

# Samuel D. Pekofsky

805-267-6067 | [sampekofsky@gmail.com](mailto:sampekofsky@gmail.com) | [linkedin.com/in/samuelpkofsky](https://linkedin.com/in/samuelpkofsky) | [pekofsky.github.io](https://pekofsky.github.io)

## EDUCATION

### University of Wisconsin-Madison

Sep 2021 – May 2025

B.S. in **Data Science** • Minor in **Computer Science** • Dean's List

Madison, WI

## EXPERIENCE

### Package Development Intern – clubSandwich

Feb 2025 – Aug 2025

University of Wisconsin-Madison Open Source Program Office

Madison, WI

- Developed an **R package** by **implementing new features**, writing **unit tests**, and updating documentation
- Authored and deployed **online training material** for future interns to quickly set up their systems and learn the basics of R package development
- Presented contributions using **slides**, **live demo**, and a **webpage deployed via GitHub Pages**

### Senior Camp Counselor

Jul 2023 – Aug 2023; Jun 2024 – Aug 2024

YMCA of Greater New York

New York, NY

- **Led and collaborated** in a dynamic, team-focused environment, **mentoring junior counselors** while fostering communication and teamwork

### Office Staff

Jan 2019 – Apr 2021

TaxpayerServices NYC, Inc.

New York, NY

- **Organized client documents**, answered calls, and responded to emails
- Set up and maintained **hardware and software systems**

## PROJECTS

### Deep Learning for Classifying Images | *Python, PyTorch, Git, GitHub*

Mar 2025 – May 2025

- Built a **hybrid CNN–Transformer model** for image recognition
- Implemented an **automated ETL pipeline** to **clean, load, and augment data** before feeding it into the model
- Trained and evaluated multiple models, **cross-comparing results** in a research paper

### Twitter Sentiment & Topic Modeling | *Python, Pandas, VADER, pyLDAvis, LaTeX*

Oct 2024 – Dec 2024

- Analyzed **2M tweets** using **NLP** and **topic modeling techniques**
- **Co-authored research paper** documenting methods, findings, and visualizations

### Volcano Database | *SQL, MongoDB, JSON*

Nov 2024 – Dec 2024

- **Designed, implemented, and queried** a relational database in MySQL
- **Migrated data and structure** to MongoDB, updating JSON schemas for compatibility

### US Migration Flows | *R, R Markdown, Tidyverse, Shiny, tigris, sf, Leaflet, Git, GitHub*

Feb 2025 – May 2025

- Developed a **reproducible ETL pipeline** that accessed migration data via an **API**
- Conducted **multi-year analyses** with extensive data processing, mapping, and **interactive visualizations**
- Produced a **well-formatted report** combining **statistical analysis** and visualizations

### NYC Arrests Analysis | *R, R Markdown, Tidyverse*

Mar 2024 – May 2024

- **Cleaned and analyzed 2M+ NYPD arrest records** through feature engineering and data visualization
- Created a **comprehensive report** with detailed **statistical analyses** and visuals

### Car Inventory App | *Java, Git*

Mar 2023 – Apr 2023

- **Collaborated with a large team** to design and build a Java app for organizing and searching car data
- Implemented **red-black trees** to enable efficient reordering and searching

Additional projects are available at [pekofsky.github.io/projects](https://pekofsky.github.io/projects)

## TECHNICAL SKILLS

**Languages:** Python, SQL, R, Java, JavaScript, HTML, CSS, Julia, MQL

**Libraries & Frameworks:** pandas, PyTorch, Django, Flask, Selenium, jekyll, testthat, JUnit, ggplot2, dplyr, devtools

**Developer Tools:** Git, GitHub, VSCode, RStudio, Google Colab, Eclipse

**Platforms & Services:** Google Cloud Platform, Tableau