

Jack Grimes

jack@jackgrimes.net

London, UK

jackgrimes.net

GitHub: [jackgrimes](#)

LinkedIn: [jacketgrimes](#)

Google Scholar: [HZLogpQAAAAJ](#)

EDUCATION

IMPERIAL COLLEGE LONDON

PH.D. ENGINEERING/EPIDEMIOLOGY

March 2012 - September 2015

IMPERIAL COLLEGE LONDON

M.Sc. ENVIRONMENTAL
ENGINEERING AND SUSTAINABLE
DEVELOPMENT

September 2010 - September 2011

UNIVERSITY OF OXFORD

M.A. (OXON) PHYSICS

September 2007 - September 2010

SKILLS

PROGRAMMING

Very familiar: Python • R • Git/GitHub
Matlab/Octave

Familiar: SQL • JavaScript • HTML
CSS • \LaTeX

PLATFORMS

Google Cloud Platform • Domino

COURSES

(Ongoing:) Machine Learning on Google
Cloud Specialization (Coursera, Google)

Bayesian Statistics: From Concept to
Data Analysis (Coursera, University of
California, Santa Cruz)

Deep Learning Specialization (Coursera,
[deeplearning.ai](#))

Machine Learning (Coursera, Stanford
University)

INTERESTS

Cycling • Walking • Travelling • Reading

EXTENDED CV

A more detailed version of my CV can be
found on my [website](#)

EXPERIENCE

MORSUM London, UK

March 2021 - Present: SENIOR DATA SCIENTIST

- [.Anchovi](#): A food marketplace for food distributors, restaurants, and shops
- Product name cleaner: classifiers for filling out abbreviations, identifying descriptions, and regex for removing unwanted string patterns, hosting on Flask API
- Machine learning framework for productionising model (re-)training, serving and monitoring
- Product recommender system: design and build of a system including a framework for A/B testing
- Design of databases and pipelines for data ingestion
- Driving the use of unit tests, code linting and formatting, and sonarcloud and GitHub Actions builds to ensure high coding standards
- Building the Data Science team and outlining the teams' responsibilities
- Git, Google Cloud Platform, MySQL, pandas, PyCharm, pytest, scikit-learn

LLOYDS BANKING GROUP London, UK

January 2019 - March 2021: DATA SCIENTIST, APPLIED SCIENCE

- Efficient, readable, unit-tested data science code for production
- Classification, clustering, network analysis, NLP: identification of incorrectly applied account and customer indicators
- Running the *Modelling Interest Group*, organising regular talks
- *Data Science Graduate Scheme Training Programme*: delivering the *Advanced Modelling* and *Natural Language Processing* sections
- Git, Jenkins, NetworkX, pandas, PyCharm, pytest, scikit-learn, spaCy

August 2017 - December 2018: DATA SCIENTIST, RISK INNOVATION

- Predictive modelling of fraudulent transactions, and of financial difficulty
- Topic modelling of regulatory data
- Presenting at internal conferences throughout the UK
- Data science solutions packaged into Flask apps for other teams
- face_recognition, Flask, Gensim, Git, pandas, SAS VA, scikit-learn

IMPERIAL COLLEGE LONDON London, UK

January 2016 - July 2017: POSTDOCTORAL RESEARCHER

- Mixed effects logistic regression (R and lme4) to compare environmental risk factors, parasites, and overall health, in a clustered dataset ([paper](#))
- Box-Cox transformations and linear regression (in R) to account for the 'long tail' of light parasite infections missed due to insufficient diagnostic sensitivity

March 2012 - December 2015: PH.D. RESEARCH

- Thesis: *An investigation into the roles of water, sanitation, and hygiene in the control of schistosomes and other helminths*
- R, ArcGIS for statistical comparisons between risk factors and infection
- Fieldwork: around six months in Ethiopia
- Presented at 6+ international conferences, led 11 journal articles
- Supervision of two M.Sc. students (plus an M.Eng. student in 2020)