Samuel Tucker

Data Engineer

Phone: (515) 864-1102
Email: SamuelT29@outlook.com
GitHub: https://github.com/SamTuck29

Professional Summary:

Detail-oriented data engineer with a strong foundation in problem-solving and a passion for continuous learning. I am skilled in ETL/ELT data pipeline development, systems architecture, machine learning, statistical modeling, and data visualization. With experience in project management, I can effectively lead projects from inception to completion, deliver strategic and actionable insights in a time-efficient manner, and communicate technical information to diverse audiences. Ultimately, I seek to leverage my expertise and passion for data to drive innovation and provide data-driven solutions that add meaningful business value.

Technical Skills:

Programming – Proficient understanding of R, Python, C#, HTML, SAS, SQL.

Data Engineering – Skilled in ETL/ELT data pipeline development: Azure Functions, Container Apps, Terraform, PubSub, GitHub, CI/CD Pipelines, REST APIs, Docker, Databases, FastAPI, matplotlib.

Machine Learning – Broad use of ML packages including Scikit-Learn and Pytorch with Pandas leveraging algorithms including: k nearest neighbors, random forests, stochastic gradient descent and more.

Web App Development – Experience creating user-friendly apps with Blazor, RShiny, and React.

Education:

AUGUST 2025 – PRESENT

Georgia Institute of Technology, Online
Masters of Science in Computer Science with Concentration in Machine Learning

AUGUST 2020 – MAY 2024

Drake University, Des Moines, IA
Bachelors of Science in Mathematics, Computer Science, and Data Analytics

References:

Available upon request.

Experience:

MARCH 2025 - PRESENT

Data Engineer | The Waldinger Corporation

Data Engineering: Create custom data pipelines that extract paginated data in real-time from external APIs, load to data lakes (Azure Blob Storage) and databases (Snowflake, Postgres, Oracle), and perform transformations that align with the business's needs. Support other teams by making data available through custom Python APIs. Automate, via CI/CD pipelines, code pushes from GitHub to Azure. Establish event-driven architecture via PubSub for C# app data crossover and Snowflake ingestion.

Systems Architecture: Collaborate with a data architect to design a flexible, scalable, reliable data pipeline architecture that is modularized and reproducible using Terraform modules and Python packages. Reduced pipeline creation timeline from months to days.

JUNE 2022 – MARCH 2025

Data Scientist | USDA-NASS

Machine Learning: Created, trained, and tested ML models using Sci-kit Learn in Python and ML.NET in C#. Models used NLP and classification to simplify tasks in an internal application.

Web Application Development: Collaborated on the creation of Blazor and RShiny applications. Responsibilities included full stack development, meeting stakeholder requirements, and deploying code with CI/CD pipelines.

AUGUST 2021 – DECEMBER 2023

Tutor/TA for Dr. Andrew Becklin | Drake University

Teaching: Coach and teach students in the classroom and during out-of-class office hours and review sessions, covering courses in Mathematics and Computer Science.

Projects:

R-Shiny Dashboard: Developed and presented data analytics capstone project findings through an interactive RShiny dashboard.

Machine Learning Classification Project: Implemented machine learning algorithms in Python to classify shoe brands using Pandas and Scikit-learn.

Logistic Regression Analysis: Conducted logistic regression analysis to assess the success factors of videos by YouTuber Mr. Beast, detailed in a whitepaper.