👤 simon-ni

Education

New York University, Center For Data Science, College of Arts and Science *Sep 2023 – May 2027*

Major: Computer Science and Data Science

GPA: 3.94/4.0

Relevant Coursework: Data Structures and Algorithms, Computer System Organization, Causal Inference

Experience

Class Tutor and Grader

Department of Computer Science, New York University

- New York, NY*
Jan 2024 – Present
- Led **3** hours of in-class tutoring each week, supporting **200+** students during lectures and labs with assignments and debugging.
 - Held **7** hours of weekly office hours, assisting students with Python programming (data types, control flow, OOP, File I/O, and core data structures).
 - Reported common student challenges to the instructor and assisted in improving assignment design.
 - Graded **11** programming assignments for ~**50** students, delivering detailed and constructive feedback.

Projects

AI-Assisted Autograder for Introductory CS Course at NYU

[GitHub Repository](#) ↗

- Designed an AI-powered autograder evaluating 11 programming assignments for 100+ students, reducing grading time by over 50% by integrating OpenAI GPT models via LangChain to assess code quality, formatting standards, and programming style beyond traditional unit tests.
- Engineered dynamic testing pipeline using `subprocess` for code execution, `re` for output validation, and custom parsers to extract numerical values from formatted tables, with adaptive input handling that automatically feeds test cases and flags edge cases requiring human review.
- Architected modular grading system with reusable utility functions and assignment-specific graders, combining static code analysis with runtime evaluation to overcome Gradescope's function-call dependency limitations and assess early assignments before students learned advanced programming concepts.
- Technologies Used: Python, OpenAI API, LangChain, Bash, Regular Expression, Git, VS Code.

NGrams – (Based on UC Berkeley CS61B: Data Structures)

[GitHub Repository](#) ↗

- Audited and independently completed UC Berkeley's CS61B Project 2, implementing NGrams, a tool for analyzing word frequency trends over time using the Google N-Grams dataset.
- Designed and implemented efficient data structures in Java (`HashMap`, `TreeMap`) to store and query millions of n-gram records, enabling fast retrieval of frequency counts across specified time ranges.
- Built methods to generate temporal trends and compute aggregate statistics, supporting queries on multiple words simultaneously.
- Technologies Used: Java, `HashMap`, `TreeMap`, JUnit, Git, IntelliJ.

Data Science Capstone: Professor Rating Analysis

[GitHub Repository](#) ↗

- Cleaned and processed RateMyProfessor data with NumPy and Pandas to explore how gender, experience, and course difficulty affect professor ratings.
- Uncovered measurable gender bias using hypothesis testing, with male professors rated significantly higher ($p < 0.005$). Built predictive models using Linear, Ridge, and logistic regression (AUROC=0.80) and improved interpretability with PCA.
- Delivered insights through clear visualizations and performance reports using Matplotlib and Scikit-learn.
- Technologies Used: Python, NumPy, Pandas, SciPy, Scikit-learn, Matplotlib.

Technologies

Languages: Java, Python, JavaScript (ES6+), C, SQL

Libraries: NumPy, Pandas, Scikit-learn, LangChain

Tools & Platforms: Git, VS Code, Anaconda, Spyder, Jupyter Notebook, IntelliJ IDEA, HTML5, CSS3, MicroSoft Word, Powerpoint, Excel