

Sebastian Capellan

New York, NY, 10003 • scapellan55@gmail.com • [linkedin.com/in/sebacape/](https://www.linkedin.com/in/sebacape/) • github.com/SebaCape • sebacape.github.io

EDUCATION

New York University

New York, NY

B.A. Data Science & Computer Science, GPA: 3.6

Expected May 2028

Relevant Coursework: Data Management, Operating Systems, Systems Programming, Computer Architecture, Data Structures, Data Science, Open Source Software Development, Basic Algorithms, Multivariable Calculus, Discrete Math, Linear Algebra

TECHNICAL SKILLS

Programming Languages: Python, SQL, C, Java, HTML, CSS

Technologies: MySQL, DuckDB, Linux, Docker, AWS, Git, GitHub

Libraries & Frameworks: Spring, NumPy, Pandas, Numba, Pytest

WORK EXPERIENCE

Duolingo

Pittsburgh, PA

Incoming Software Engineer Intern

June 2026

University of California, San Francisco

Remote

Research Assistant

October 2025 - April 2026

- Migrated a Next.js microplastics research dashboard from per-request JSON/CSV file parsing to an indexed SQLite database via custom API routes and seeding, reducing API response times by 70+%.
- Engineered a universal Python ETL pipeline that auto-infers schemas from heterogeneous JSON sources, normalizes nested objects and arrays into relational SQLite tables, and validates row counts post-migration across 10+ tables.
- Fact-checked and corrected physical property data across 18 polymer records against peer-reviewed literature and manufacturer datasheets.

PROJECTS

Quantitative Backtesting Engine | Python, SQL, DuckDB, Pandas, Pytest

- Engineered an asynchronous ETL pipelined, event-driven backtester using DuckDB and Alpaca Markets, utilizing columnar data storage for 10x faster analytical queries than standard relational databases.
- Implemented a suite of strategies such as vectorized crossover & Bollinger Bands with RSI verification, visualizing respective portfolio qualities with NAV curves, and comparing against SPY and buy-and-hold benchmarks.
- Deployed live paper order execution via Alpaca API with position guards and buying-power-based position sizing.

Multithreaded Key-Value Server | C, POSIX, Ubuntu Linux

- Designed and implemented a datastore in C, supporting read-write functionality via CLI parsing and signaling.
- Engineered core functionality using POSIX system calls, enabling interaction with the operating system for file persistence, concurrency, and inter-process communication.
- Implemented datastore networking, supporting multiple clients through UNIX domain sockets and multithreading.

Cryptocurrency Portfolio Tracker | Java, Spring, MySQL

- Engineered "Coinleaf," a backend cryptocurrency portfolio tracker with RESTful API endpoints using Java, Spring Boot, and MySQL, enabling full CRUD operations and database persistence.
- Integrated live market data by consuming Coinbase API services, building a service layer to fetch and update real-time cryptocurrency prices with <200ms response times, improving accuracy of portfolio valuation.

Mandelbrot Fractal Visualizer | Python, Streamlit, NumPy, Numba

- Developed and optimized fractal generation algorithm using Numba JIT compilation, NumPy vectorization, and dynamic iteration control based on zoom depth, achieving over 700% image rendering speedup.

LEADERSHIP AND PROFESSIONAL DEVELOPMENT

ColorStack @ NYU

New York, NY

President & Co-Founder

January 2025 - Present

- Foster an NYU specific chapter of an organization with 16,000+ members nationally, providing peer support, career development, and technical skill-building for the empowerment of Black and Latino students in tech.
- Spearhead outreach activity, gathering 60+ student members through targeted engagement and event hosting.