Kevin Wang Ph.D.

Data Scientist and Statistician

📱 *(+61) 404 955 255*
✉ *kevinwangstats@gmail.com*
🌐 *Portfolio: kevinwangstats.com*
Nationality: Australian

Experience

Now
2021 Nov — **Senior Data Scientist**, *Illumina*
- Leading strategic bioinformatics development for three pioneering IVD cancer companion diagnostic products on the TruSight Oncology DRAGEN platform. Delivered critical insights that enhanced efficiency and enabled timely completion of high-value deliverables, directly contributing to multi-million-dollar project outcomes.
- Collaborating with interdisciplinary teams of R&D scientists, software developers, regulatory bodies, and external pharmaceutical partners to achieve regulatory milestones and meet critical multi-million-dollar contractual payments. Demonstrated expertise in cross-functional leadership and stakeholder management.
- Standardizing internal clinical and medical data management processes, creating a scalable framework that improved data consistency, reliability, and accessibility across teams, streamlining downstream workflows.
- Developing and integrating machine learning algorithms to analyze large-scale genomics datasets, uncovering actionable insights that significantly enhanced the performance of oncology diagnostics and prenatal testing products. Highlighting technical expertise and innovation in developing commercially viable solutions.
- Designing and deploying production-grade bioinformatics tools and pipelines in Python and R, improving data processing efficiency and ensuring scalability for future commercial applications. These tools bolstered the company's ability to handle large datasets reliably in high-throughput settings.

2021 Oct
2020 Jul — **Statistician**, *CSL Behring*
- Used statistical process control techniques to standardise manufacturing and quality control laboratory procedures with engineers and biochemists.
- Designed experiments and used innovative machine learning models and R-Shiny dashboards to understand the chemical stability of products, conform to regulatory requirements and ensure product effectiveness.
- Worked with high-value products like COVID vaccines and treatments and provided actionable insights to achieve savings of multiple product batches, some up to $1M.

2020 Jun
2019 Sep — **Research Associate**, *University of Sydney*
- The major research project developed an interpretable machine learning framework for predicting patient clinical outcomes using omics data. Melanoma was a key focus of this research and the key biomarkers were validated and implemented on a cost-effective bioassay.
- R software packages associated with the research were developed and maintained in a reproducible environment.
- Visiting scholar at Cornell University for 2 months and delivered a workshop in open-source single-cell analytics using Cloud technology.
- Published peer-reviewed articles in top-tier journals, including PNAS, Nature Digital Medicine and Nature Computational Science.

2019 Mar
2016 Mar — **Postgraduate Teaching Fellow**, *University of Sydney*
- Delivered lectures (200+ students) and tutorials, covering 15 different courses, including statistics, mathematics and data science. Achieved 85% satisfactory rating for end-of-course surveys.
- Authored and designed course content at both the undergraduate and postgraduate levels.
- Mentored and trained junior statistics tutors.

Education

2016–2020 — **Doctor of Philosophy in Statistical Bioinformatics**, *University of Sydney*
- Developed machine learning methods to solve statistical challenges in patient clinical outcome predictions and thereby enabling personalised medicine.
- Real-world biomedical data from collaborators were used to motivate and articulate important challenges in current statistical research. The value of this thesis is the synthesis of statistics, clinical implementation and practical deployment.

2012-2015	**Bachelor of Science (Adv. Mathematics) (Hon. I)**, *University of Sydney*

- Major in statistics and financial mathematics.
- The Honours thesis examined functional MRI data and the inferred brain connectivity networks through the use of different statistical association measures.

Skills

python Developed computational modules for regulated products at Illumina to support clinical studies. Highly proficient in writing scripts and notebooks to solve analytical problems for stakeholders. Experience in using and applying `pandas`, `numpy`, `scikit-learn`, `keras`, `pytorch`, and `openai` (ChatGPT) for a range of projects.

R Highly proficient in scripting and developing packages. Strong ability to use `tidyverse` and `shiny` for the data ETL process to produce visualisations and presentable insights. Author and maintainer of several open-source R packages, including `APES`, `learningtower`, `mcvis` and `scMerge`.

git & CI/CD & docker Highly proficient in a reproducible environment that supports business operations and software development requirements. Experience in holding reproducible workshops with excellent participant feedback.

Applied statistics Highly experienced at the research level and consulting level. Ability to transform business and cross-disciplinary problems into a statistical framework and deliver solutions. Strong experience with model selection to extract explainable insights from data.

Statistical modelling Highly experienced at the research level and consulting level. Specialisation in predictive modelling for high-dimensional data.

Other skills SQL, cloud and dashboard development for business use.

Volunteering

2021-Present **Mentor**, *Monash University*
Supervised and mentored three Master of Data Science students with the students successfully completed their research projects. Developed high-quality R-CRAN package and published academic journal article.

2021-2023 **Councillor**, *Statistical Society of Australia*
Served as the Assistant Secretary and Communication Officer in the NSW and Victorian Branch of the Society. Organised public events, designed automated management systems, responsible for member communications and social media account maintenance.

Selected Awards

2019 **Statistical Society of Australia, NSW**, *JB Douglas Award (runner-up). Annual competitive prize awarded to the top statistics research PhD candidate in the State.*

2019 **Statistical Society of Australia**, *Golden Jubilee Travel Grant. Annual competitive travel funding for a statistics PhD candidate in Australia.*

2017 **International Biometric Society - Australasian Region**, *Best student talk at the biennial conference.*

2016-2020 **Australian Postgraduate Award**, *for the duration of the PhD program at the University of Sydney.*

2015 **International Biometric Society - Australasian Region**, *Honours scholarship award. Annual award to the top biostatistics student undertaking an Honours program in AU/NZ region.*