Rob Hoelzle, PhD

Data Scientist and Engineer

Results-driven data scientist and engineer with 8 years' experience in big data analytics, 13 years' experience in high paced R&D environment, and a life-long passion for finding creative solutions to complex problems. Proficient in **Python**, **R**, and **SQL** for data mining, statistical analysis, visualization, and various machine learning models.

Skills & Techniques

Languages & Libraries	Python • SKLearn, SciPy, Folium, DB-APIs, SQLAlchemy, PyODBC, RegEx, BeautifulSoup, Dash, Matplotlib, Plotly, Seaborn R • Vegan, NBClust, RCy3, RStatix, mvabund, lsmeans, dplyr, tidyr, spaa, SpiecEasi, ggplot, sciplot, corrplot, igraph, stringr SQL • MS Server, Lite, Transact • Bash • C++
Modeling & Statistics	Linear/Non-Linear/Logistic Regression • Monte Carlo • K-Means Clustering • Hierarchical Clustering • KNN • Decision Trees • Support Vector Machines • Network Analysis • Dimensional Reduction • Ordination • Multivariate Statistics • ANOVA • PERMANOVA • T-test • Z-test • Chi-square Test
Tools	GitHub • Power BI • VS Code • Jupyter Notebook • R Studio • Adobe Suite • M365 Suite
Cloud & Systems	IBM Watson Studio • IBM DB2 • Galaxy Australia • Linux HPCs
Professional	Project Management • Teamwork • Team Coordination • Leadership • Creative Problem Solving • Technical Writing & Communication • Consulting • Engineering • Budgeting • Strategic Planning • Statistical Analysis • Data Mining • Predictive Modeling • Research • Experimental Design

Professional Experience

Senior Quantitative Analyst *Jan 2023 – Present*

Clean Energy Transfer Fund Pty Ltd, Brisbane, Queensland, Australia

- **Python**, **SQL**, **Power BI**, **Regression**, **Monte Carlo**, **Clustering**, **Statistical Validation**
- Market research and analysis of Australian VRE generators for originating PPAs, energy trades, and weather hedges via forecasting interactions between energy market, demand, and weather data
- Build **Power BI** dashboards for the trading desk, board of directors, and investors describing market conditions and asset performance, with highlights on portfolio risk conditions and new investment opportunities
- Identified and found a solution to $10M in forecast losses by **regression** and **cluster** modeling the performance of all NEM VRE assets against our own risk profile, generating ideal generic asset profiles, as well as a target short list
- Created automated **Python** pipelines to extract, transform, and load (ETL) key data streams (ASX futures, Copernicus weather, and UtiliView meter data) into **SQL** databases
- Wrote data management and security policy and updated company data management system to Australian Energy Sector Cyber Security Framework (AESCSF) standards

Data Scientist (Bioinformatics) – Team Lead *Sep 2021 – Dec 2022*

Data Scientist (Bioinformatics) *Aug 2016 – Aug 2021*

The University of Queensland (Faculty of Science), Brisbane, Queensland, Australia

- **Python**, **R**, **SQL**, **Bash**, **Dimensional Reduction**, **Regression**, **Clustering**, **Classification**, **Statistical Validation**
- Designed and ran data analysis and modeling pipelines in **Python**, **R**, **SQL**, and **Bash** on structured and unstructured genomic, environmental, agricultural, immunological, mining, electrochemical, and environmental engineering data
- Communicated findings and deliverables through consulting reports, presentations, and peer-reviewed publications
- Mentored and coached professional and student researchers on experimental design, data management and processing, predictive modeling, and statistical validation. Included running code reviews and workshops
- Solved groundwater cyanide issue for industry client using **dimensional reduction**, **KNN classification**, and **PERMANOVA** on genomic and environmental datasets
- Identified probiotic bacteria for an industry client using **dimensional reduction**, **multivariate regression**, **k-means clustering**, and **PERMANOVA** on immunological and genomic datasets
- Modeled future greenhouse gas emissions from thawing arctic permafrost using **k-means clustering**, **PERMANOVA**, and **multivariate regression** on environmental, genomic, and protein expression datasets

Brisbane, Qld, Australia | 04 2251 5081 | robert.hoelzle@gmail.com

linkedin.com/in/robert-hoelzle/ | rdhoelzle.github.io/

Bioprocess Engineer *Jun 2012 – Jul 2016*
The University of Queensland (Advanced Water Management Centre), Brisbane, Queensland, Australia

- **R**, **C++**, **Bash**, **Dimensional Reduction**, **Regression**, **Statistical Validation**
- Coordinated with research team to design experiments and communicate findings in peer-review publications
- Built **C++** programs to automate experiments and collect time-series data into pipeline-formatted data tables
- Processed, modelled, and statistically validated large, multivariate datasets to validate carbon biorecycling to bioplastics process, using **dimensional reduction**, **PERMANOVA**, and **multivariate regression**

Bioprocess Engineer *Jun 2010 – Mar 2012*
Green Biologics Ltd, Columbus, Ohio, USA

- **C++**, **Excel**, **Regression**, **Statistical Validation**
- Built **C++** programs to automate experiments and collect time-series data on pilot-scale biofuel system
- Designed successful commercial biofuel reactor from time-series data using **regression** modelling and **ANOVA**

Education & Certificates

PhD, Bioprocess and Chemical Engineering, The University of Queensland *2017*
Thesis: *Metabolic mechanisms and regulation of mixed culture fermentation*

BSc w/ Hons, Bioprocess Engineering, The Ohio State University *2010*
Hons Thesis: *Genetic improvement of* Clostridium tyrobutyricum *for butanol production by insertion of adhE from* Clostridium acetobutylicum

Professional Certificate, Data Science, IBM *2022*

Community Engagement

Data Science Blog on Medium: medium.com/@RDHoelzle *2022 – Present*
Personal blog for publishing data science projects and how-to articles

Mental Health Working Group, UQ Sch of Earth & Environmental Sci: Mentorship Committee *2020 – 2022*
Developed school-wide mental health, wellness, and mentorship programs

Joint Academic Microbiology Seminars (JAMS): Organizing Committee *2018 – 2022*
SE Queensland university and industrial microbiology networking and seminar series

March for Science, Brisbane: Organizing Committee *2017*
Coordinated international demonstration promoting scientific and data-driven governmental policies

Engineers Without Borders: UQ Institute Engagement Committee *2015 – 2016*
UQ research and student engagement arm of international engineering aid organization

Top Peer-Reviewed Publications

Genome-resolved metagenomics of milk microbiomes reveals the influence of maternal dietary fiber on neonatal inheritance of immunoregulatory traits. Primary author. **Research Square** (2023)

Substrate availability drives mixed culture fermentation of glucose to lactate at steady state. Primary author. **Biotechnology and Bioengineering** (2021)

A facile method to enhance the performance of soil bioelectrochemical systems using in situ reduced graphene oxide. Secondary author. **Electrochimica Acta** (2019)

Maternal diet modulates the infant microbiome and intestinal Flt3L necessary for dendritic cell development and immunity to respiratory infection. Contributing author. **Immunity** (2023)

Genome-centric view of carbon processing in thawing permafrost. Contributing author. **Nature** (2018)

Regulation mechanisms in mixed and pure culture microbial fermentation. Primary Author. **Biotechnology and Bioengineering** (2014)

Personal Interests

Rock Climbing • Ultimate Frisbee • Home Brewing • Escape Rooms • Photography • Gardening

Brisbane, Qld, Australia | 04 2251 5081 | robert.hoelzle@gmail.com
linkedin.com/in/robert-hoelzle/ | rdhoelzle.github.io/