

# Kevin Wang Ph.D.

*Data Scientist and Statistician*

📞 US: (+1) 619-215-3725

📞 AU: (+61) 404-955-255

✉️ kevinwangstats@gmail.com

🌐 Portfolio: kevinwangstats.com

## Experience

2021 Nov **Senior Data Scientist, Illumina**

- Now
  - Led strategic informatics initiatives for high-impact cancer diagnostics products, delivering critical insights that contributed to multimillion-dollar partnerships and regulatory approvals.
  - Built and maintained company-wide data storage and computational tools in Python and R, significantly improving product development speed and reducing manual review processes from months to hours.
  - Developed scalable frameworks to improve how teams across medical, clinical, and operations departments manage and validate data - boosting decision-making efficiency.
  - Coordinated cross-functional projects with scientists, software teams, regulatory bodies, and external partners to align product development with global compliance goals.
  - Applied machine learning to uncover insights in large datasets, enhancing the performance of products used in cancer detection and prenatal health.
  - Led a peer-reviewed publication highlighting business value of a diagnostic product for blood cancer.

2020 Jul **Statistician, CSL Behring**

- 2021 Oct
  - Contributed to key R&D initiatives during the COVID-19 pandemic, supporting development of vaccines and treatments.
  - Built real-time monitoring tools for manufacturing quality control, helping streamline compliance and reduce production risks.
  - Delivered data-driven insights that averted potential regulatory delays and safeguarded product revenue, with direct impact on large-scale production batches.

2019 Sep **Research Associate, University of Sydney**

- 2020 Jun
  - Created an easy-to-understand predictive analytics framework to forecast patient outcomes using medical data, resulting in a cost-effective test for melanoma risk.
  - Maintained open-source software packages used by other research and health institutions.
  - Hosted a research workshop at Cornell University on advanced cloud-based data analysis tools.
  - Published in globally recognized journals including Nature Digital Medicine and PNAS.

2016 Mar **Postgraduate Teaching Fellow, University of Sydney**

- 2019 Mar
  - Taught over 200 students per course across statistics, data science, and mathematics subjects.
  - Designed coursework at both undergraduate and postgraduate levels and received an 85
  - Mentored junior instructors in teaching and technical skills.

## Education

2016 **Doctor of Philosophy in Statistical Bioinformatics, University of Sydney**

2020 Focused on machine learning applications for personalized medicine, directly relevant to the commercial sector's interest in tailored healthcare solutions.

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

2015 Major in statistics and financial mathematics. Thesis on brain imaging analysis using advanced statistical modeling.

## 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 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 under a reproducible environment that support business operations and software development requirements. Proficient in building docker images and deployment through Google Cloud. 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 such as omics data with 20k+ features.
- 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 a Honours program in AU/NZ region*