gavin leech

data scientist, software developer

about

London based.

Email
Github
Blog
LinkedIn

research

Speaker similarity
Aid impact bounds
Automation & welfare
On data science
Maximum wage
Controversy vs impact
Formal analogies

programming

Python/SciPy/Spark,
Keras, *nix, C#, Scala,
JS (ES6, node, TS),
PHP, Pig, LaTeX
...Powershell.

metadata

vs publication bias
Career JSON

interests

effective altruism, failures of human reasoning, failures of human reasoning about human reasoning, statistics as applied philosophy, data mining as conceptual analysis, lists.

experience

Sep'16 – Now — **Data Scientist** — AXA Insurance, London
Creating infrastructure, tools and machine learning models (GLMs, GBMs, ANNs) for pricing, medical risk, & image recognition. Also web scraping at scale & managing a TB/day Hadoop pipeline. Tech lead for 4 other 'scientists.
Buzzwords: Python, Spark 2.2, Keras, xgboost, Jupyter, Oozie, Avro, Pytest.

May – Aug'16 — **Frontend Developer** — Furo, London
Developed an Angular SPA for a procurement startup. Sole responsibility for the design & implementation of tendering. TS controllers from C# APIs.
Buzzwords: C# 6, Typescript, Angular, Visual Studio/SQL Server, Bitbucket.

Oct'15 – Jun'16 — **Software Developer** — Freetobook, Glasgow
Improved a large legacy accommodation engine: PHP back-end, JS front-end, REST APIs. Sole design of secure, PCI compliant, credit-card storage.
Buzzwords: PHP (5 & 7), LAMP, SQL, Laravel, cURL, Gitlab, New Relic.

Jun – Sep'15 — **Graduate Software Developer** — Stugo, Glasgow
Jumped into web tech from scratch; wrote first production code within 2 weeks, deployed to pharmacies worldwide. Worked on a medicine pump manager (C#) & a secure pharmacy record system (nodeJS). 4th employee.
Buzzwords: C#, WPF, node.JS, MEAN, Ractive, Mongo, Vagrant, VMWare.

Dec'14 – May'15 — **Database Developer** — STAND International, Edinburgh
Built a small ops database for an international NGO. Greenfield design & deployment; delivered at zero infrastructure and licence tcost.
Buzzwords: MS Access, Office365, SQL, UML, data protection.

Jan – Apr'14 — **Assistant Statistician** — The Scottish Government, Edinburgh
Derived & explained official stats in Education Analytical Services. Worked on the 'Insight' reporting platform, giving teachers hard data & viz on their classes & themselves. Prepped the national Social Care Survey 2015.
Buzzwords: SAS, SQL, OLAP, PCA, data protection, official statistics.

software

2017	**sesame: SEarch for StAtistical MEthods**	work in progress

Decision trees / search engine for which stats you can/should use. Scala.js

2017	**The First Computers**	deployed here

Tiny app (and a conceptual analysis) to solve a historical question. Uses the Transcrypt Python transpiler - unfortunately.

2016	**modernlove**	work in progress

A visualiser for sexualities. Python / Django / nose / seaborn

2015	**Py2HTK**	github here

Wrapper for venerable (creaky) ML library, the Hidden Markov model ToolKit.

education

2015 – Now	**BSc Mathematics & Statistics**	The Open University

Proper foundation: Analysis, probability theory, functional analysis, linear algebra, ODEs & PDEs, experiment design and RCTs, sampling theory, classical mechanics & field theory, stochastic processes, medical statistics, Bayesian modelling. Part-time during full-time work.

2014 – 2015	**MSc Software Development**	University of Glasgow

Conversion Masters: Java, SQL, HTML/CSS/PHP, MC6800 assembly plus 'patterns, data structures & algorithmics. The psychology of security, law of safety-critical systems, design of UX. Thesis in computational linguistics: 'Hidden Markov Models in the study of accommodation'.

2008 – 2012	**MA (Hons) Economics with Philosophy**	University of Aberdeen

Topics covered included formal logic, micro- & macro- economics, rudimentary statistics, health policy, development policy. Received a four-year Alumni Fund scholarship on merit. Dissertation in foreign aid effectiveness: 'Evaluating aid abolitionism'.