

Jason Chen

Boston, MA | 718-866-7738 | jachen7@bu.edu | [Personal Website](#)

Skills:

Programming Languages: Python, SQL, R, Rust, HTML/CSS/JavaScript

Frameworks & Tools: Git, ArcGIS, Power BI, Tableau, Microsoft Azure, MySQL, PostgreSQL, Google BigQuery, Looker Studio, Docker, LaTeX

Professional Experience:

Second Nature, *Data Science Intern* – Boston, MA

January 2025 - Present

- Streamlined carbon emissions reporting by transforming unstructured goal data into structured formats, enabling efficient tracking of 193 goals across 119 institutions and improving data usability by **23%**.
- Optimized outlier detection and data cleaning processes, reducing data inconsistencies by **35%** and ensuring accurate analysis of Scope 1, 2, and 3 emissions for improved decision-making in sustainability strategies.
- Visualized regional emission hotspots via ArcGIS spatial analysis, revealing a **28%** faster progress lag in Southeastern states post-COVID and prompting state-level policy advocacy for renewable energy incentives.

Kolachalama Lab, *Machine Learning Research Assistant* – Boston, MA

September 2024 – Present

- Engineered a transformer-based AI model to analyze real-world EHR data, optimizing masking strategies to address **40%** missing data inconsistencies, achieving **92%** accuracy in early dementia detection.
- Conducted parameter experimentation, evaluating masking strategies across 30 features, which reduced diagnostic delays by **27%** and improved AUC scores by **20%** across diverse datasets.
- Synthesized insights from machine learning research papers to design a novel permutation-masking hybrid technique, accelerating inference speed by **1.7x**.

DC City Government, *Data Visualization Intern* – Remote

September 2024 – December 2024

- Automated data extraction by leveraging Python, SQL, and Google BigQuery to process over 500k records monthly that resulted in **18%** boost in operational efficiency for service request analysis.
- Pinpointed high-demand service areas using geographic data, identifying hotspots for repair requests, which led to the implementation of a preventative maintenance schedule, decreasing complaints by **20%**.
- Built interactive dashboards to track 311 service request response times and volumes, revealing the most unresolved requests and requests counts, enabling resource allocation to reduce backlogs by **22%**.

National Science Foundation, *CEET REU Research Assistant* – Amherst, MA

June 2024 – August 2024

- Designed an interactive visualization tool showing real-time carbon intensity, enabling users to optimize energy usage, resulting in a **43%** reduction in carbon emissions for participating regions.
- Developed Python scripts to automate data merging, daily file extraction, and missing data imputation, reducing manual processing time by **30%** and ensuring seamless integration across **50+** regions.
- Identified and implemented new evaluation metrics to assess model performance, enhancing forecast reliability and reducing discrepancies in carbon intensity predictions by **15%**.

Authentic Caribbean Foundation Inc, *Data Analyst Intern* – Boston, MA

November 2023 – February 2024

- Analyzed **10K+** monthly restaurant transactions using Python and Excel, pinpointing underperforming menu items and driving a **13%** sales increase through targeted promotions.
- Mapped U.S. Census income and population data to restaurant locations, identifying 3 high-potential suburban markets that generated **\$1.2M** in first-year revenue post-expansion.

Project Experience:

Handwriting Recognition (Python) – *Data Science Lead*

- Preprocessed and augmented a dataset of **10,000** handwritten word images, enhancing model generalization and reducing validation loss by **20%**.
- Evaluated model performance using industry-standard metrics, achieving a CER of 5% and WER of 10%, surpassing baseline benchmarks by **25%**.

Education:

Boston University

Expected: January 2026

BS in Data Science; Minor in Business Administration & Management

Courses: Data Visualization; Data Science; Data Engineering; Data Privacy; ML & AI; Algorithms