

Gordon McDonald

data science | statistics | visualisation | physics | mathematics | chemistry | teaching

Hi, I'm Gordon - a data scientist and visualisation expert working at the Sydney Informatics Hub, at the University of Sydney. I have extensive simulation and computational data analysis experience, academic publication experience, and a passion for informative data visualization techniques, which I have developed during three years of project based statistical and data science consulting for internal and external clients, as well as seven years of cutting-edge physics research. I have taught university level physics and data analysis courses in Australia and Indonesia, as well as tutoring students in Statistics, Mathematics, Coding, Physics and Chemistry. I have presented my research at other universities and conferences in Australia, the US, Europe and Indonesia; and I have co-authored 18 peer-reviewed papers in academic journals, as well as writing my honours and doctoral theses.

I am experienced in working in a small team environment using agile management techniques; I mentor three staff in my current role, having previously supervised two honours students and mentored four PhD students. In this role I have developed experience with project management and stakeholder engagement across multiple organisations in my role at the University of Sydney, both with internal researchers and external clients in government departments, in the hospital system and at medical research institutes across Sydney.

experience

2016-Now

Sydney Informatics Hub / Centre for Translational Data Science

The University of Sydney

Data Scientist

Since working as a data scientist at the Sydney Informatics Hub, I have been applying frequentist, machine learning and Bayesian statistical techniques to:

Health

- An analysis of the relative risk of discharge against medical advice within the Sydney Children's Hospital Network over five years of historical admissions records and 250k admissions, which I presented at the Health Data Analytics conference in Brisbane, October 2017. *Methods:* Bayesian logistic regression. *Software:* R.

- Clinical studies at the Woolcock Institute into sleep disorders such as insomnia and how they can be monitored with actigraphy measurements (i.e. a fitness wristwatch). *Methods:* Hierarchical and k-means Clustering, principal component analysis, descriptive statistics and visualisation. *Software:* R and R Shiny.

- A comparison of cost-effectiveness between two competing dental health programs for remote indigenous communities in NSW. *Methods:* Regression, hypothesis testing and visualisation. *Software:* Matlab and Tableau

Science

- Creating a software tool to streamline the process of analyzing metabolites through High Pressure Liquid Chromatography Mass Spectroscopy (HPLC-MS) at the Charles Perkins Centre. *Methods:* Time series analysis, wavelet transform, peak detection, hierarchical clustering, regression and multiple hypothesis testing. *Software:* R and R Shiny.

- Developing a software tool to enable researchers to calculate chemical concentrations and kinetics in complex biological reactions involved in cell differentiation in developing spinal cords. *Methods:* numerical solutions of differential equations *Software:* Matlab and Mathematica.

contact

+61 424 344 560

mail@gordonmcd.com

gtdmcdonald

gtdmcdonald

Education

· Financial modelling for the NSW Department of Industry's Smart and Skilled program for Vocational education and training, a program which allocates more than \$600 million a year in subsidies for NSW students. We have developed a predictive tool to forecast and drill down into spend within the program. *Methods:* Conditional inference trees, survival modelling, rank regression, correlation, logistic regression, Sequential event modelling. *Software:* R, Oracle SQL, Tableau, Prophet.

Social Science

· Correlating election results for the 2016 US presidential election and the 2016 UK Brexit election with demographics of each electoral region. *Methods:* Partial Correlations, Linear regression with bayesian variable selection, Variational Inference. *Software:* R, Matlab, Tableau

· Analysing crime data across NSW looking for demographic, spatiotemporal, and weather-related patterns which can be exploited to improve policing efficiency. *Methods:* Bayesian generalized linear regression, correlations and mutual information. *Software:* Python

· Investigating skills transferability within occupational subgroups to identify near-fit training opportunities, in a project with the NSW Department of Industry. To enable this analysis we are combining change-of-occupation data from the Australians At Work longitudinal survey with skills similarity of those same occupations derived from millions of web-scraped job ads in data collected by Burning Glass. By identifying pairs of occupations between which there could theoretically be a labour flow, and comparing that to actual labour flow, we can identify mismatches where there is the potential for near-fit training to unblock career pathways. *Methods:* Dynamic network visualisation, hierarchical clustering, silhouette scores, Jaccard indices. *Software:* R, plotly, D3.

Since partly taking on the role of the university's statistical consulting service, I have also been providing statistical assistance to researchers across the university including meta-analyses, survey analysis, experimental design and statistical methods.

I designed and built the Sydney Informatics Hub website at <https://informatics.sydney.edu.au> using Hugo.

2015

Quantum Sensors and Atom Laser Group

Australian National University

Postdoctoral Researcher

I developed software to leverage principal component analysis, non linear fitting and Fourier transforms to automate our image processing pipeline and extract relevant measurements for our physics research.

education

2010–2015

PhD in Experimental Quantum Physics

Australian National University

Thesis: *"Cold Atom Interferometry in Optical Potentials"*

Worked in a small team to develop and prototype the next generation of quantum sensors for fundamental science, mineral exploration and military applications.

2009

First class Honours with University Medal in Physics,

Australian National University

Thesis: *"Detecting Atomic Shot Noise On Ultra-cold Atom Clouds"*

2006–2008

Bachelor of Science

Australian National University

Majoring in Physics and Chemistry, minor in Mathematics. High Distinction average of 90% (Grade point average of 6.87 out of 7).

teaching experience

2018	Data Carpentry R Ecology course <i>Teaching a 2-day workshops on basic data analysis and reproducible research</i>	ResBaz2018 @ Macquarie University
2017	Sydney Machine Learning <i>Teaching basic data science skills in Python in a course following the CS109 syllabus from Harvard.</i>	Amazon Web Services, Sydney
2017	Certified Software Carpentry Instructor <i>Teaching workshops on basic data analysis and reproducible research</i>	Software Carpentry
2017	Introduction to data analysis in R <i>Course Coordinator</i>	Brain and Mind Centre, University of Sydney
2013	International Engineering Program <i>Guest Lecturer in Electromagnetism</i>	University of Indonesia
2012–2015	ANU Dance and Salsabor Dance Studio <i>Dance Teacher</i>	Canberra, Australia and Washington DC, USA
2010–2014	Advanced Physics Tutor I tutored first-year Foundations of Physics, third-year Advanced Theoretical Physics and fourth-year Honours Electromagnetism	Australian National University

selected awards

2014	Vice Chancellor's Award for Outstanding Contribution to Student Learning	ANU
2014	Editor's choice for a publication in Europhysics Letters	Europhysics Letters
2009	University Medal	ANU

selected publications

My work at the University of Sydney has resulted in a co-authorship on one published paper, with four more in preparation. I gave a talk at the Health Data Analytics conference in Brisbane in October 2017 on our work on discharge against medical advice in the Sydney Children's Hospital Network. My doctoral research resulted in 18 scientific publications in peer-reviewed journals from 2011 to 2017 (garnering over 600 citations so far); as well as 12 presentations delivered around the world, in addition to my doctoral thesis, all of which are available upon request. My honours thesis is also available. Some selected publications are highlighted below.

2018	Disparate binding kinetics by an intrinsically disordered domain enables temporal regulation of transcriptional complex formation Chemical reaction rate in gene transcription in motor neurones is affected by domains of protein which lack a well-defined 3D structure. This in turn affects cell differentiation. Proceedings of the National Academy of Sciences 115 (18) 4643-4648 (2018)	Proceedings of the National Academy of Sciences
2016	Simultaneous Precision Gravimetry and Magnetic Gradiometry with a Bose-Einstein Condensate: A High Precision, Quantum Sensor We measure local gravity to a part per billion, and measure the local magnetic gradient to 0.1 nT/m, using a Bose-Einstein Condensate based atom interferometer. Phys. Rev. Lett. 117 138501 (2016)	Physical Review Letters
2015	Fast machine-learning online optimization of ultra-cold-atom experiments We use machine learning to optimize laboratory production and tuning of Bose-Einstein Condensates. Scientific Reports 6 25890 (2016)	Scientific Reports
2014	A Bright Solitonic Matter-Wave Interferometer A soliton is a non-dispersive cloud of atoms. We present the first soliton-based matter-wave interferometer, and show that using a soliton increases the visibility of the interference fringe. Phys. Rev. Lett. 113 013002 (2014)	Physical Review Letters