125 St Georges Terrace, Perth WA 6000

2023-01-25

To the team at BHP,

I wish to apply for your advertised role of a Specialist Data Analyst in Perth, WA (or Brisbane). I can provide your team with my years of experience in research. I am experienced in data analysis in a research context and am an enthusiastic student of biostatistics. I can provide BHP with my experience in processing biological data. I am capable of processing datasets in Excel, R, python, and am very comfortable using the unix command line and various CLI tools such as grep/ripgrep, awk, and vim.

The skills I can bring to BHP include:

• Experience in using R, python, the unix terminal, and building simple static websites  
• Experience accessing and processing data from online databases using APIs  
• Experience of processing complex population genetics datasets  
• Knowledge of Biostatistics and associated technical knowledge of study design.   
• Experience in translational medical research including cell culture, molecular biology, and animal husbandry

I am interested in this role as I wish to pivot into a data focused role.

I have experience working with university, research institute, and NHMRC bureaucracies.

Your team can also draw on my diverse experience as a science communicator, quality control technician, and as a biology tutor.

Additionally, I was recently responsible for setting up and managing a research program in precision medicine at the Hunter Medical Research Institute. This included setting up support programs and a conference (Cancer and Precision Medicine Program 2022).

See below for my response to criteria.

Yours Sincerely,

Adam Graham

ph: 0431587399

adam.graham@uon.edu.au

## Bachelor degree in a quantitative field

I am currently completing a Masters in Biostatistics. I also have a Bachelor of Science (Honours) in Biotechnology.

## Proficient in at least one programming language such as SQL, R, Python etc.

I am experiened in using both Python and R. I have some training in how to use SQL and am familiar with the SQL tables paradigm thanks to the equivalent [R::dplyr join functions](https://dplyr.tidyverse.org/reference/mutate-joins.html).

### python experience & projects

I have used python professionally and for self-learning. I have used python both professionally and for personal projects. A few recent projects are listed below:

* With permission, I have analysed the todo list I kept for a previous wet-lab biology research assistant role. I have used python to analyse the data and produced custom graphs using matplotlib and seaborn. You can view the project [here](https://adamdoescode.github.io/todo-parsing/)
* A webpage for my quotes collection, which required processing semi-structured text data and outputing a html table (which I used pd.DataFrame.to\_html() for). This project is available [here](https://adamdoescode.github.io/AdamsQuotes/)
* I have been sharpening my python skills by completing the [Advent of Code](https://adventofcode.com/) challenges. For this, I have mostly stuck to python’s standard libraries and have worked to make myself more comfortable with python’s class structures. You can view my solutions [here](https://github.com/adamdoescode/AdventOfCode2022Scribbles)

### R experience & projects

I have recently been exploring weather data around Perth using R. You can view this project [here](https://www.seek.com.au/job/59804986?type=standard" \l "sol=6ab57c110f832ff881da8c614b8af6a9c0d54ac2). In this project I have utilised ggplot and dplyr to perform visual exploration of the data. I have also run several simple linear models to explore the relationship between rainfall, temperature, and date. Beyond this, I have used R in my recent Research Assistant role to perform analyses of complex genetics datasets and other smaller datasets relevant to grants and administation. I also helped debug R code written by PhD students in my research group. Beyond this, I have used R extensively in my Masters of Biostatistics course. I have made extensive use of R’s built in statistical functions as well as utilising statistical packages to perform spline models and simple bayesian regression modelling.

## Experience delivering data storage solutions, data cleansing and analysis

I have experience dealing with data cleaning and analysis.

The research datasets I have used were stored as simple tables, text, or binary files, so I have limited experience with sophisticated data storage solutions. I do have experience accessing data via web APIs and have used resources such as [NCBI E-utilities](https://www.ncbi.nlm.nih.gov/books/NBK25500/), [IEU open GWAS data and related scripts](https://gwas.mrcieu.ac.uk/), and the NCBI SRA toolkit.

## Ability to communicate technical contents and complex ideas to non-technical audiences at all levels of the organisation including those in different time zones or locations

I have experience working remote, from Perth, for a research team in Newcastle, Australia. In this role it was important that I be able to convey what work I had done and produce clear, concise emails as I was unable to physically contact people. I did presentations on various small research projects that I did to help PhD students as well as an independent project. I have also given presentations to professional and lay-person audiences on my scientific research as well as engaging educational science shows designed to engage primary and high school students in science with liquid nitrogen and explosions.

## Ability to perform data quality and data sense checking with good attention to detail

I have experience working with diverse datasets ranging from multi terabyte genetics datasets to small datasets generated by my own output. These datasets often require various data cleaning steps to remove missing data, remove outliers, and to ensure that the data is in the correct format. I have experience performing data cleaning using python (missingno, pandas) and R (data.table, dplyr).

## Well-developed written and verbal communications skills

I have utilised my written and verbal communication skills during my time as a research student and as a science show presenter. I have helped proofread and review several published publications and successful grants. I have given presentations to professional and lay-person audiences on my scientific research as well as engaging educational science shows designed to engage primary and high school students in science with liquid nitrogen and explosions.

## Experience with a visualisation software (e.g. Spotfire, Power BI)

I am willing to learn appropiate visualisation software such as Spotfire of Power BI. Aside from these, I have experience using python’s matplotlib and seaborn libraries and also R’s ggplot2 library to create custom graphs for research projects, coursework, and personal projects.

## Experience in Microsoft PowerApps and Power Automate or a willingness to learn

I am willing to learn MS PowerApps. I am an enthusiastic self-learner and my python knowledge is entirely self-taught.

## Proficiency in MS Office, Excel, Word, PowerPoint

I have used MS Office programs extensively for administration and research. I have used Excel to perform data analysis and to create custom graphs. I have used Word to write reports and to create custom templates. I have used PowerPoint to create presentations and to create custom figures for my masters thesis and presentations.