

DR JOEL MILLER

Data Scientist

@joelsamiller@gmail.com
in joel-miller-phd-0630a21aa

+447787808981
joelsamiller

London, UK
0000-0001-7775-8803

ABOUT ME

I have over seven years of experience applying data science techniques across academia and public health, and nearly 15 years of experience in computer programming over eight different languages. I am a self-driven individual with a scientific mindset, who loves building tools, solving problems and learning. I am also an international athlete who has represented Great Britain at the senior level, competing in many high-pressure scenarios.

EXPERIENCE

Principal Data Scientist

UK Health Security Agency

July 2023 – Ongoing

London

Python R SQL Pandas GeoPandas Matplotlib Plotly AWS Git

- Lead developer on a Python package for creating automated slide decks. Data is pulled from S3, SQL databases or local files, with pandas used for data manipulation and Plotly for data visualisations. This package has enabled hours of recurring work to be automated.
- Created a Python package to simplify authentication when connecting to AWS from our analytics platform. Reduces terminal commands and many lines of code to a single line. This has been well-received by other data scientists and has increased efficiency.
- Took an existing internal report that was manually created each month using Excel spreadsheets, with data received by email. Designed and built a new data pipeline and automated generation using Python. This saves three days of work each month.
- Automated two GOV.UK reports using R markdown, saving 10 days of analyst time per production cycle and improving quality.
- Working on mapping catchment areas of water treatment plants to inform site selection for wastewater monitoring using GeoPandas. Looking primarily at estimating population coverage vs the number of sampling sites chosen.
- Conducted a data quality review of a SQL database following the identification of an ingestion error. Measured impacts and made recommendations to key stakeholders.
- Created and delivered training sessions on Python and Git to interns. These were rated in the top three sessions they received.
- Managed and trained three employees who were promoted shortly afterwards, with two going from non-coding to coding roles.

Senior Data Scientist

UK Health Security Agency

April 2022 – July 2023

London

Python R SQL Pandas Plotly ggplot2 PyTorch AWS Git

- Improved a high-profile COVID-19 product built using a large object-oriented Python codebase and AWS. Reduced 60k lines of code to 10k, significantly increased speed and reduced operating costs by 85%. I also improved the visualisations to improve clarity and provide a consistent look and feel.
- Supported incident responses, delivering data analysis and visualisations on high-priority public health situations at pace.
- Worked to a tight deadline on a greenfield project using Python to extract data from PDF forms into a tabular form. This provided a solution where the existing manual process could not process forms fast enough during an outbreak.
- Led a team of 5 data scientists on a project to automate the production of a regular GOV.UK report using R-markdown. This reduced the production time from 2 days to ~30 seconds.
- Trained an LLM using Pytorch and HPC, producing a binary classification model with >90% accuracy. This model was then applied to free text data to improve reporting on Avian Influenza.
- Taught myself R and SQL on the job after being placed on projects using these languages, quickly delivering results.
- Worked collaboratively using an Agile framework, Jira boards and GitHub/GitLab. This enabled effective work in interdisciplinary teams and time management across multiple projects.

EDUCATION AND CERTIFICATIONS

Machine Learning, Modeling, and Simulation Principles

MIT xPro

📅 April 2024

Ph.D. Physics

University of Nottingham

📅 Oct 2017 – August 2021

 Python  C  Git  Numpy  Matplotlib  Linux

- Taught myself to write C, Python and IDL code to run complex calculations on large-scale simulations, analyse the results and produce visualisations.
- Analysed my results using statistical methods such as Welch's t-tests, bootstrapping and Spearman's rank coefficients, before placing them within the broader context of the literature of my area of research.
- Gave several intra-departmental talks and a talk at an academic conference on my research. Also delivered several outreach talks to a range of audiences. I was even interviewed by the BBC about astrophotography [Link].

Publications:

- Miller, J. S. A., Bolton, J. S., & Hatch, N. A. (2021). *Monthly Notices of the Royal Astronomical Society*, 506(4), 6001–6013. doi:10.1093/mnras/stab2083
- Miller, J. S. A., Bolton, J. S., & Hatch, N. (2019). *Monthly Notices of the Royal Astronomical Society*, 489(4), 5381–5397. doi:10.1093/mnras/stz2504

M.Sci. Physics with Astronomy

University of Nottingham

📅 Sept 2013 – June 2017

First Class

PERSONAL PROJECTS

Disc Tracker  Python  OpenCV  Numpy

Computer Vision

Takes stereo input of video files where a flying disc is thrown, then implements a Gaussian mixture model and blob detection to identify foreground objects. A Jonker-Volgenant assignment algorithm, with a cost matrix given by distances between previous and new locations, is used to track objects over time. Finally, tracks corresponding to the flying disc are de-projected to give the path in three dimensions.

N-Body Simulation  Python  Numpy

Numerical Simulations

Simulates the motions of many objects interacting via gravity given their initial conditions. Equations of motion are written as a series of first-order ordinary differential equations which are then solved using the forward Euler or Leapfrog methods.

ACHIEVEMENTS

Competitive team sports (Ultimate Frisbee)

- Won gold at the 2019 European Championships with Great Britain.
- Won 5 BUCS gold medals and 2 regional titles with university teams.
- Captained my university mixed indoor 1s through an undefeated season to become national champions.
- Finished top 4 at UK National Championships in 2023 and 2024.

INTERESTS

Ultimate Frisbee Astronomy Gym Photography Playing Piano Coding Computer Games
Science-Fiction Space Exploration