Joshua Prettyman, Ph.D.

Data Scientist, Python Developer, Mathematics Ph.D.

joshua@prettyman.me
linkedin.com/in/prettyman
joshuaprettyman.com

Personal Statement

I built an ML-powered SaaS platform that increased productivity by 20×, taking Blink SEO's internal tooling from spreadsheets to a marketable product: full-stack, handling everything from NLP to cloud infrastructure.

My research at the National Physical Laboratory produced a novel scaling indicator for tipping point detection in multidimensional dynamical systems, with applications to real geospatial time series data. Published in Environmental Research Letters (2022).

Technical Skills

- Regular use (10+ years) of Python: ScikitLearn, Numpy, Pandas, NLTK, Matplotlib, Plotly.
- Full stack development with Bash, Python, Google Cloud Platform (GCP), SQL, JavaScript.
- MLOps, RAG, NLP, OpenAI/Anthropic API integrations and fine-tuning, Ollama + HuggingFace.
- Data engineering (AWS, GCP, Databricks, dbt), ETL & ELT, analysis, dashboards (Tableau, Retool).
- Git, GitHub, Docker, package management, CI/CD, async, REST, GraphQL.
- Experience with C++, MATLAB, scientific computing, geospatial data, PyTorch.

Professional Experience

Data Freelancer — joshuaprettyman.com/freelance   *Jan. 2025 to Present*

Delivered projects including: Model Deployment, Dashboard Creation (Tableau), Pipeline Automation (dbt, Databricks, Apache), Database Setup (AWS), Data Engineering and Data Analysis.

Data Scientist — Blink SEO / Macaroni Software   *Nov. 2021 to Dec. 2024*

Built a full-stack ML SaaS platform from scratch, increasing team productivity by 20× and enabling year-long SEO campaigns to be delivered in weeks.

- Developed NLP-powered keyword clustering combined with quantitative data, reducing work from days to minutes.
- Architected Python backend (*GCP Compute Engine*) integrated with *BigQuery* to process 50M+ data points daily from external services' APIs.
- Translated workflows into automated pipelines through close collaboration and continuous deployment.
- Created a job-queue system (*PostgreSQL*) to run client onboarding, data imports, data analysis and ML tasks asynchronously in the backend.
- Built interactive dashboards (*Plotly + JavaScript*) enabling clients to identify and prioritise revenue opportunities in minutes and act on recommendations directly via service integrations.
- Integrated LLMs (*huggingface*) to generate ready-to-go suggestions for content gaps, allowing customers to see immediate impact.
- Delivered ad-hoc analyses and data engineering for investor presentations and internal reporting.

Data Science Researcher — National Physical Laboratory (NPL)   *Sep. 2015 to Feb. 2021*

Produced a novel scaling indicator for tipping point detection in multidimensional time series.

- Presented results for stakeholder meetings at NPL, in published articles, and at international conferences.
- Developed methods to detect tipping points in dynamical systems, working with large geospatial time series datasets and building stochastic models to test hypotheses.
- Code written in *MATLAB* and *Python*. Visualisations produced using *Matplotlib* and *XMGrace*.
- Accessing, cleaning, interpolating and correctly interpreting large, raw meteorological datasets.

Associate Lecturer — Sheffield Hallam University *Sep. 2017 to Jul. 2019*

Taught mathematics, statistics and computing courses from Foundation Degree to Masters.

- Taught a course on *Microsoft Excel for Business*.
- Responsibilities included lesson preparation, course planning, marking, leading tutorials.
- Manned the "Maths Help" desk three times a week, providing support to students of all departments.

Informatics developer (internship) — UK Met Office Informatics Lab *Jun. to Aug. 2017*

Project assessing the viability of using amateur meteorological data to augment official observations.

- Accessed internal and external data via APIs using Java.
- Data cleaning and analyses in Python.
- Worked with the Informatics team and participated in daily stand-ups.

ACADEMIC BACKGROUND

Ph.D. Mathematics — University of Reading
 Thesis: *Tipping Points and Early Warning Signals with Applications to Geophysical Data*.

- Thesis develops spectral analysis methods to detect correlations in multidimensional time series, which are used as Early Warning Signals for tipping events in stochastic dynamical systems.
- Methods applied to real geospatial time series data: hurricane paths and greening/desertification.
- Publication of three papers in respected journals:
  - *A novel scaling indicator of early warning signals* (EPL, 2018)
  - *Generalized early warning signals in multivariate and gridded data* (Chaos, 2019)
  - *Power spectrum scaling as a measure of critical slowing down* (ERL, 2022)
- Implemented mathematically derived methods numerically by discretising and linearising.
- Participation in several international conferences and workshops.
- Three-month data-focussed internship at the UK Met Office (see *Professional Exp.* section).
- EPSRC funded research in industrial partnership with NPL. Attracted additional funding from NERC.

M.Res. Mathematics — Imperial College London (*Distinction*)
 Dissertation: *Mesh Generation Using Solutions of an Optimal Transport Problem*.

- Dissertation uses solutions to a linearisation of the Monge-Ampère equation to generate an adaptive, optimally distributed mesh for numerical models.
- Working in C++; monitoring CPU usage and convergence time over various scenarios.
- Participation in weekly soft skills workshops and monthly internal showcases.
- Attendance at weekly Mathematics of Planet Earth seminars hosted by the Centre for Doctoral Training.
- Completed taught courses in Python, R, Dynamical Systems, Stochastic differential equations, Bayesian Analysis and Probability.

MA Mathematics — University of Edinburgh (*First Class Hons.*)
 Dissertation: *Growth and its Applications in Graph Theory*.

- Dissertation explores patterns in paths through connected digraphs.
- Wrote a program with a simple UI to investigate the properties of these digraphs: this then became a teaching aid, earning a letter of thanks from the principal.
- Requisite taught courses for the Pure Mathematics degree besides optional modules in computing and mathematics education.
- Maths outreach in local schools, museum events and university seminars.