

# Henry Williams

(919) 903-1265 | [jhwill28@ncsu.edu](mailto:jhwill28@ncsu.edu) | [LinkedIn](#) | [Website](#)

---

Interdisciplinary data scientist, problem-solver, and communicator dedicated to optimal and ethical decision-making. Equally happy wrangling data, leveraging statistics and machine learning to uncover insights that reliably deliver real value, and sharing those insights with technical and nontechnical audiences alike.

**Languages and Software:** R, Python, SQL, Tableau, AWS, ArcGIS, Excel, Linux, Git.

**Certifications:** PCEP — Python Certified Entry-Level Programmer. Preparing for AWS Certified Cloud Practitioner.

## EDUCATION

---

**M.S. in Analytics** May 2024

Institute for Advanced Analytics, North Carolina State University

**B.S. in Statistics, B.A. in Philosophy** May 2021

University of North Carolina at Chapel Hill

## PROJECTS

---

**M.S. PRACTICUM PROJECT, ATRIUM HEALTH** August 2023 — Current

### Communication Lead

- Predicting ICU readmissions by building logistic regression and machine learning models in Python.
- Transforming and posturing complex datasets of electronic health records from over 25,000 patients for explanatory and predictive modeling.
- Developing a Tableau dashboard to aid clinicians in real-time patient evaluations.

**PERSONAL PROJECT, JOURNAL DIGITIZATION** February 2024 — Current

- Fine-tuning a Transformer-based AI model to enable OCR of years of hand-written journal entries.
- Prepared a complex dataset for modeling by automating pre-processing of 15,000 lines of text using Python.

## WORK EXPERIENCE

---

**U.S. ENVIRONMENTAL PROTECTION AGENCY** Research Triangle Park, NC

**Toxicology Data Curator (Contract)** May 2022 — May 2023

- Expanded public-facing toxicology records database by over 25% (65,000+ records) through a mixture of Python scripting and web crawling, manual cataloging, and collaboration with the EPA Libraries.
- Extracted, transformed, and loaded all catalogued data using R and SQL to clean and pre-process.
- Developed and maintained cloud storage API connectivity functions for an internal Python package with 20+ users that have since been used to migrate terabytes of data.

## VOLUNTEER EXPERIENCE

---

**UNC STATISTICAL CONSULTING CENTER** Chapel Hill, NC

**Statistical Consultant** January 2022 — May 2022

- Fielded and discussed project proposals from a diverse set of clients across the university to assess project scope, feasibility, and initial approaches.
- Built a data cleaning pipeline in R to extract, transform, and load over 11,000 survey responses for future analysis while maintaining dialogue with our collaborators to ensure our process fit their goals.