GUNDALAI "DALE" BATKHUU
Data Analyst

☎ 0405421982 ✉ gdbt0623@gmail.com
📍 Canberra, ACT ⚙ github
                  in Linkedin

SUMMARY

Enthusiastic Graduate Data Analyst with a passion for developing creative solutions to complex problems. Solid experience in data modelling, processing and visualisation within modern cloud environments.

SKILLS

Languages: Python, Java, R, PostgreSQL, Terraform (HashiCorp).
Technologies: AWS, Data Modelling, Data Warehouse, MLOps.

EXPERIENCE

2/2023 – 7/2023 AWS Data Engineer Intern   **Shine Solutions Group**

- Designed and developed an AWS-hosted end-to-end ETL and data analytics pipeline, enabling the streamlining of data ingestion, transformation, validation, storage, and analytics processes.
- Promoted Infrastructure as Code (IaC) best practices by utilising Terraform modules to automate AWS resource deployments, encompassing multiple development environments.
- Developed a CI/CD pipeline with CodePipeline, streamlining the validation and deployment of Terraform configurations, significantly improving the reliability and integrity of infrastructure.
- Implemented a serverless ETL pipeline using Step Functions and Lambda, focusing on automating the validation, transformation, and partitioning of large datasets for efficiently processing data for analytics use, with robust error handling and user notification features.
- Developed a PostgreSQL database on AWS RDS for storing the processed data, incorporating data modelling and normalisation principles in the design of data schemas, with a focus on query performance optimisation.
- Leveraged Athena to efficiently query data from the database and utilised QuickSight to design and develop dynamic dashboards, facilitating data-driven decision-making.

```
AWS  / Terraform / Python / PostgreSQL
```

6/2022 – 9/2022 Data Science Intern   **Curious Machines AI**

- Developed end-to-end data science pipeline with Kedro MLOps framework to analyse extensive real estate listing data to derive insightful trends and patterns.
- Performed ETL operations with Python by organising data processing logic into more manageable layers.
- Utilised fastAI's pre-built deep learning models for predictive analytics and classification tasks.

```
Python
```

11/2021 – 1/2022 Research Assistant   **National Cancer Centre**

- Assisted with collection, curation and analysis of cancer research data with Python scripts and Power BI, optimising research workflows.
- Normalised legacy spreadsheets with clinical data into a relational database with SQL Server, modernising the data environment.

```
Python  / MySQL / PowerBI
```

1/2020 – 7/2020 Data Analyst Intern   **BT Financial Group**

- Assisted with market research by leading a team of nine interns to supplement investment analysis and decisions.
- Conducted quantitative research to assess fair value, risk and economic capital of an investment portfolio.

```
Python  / Excel
```

12/2018 – 2/2019 Undergraduate research assistant   **Nagoya University**

- Participated in international research on hazardous air pollution in urban areas to identify sources of pollutants.
- Assisted with analysing the environmental data and prepared visualisations for scientific reports.

```
Excel
```

EDUCATION

7/2021 - 12/2023 **Master of Computing**   **The Australian National University**
Machine learning and software development

7/2017 - 12/2019 **Bachelor of Applied Finance**   **Macquarie University**
Business statistics, econometrics and financial modelling

CERTIFICATES

AWS Cloud practitioner, Azure Data Fundamentals

ACHIEVEMENTS

2nd place **Monash University and NAB Data Hackathon 2023**
Built a Streamlit web application that predicts the quantity of waste generated by Australian states using an ensemble learning model with 0.96 accuracy.

2nd place **Macquarie University Datathon 2023**
Created an AWS-hosted web application that predicts physical injury recovery timeline using a gradient boosting regressor model with 0.94 accuracy.