Kevin Wang PhD

Data Scientist and Statistician

(+61) 404 955 255
kevinwangstats@gmail.com
kevinwang09.github.io/
Nationality: Australian

Passionate data scientist with a PhD in statistics and a proven track record of delivering valuable analytical insights. I am a highly effective communicator with numerous awards. I am highly skilled in developing machine learning models using the R language.

Experience

2020–Now **Statistician**, *CSL Behring*, Melbourne, Australia.

- Establish new machine learning models and biostatistical protocols to address business challenges.
- Liaising with internal clients to translate practical challenges to statistical analysis and towards final implementation and ultimately drive impact.
- Providing statistical consultation services to guarantee high pharmaceutical production quality.

2019–2020 **Research Associate**, *University of Sydney*.

- Developed a clinically implementable prediction model using genomics data a team of with clinicians and biologists.
- Developed and maintained open-sourced R software packages.
- Designed and hosted cloud-based workshops.

2016–2019 **Postgraduate Teaching Fellow**, *University of Sydney*.

- Worked within a teaching team to author lecturing materials for courses at both undergraduate and postgraduate level with strong positive student feedback.
- Delivered lectures (200+ students) and tutorials, covering 15 different courses, including statistics, mathematics and data science.
- Mentored and trained new statistics tutors.

Education

2016–2020 **Doctor of Philosophy in Science**, *University of Sydney*.

- Research in statistics and bioinformatics.
- The PhD thesis develops methods to enable prediction of patient clinical outcomes using omics data under a rigorous statistical framework.
- Produced peer-reviewed articles with collaborators, covering a wide range of cancers, diseases and statistical topics.

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

- Major in statistics and financial mathematics.
- The Honours thesis examines the effect of different measures of association in human brain connectivity modelling utilising functional MRI data.

Skills

Applied statistics	Highly experienced at the research level and consulting level. Specialisation in inferential and predictive modelling.
`R` & `git`	Highly proficient in scripting, `tidyverse`, `shiny` and package development with unit testing, continuous integration and coverage testing. Author and maintainer of three open-source `R` packages.
`python` & `SQL`	Proficient in writing working scripts.
`docker` & Google Cloud	Proficient in building docker images and deployment through Google Cloud.

Volunteering

2021 **Deputy communication officer**, *Statistical Society of Australia (Victorian branch)*.
Managing communications and organising events for the society.

2016-2019 **Outreach volunteer**, *University of Sydney*.
Delivered seminars and hands-on workshops for visiting high school and university students.

2017 **President**, *Sydney University Mathematics Society*.
Oversaw finances of the Society, organised academic/industry events and and negotiated coporate sponsorships for three years.

2015 **President**, *Sydney University Statistics Society*.
Organised academic/industry events and and negotiating coporate sponsorships.

Scholarships & Awards

2019 **Statistical Society of Australia**, *JB Douglas Award (joint runner-up). Prize awarded to the top statistics reseach PhD candidates in New South Wales. AUD 250*.

2019 **Statistical Society of Australia**, *Golden Jubilee Travel Grant. Competitive funding for a statistics PhD candidate in Australia. AUD 1,000*.

2018 **Sydney Bioinformatics Research Symposium**, *Best poster presentation at the conference. AUD 50*.

2017 **International Biometric Society - Australasian Region**, *Best student talk at the conference. AUD 500*.

2017 & 2018 **Australian Bioinformatics & Computational Biology Society**, *Travel scholarship to ABACBS annual conference. AUD 250*.

2016-2020 **Australian Postgraduate Award**, *For the duration of PhD program at the University of Sydney. AUD 26,288 p.a.*

2016 **Summer Research Scholar in Bioinformatics**, *Charles Perkins Centre*.
Research in data visualisation, focus on gene interaction networks for cancers. AUD 2,400.

2015 **International Biometric Society - Australasian Region**, *Awarded to the top biostatistics student undertaking a Honours program. AUD 1,500*.

2015 **Summer Research Scholar in Mathematical Statistics**, *Australian National University*.
Research in statistics, particularly on model selection and averaging techniques. AUD 3,000.

2013 **Research Internship in Mathematical Biology**, *Winston Charitable Foundation*.
Research in mathematical modelling in the spread of infectious parasites in beehives. AUD 2,000.

Publications

1 Kim J.H., **Wang, K. Y. X.**, Chen, C., Lin, Y., Tam, P.P.L., Lin, D.M., Yang, J.Y.H., & Yang, P. **2021**. Cepo uncovers cell identity through differential stability. (under review)

1 Lin, C., **Wang, K. Y. X.**, & Mueller, S. **2020**. mcvis: A new framework for collinearity discovery, diagnostic and visualization. Journal of Computational and Graphical Statistics, 1-13.

2 Hewavisenti, R., Ferguson, A., **Wang, K. Y. X.**, Jones, D., Gebhardt, T., Edwards, J., Zhang, M., Britton, W., Yang, J., Hong, A., & Palendira, U. **2020**. CD103+ tumour-resident CD8+ T cell numbers underlie improved patient survival in oropharyngeal squamous cell carcinoma. Journal for ImmunoTherapy of Cancer, 8:e000452.

3 **Wang, K.Y.X.**, Tarr, G., Yang, J.Y.H., Mueller, S. **2019**. Fast and approximate exhaustive variable selection for generalised linear models with APES, Invited paper to *Australia & New Zealand Journal of Statistics*, 61 (4) 445-465.

4 Lin, Y., Ghazanfar, S., **Wang, K.Y.X.**, Gagnon-bartsch, J.A., Lo, K.K., Han, Z., Ormerod, J.T., Speed, T.P., Yang, P., Yang, J.Y.H. **2019**. scMerge: Leveraging factor analysis, stable expression and pseudo-replication to merge multiple single-cell RNA-seq data, *Proceedings of the National Academy of Sciences of the United States of America*, 116 (20) 9775-9784.

5 Pires da Silva, I., **Wang, K.Y.X.**, Wilmott, J.S., Holst, J., Carlino, M.S., Park, J.J., Quek, C., Wongchenko, M., Yan, Y., Mann, G., Johnson, D.B., McQuade, J.L., Rai, R., Kefford, R.F., Rizos, H., Scolyer, R.A., Yang, J.Y.H., Long, G. V, Menzies, A.M. **2019**. Distinct molecular profiles and immunotherapy treatment outcomes of V600E and V600K BRAF-mutant melanoma. *Clinical Cancer Research*, 25 (4) 1272-1279.

6 **Wang, K.Y.X**., Menzies, A.M., Silva, I.P., Wilmott, J.S., Yan, Y., Wongchenko, M., Kefford, R.F., Scolyer, R.A., Long, G. V, Tarr, G., Mueller, S., Yang, J.Y.H. **2019**. bcGST - an interactive bias-correction method to identify over-represented gene-sets in boutique arrays. *Bioinformatics*, 35 (8) 1350-1357.

7 Strbenac, D., **Wang, K.Y.X.**, Wang, X., Dong, J., Mann, G.J., Mueller, S., Yang, J.Y.H. **2019**. Melanoma Explorer: a web application to allow easy reanalysis of publicly available and clinically-annotated melanoma omics datasets. *Melanoma Research*.

Seminars

Workshop
- *"Single-cell RNA-Seq analysis from beginning to end"* at Hong Kong University (2019), Sydney University (2019) and Cornell University (2020).
- *"Enter the Tidyverse with R and RStudio"* at BioinfoSummer (2019).
- *"Fast algorithms and modern visualisations for feature selection"* at Joint International Society for Clinical Biostatistics and Australian Statistical Conference (2018).

Talks
- 2019 International Conference on Econometrics and Statistics (invited)
- 2019, 2017 International Biometrics Society - Australasian Region Conference
- 2018 Australian Bioinformatics & Computational Biology Society Conference
- 2018 Joint Statistical Meeting
- 2017 International Conference in Robust Statistics
- 2016 Australian Statistical Conference