

Kevin Wang Ph.D.

Data Scientist and Statistician

☎ (+1) 619-215-3725

✉ kevinwangstats@gmail.com

🌐 Portfolio: kevinwangstats.com

Experience

2021 Nov **Senior Data Scientist, Illumina**

- Now ○ Pioneered data science and analytical strategies for three cancer diagnostic products, demonstrating a deep understanding of data-driven efficiency and cost reduction in a product development setting.
- Spearheaded cross-functional collaborations, directly contributing to multi-million dollars milestone achievements and exemplifying leadership in stakeholder management with pharmaceutical partners and regulatory bodies.
- Applied machine learning to enhance product efficacy, showcasing the ability to translate technical expertise into commercial success.
- Crafting and deploying robust, production-grade analytical tools and software using Python and R. Contributions towards efficient data management strategies were pivotal in securing the reliability and scalability of the company's data solutions and commercial operations.

2020 Jul **Statistician, CSL Behring**

- 2021 Oct ○ Utilized statistical process control methodologies and experimental design to optimize product stability and manufacturing processes, showcasing a strong blend of technical skill and commercial insight. This role involved critical analysis and application of statistical models to ensure product efficacy and compliance with regulatory standards.
- Delivered substantial cost savings and streamlined formulation through innovative analytical approaches, underlining a strategic and impactful approach to data analysis in the manufacturing of high-value biotechnology products.
- Designed and implemented R-Shiny dashboards for real-time data analysis, significantly enhancing decision-making processes and operational efficiency. This work demonstrated a high level of proficiency in data visualization and user interface design to support business operations.

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

- 2020 Jun ○ Developed an interpretable machine learning framework for clinical outcomes, highlighting the practicality of translating predictive analytics into actionable healthcare decisions. These projects showcased the ability to translate complex data into actionable insights, contributing to the advancement of personalized medicine.
- Lead the development activities and actively maintained open-source R software packages in a reproducible environment, demonstrating a commitment to the scientific community and advancing the field of bioinformatics data science through accessible tools and resources.

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

- 2019 Mar ○ Developed and delivered a comprehensive curriculum in data science across more than 15 courses, achieving high student satisfaction ratings. This role highlighted strong communication skills and the ability to convey complex statistical concepts in an accessible manner to a broad audience.
- Mentored and trained junior statistics tutors, demonstrating leadership and a commitment to educational excellence. This experience showcases an ability to develop talent and foster a learning environment that encourages curiosity and critical thinking.

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.

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 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

- 2021-2023 **Councillor, Statistical Society of Australia**
Demonstrated leadership and communication skills in a professional society, contributing to the advancement of the statistical community and its application in various industries.

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