Stacy Che

(980)-339-6330 stacy.c2000@gmail.com Stacy Che staccjch

Education

Georgetown University **Expected May 2026**
M.S. in Data Science and Analytics; GPA: 4.00/4.00 *Washington D.C.*
Coursework: Statistical Computing, Database Management & SQL, Data Science Fundamentals

University of Washington **Aug. 2023**
B.A. Economics and Education; Minor in Data Science; GPA: 3.76/4.00 *Seattle, WA*
Coursework: International Finance, Econometric Applications, Population & Development

Technical Skills

Programming Languages: R, Python, SQL, HTML, CSS
Data Visulization: Matplotlib, Seaborn, ggplot2, Shiny, Latex, Quarto
Data Science & Machine Learning: Excel, Python(pandas, numpy, statsmodel, scikit-learn, matplotlib), ETL, Data Science pipeline (cleaning, wrangling, visulization, modeling, analyzing), Git, Hypothesis Testing, Regression & Classification

Experience

Zoomlion Agriculture Machinery **Jun 2025 – Aug 2025**
Data Analyst Intern *Beijing, China*

- Improved **XGBoost models** for maize fungus infection prediction by adding **cross-validation and robust feature engineering**, raising generalization accuracy on non-lab field data by **20%**.
- Delivered optimized **modeling code** to the software engineering team, supporting deployment in a farmer-facing **mini-app** for forecasting and agricultural practice guidance.
- Benchmarked **Chinese vs. international calibration standards** for soil moisture sensors, developing reproducible workflows that streamlined validation by **40%**.
- Built **visual analytics dashboards** (Python, Plotly) to communicate ML outputs and sensor insights, enabling agronomists and engineers to interpret results for R&D and product design.

McCourt School of Public Policy **Feb. 2024 – Present**
Graduate Research Assistant – Colombian-Venezuelan Migration & Media Analysis *Washington, D.C.*

- Designed and implemented large-scale **web scraping pipelines** to collect Venezuelan migration coverage from 10+ major Latin American media outlets (e.g., El Tiempo, El Nacional, El Colombiano, El Pitazo, La Silla Vacía, Efecto Cocuyo, La República).
- Automated data extraction from **dynamic websites and APIs** (Next.js JSON objects, WordPress Jetpack API, Selenium-based crawlers), retrieving **5,000+ articles and social media posts**.
- Optimized pipelines with **parallelization (ThreadPoolExecutor)** and robust error handling, improving throughput by **3x** compared to sequential scrapers.
- Built reproducible **data pipelines in Python and R** (BeautifulSoup, requests, pandas, tidyverse), integrating structured metadata (date, section, URL) with full-text content for downstream analysis.
- Conducted **narrative and sentiment analysis** across media sources, producing visualizations of discourse trends to support research on **migration attitudes and exposure effects**.

University of Michigan Ross School of Business **Oct. 2023 – Apr. 2024**
Economic Research Assistant *Ann Arbor, MI*

- Analyzed decades of U.S. childcare regulations to identify tightening policy trends and potential impacts on childcare supply.
- Cleaned and integrated **50+ datasets** (regulations, center counts, demographics) into a panel dataset for econometric analysis.
- Created **visualizations and statistical summaries** linking regulation changes to childcare availability, supporting policy-focused research outputs.

Projects

Whose Civilization Is On Display? – MET Dashboard Project | *R Shiny, Data Visualization, UX Design* **Spring 2025**

* Designed a multi-tab R Shiny dashboard visualizing over 2 million MET artifacts, with interactive maps, acquisition timelines, treemaps, and word clouds exploring representation across four ancient civilizations.
* Led the UX design of the dashboard, optimizing layout, navigation, and visual hierarchy to support storytelling and user engagement.
* Developed narrative visualizations focused on Ancient Egypt, including dynastic timelines, material classifications, and artifact acquisition patterns.
* Refined team visualizations for clarity and cohesion, ensuring alignment with the dashboard's curatorial themes and historical insights.

Disney Princess Movie Sentiment Analysis | *Python (Selenium, NLP, Scikit-learn), Quarto, CSS* **Winter 2024**

* Built a Selenium WebDriver scraper to extract and clean **50,000+ movie reviews** from Rotten Tomatoes, handling **pagination, dynamic content**, and regex-based text processing.
* Developed an interactive **Quarto & CSS website**, visualizing **cultural and gender sentiment trends** in Disney Princess movies.
* Applied **NLP and supervised machine learning** to analyze **cultural and feminist sentiment**, leveraging text-analysis techniques on audience reviews.