

Robert Hoelzle, PhD

Data Scientist

Results-driven data scientist and engineer with 7 years' experience in big data analytics, 12 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 **Bash** for data extraction/cleaning/analysis, visualization, and various machine learning models.

Skills & Techniques

Languages & Libraries	Python • SKLearn, SciPy, Pandas, NumPy, BeautifulSoup, FuzzyWuzzy, Dash, Matplotlib, Plotly, Seaborn, Folium R • Vegan, NBCLust, RCy3, RStatix, mvabund, lsmeans, spa, dplyr, tidyr, SpiecEasi, ggplot, sciplot, corplot, igraph Bash • T-SQL • SQLite • C++
Modeling & Statistics	Linear/Non-Linear Regression • Logistic Regression • 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
Cloud & Systems	IBM Watson Studio • IBM DB2 • Galaxy Australia • Linux Architecture
Tools	GitHub • Atom • Spyder • Jupyter Notebook • R Studio • REST API • Adobe Suite • MS Office Suite
Professional	Predictive Modeling • Forecasting • Creative Problem Solving • Technical Writing & Communication • Project Management • Team Coordination • Leadership • Teamwork • Research & Analysis • Experimental Design • Consulting • Strategic Planning • Engineering • Budget Management

Professional Experience

Data Scientist (Bioinformatics) – Team Lead

Sep 2021 – Present

Data Scientist (Bioinformatics)

Aug 2016 – Aug 2021

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

- **Python, R, Bash, Dimensional Reduction, Regression, Clustering, Classification, Statistical Validation**
- Design and run data analysis and modeling pipelines in **Python, R**, and **Bash** on environmental, agricultural, immunological, mining, electrochemical, environmental engineering, and microbial genomic data
- Coordinate interdisciplinary research programs involving researchers and industry stakeholders to define key project deliverables and design experiments for targeted data collection
- Solved groundwater cyanide issue for industry client using **dimensional reduction**, **KNN classification**, and **PERMANOVA** on complex genomic and environmental data sets
- Identified probiotic bacteria for an industry client using **dimensional reduction**, **multivariate regression**, **k-means clustering**, and **PERMANOVA** on complex immunological and genomic data sets
- Built a **Bash**-executable **Python** program to extract HTML-formatted metadata from open-source databases into pipeline-formatted data tables, given a user-specified list of target variables
- Communicate findings and deliverables through consulting reports, presentations, and peer-reviewed publications
- Mentor and coach professional and student researchers on experimental design, data management, data processing, predictive modeling, and statistical validation. Includes code reviews, workshops, and brainstorming sessions

Bioprocess Engineer

Jun 2012 – Jul 2016

The University of Queensland (Advanced Water Management Centre), Brisbane, Queensland, Australia

- **R, Bash, C++, Dimensional Reduction, Regression, Statistical Validation**
- 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**, **multivariate regression**, and **PERMANOVA**
- Coordinated with research team to design experiments and communicate findings in peer-review publications

Bioprocess Engineer

Jun 2010 – Mar 2012

Green Biologics Ltd, Columbus, Ohio, USA

- **C++, 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**

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

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

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** 2022 – Present
medium.com/@RDHoelzle: Personal blog for publishing data science projects and how-to articles.
- Mental Health Working Group: UQ School of Earth & Environmental Sciences** 2020 – Present
Mentorship and Activities Committees: Working group to develop and implement positive mental health programs within the School of Earth & Environmental Sciences, including a mentorship program for early career researchers, networking events, and lunch and sporting activities.
- Joint Academic Microbiology Seminars (JAMS)** 2018 – Present
Brisbane Organizing Committee: Bimonthly networking and seminar series aimed at developing collaboration between academic, industry, and medical microbiologists across Southeast Queensland.
- March for Science: Brisbane** 2017
Organizing Committee: Coordinated international demonstration calling for increased scientific and data-driven governmental policy making, especially on issues of climate change and public health.
- Engineers Without Borders** 2015 – 2016
UQ Institute Engagement Committee: International engineering aid organization. The UQ Institute Engagement program aligned PhD candidates with UQ researchers working on infrastructure projects for developing nations, especially low-maintenance wastewater and drinking water facilities.
- Navigating the Realities of a Diverse Engineering Workplace: UQ School of Chemical Engineering** 2015
Organizing Committee: Panel Q&A session on workplace culture in professional engineering, with a focus on diversity of race and gender identity. The panel consisted of a highly diverse range of early, mid, and late career engineering professionals, and facilitated networking between engineering students and professionals.

Top Peer-Reviewed Publications

- Substrate availability drives mixed culture fermentation of glucose to lactate at steady state.* RD Hoelzle, D Puyol, B Viridis, D Batstone. **Biotechnology and Bioengineering** (2021)
- A facile method to enhance the performance of soil bioelectrochemical systems using in situ reduced graphene oxide.* C Camedda, RD Hoelzle, A Carucci, S Milia, and B Viridis. **Electrochimica Acta** (2019)
- Genome-centric view of carbon processing in thawing permafrost.* BJ Woodcroft, CM Singleton, JA Boyd, PN Evans, JB Emerson, AAF Zayed, RD Hoelzle, TO Lamberton, CK McCalley, SB Hodgkins, RM Wilson, SO Purvine, CD Nicora, C Li, S Frolking, JP Chanton, PM Crill, SR Saleska, VI Rich, GW Tyson. **Nature** (2018)
- Influence of pH regulation mode in glucose fermentation on product selection and process stability.* Z Mohd-Zaki, JR Bastidas-Oyanedel, Y Lu, R Hoelzle, S Pratt, FR Slater, and DJ Batstone. **Microorganisms** (2016)
- Regulation mechanisms in mixed and pure culture microbial fermentation.* RD Hoelzle, B Viridis, and DJ Batstone. **Biotechnology and Bioengineering** (2014)

Personal Interests

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