

Jack Davidson

Brisbane, QLD 4006 • (+61) 403 178 709 • Jacknikolaidavidson@gmail.com
<https://www.linkedin.com/in/jacknikolaidavidson/> • <https://www.jackndavidson.com/>

Technologies and Skills

- **Languages:** Python, R, SQL
- **Databases:** OLTP (MySQL, MS SQL Server, PostgreSQL), OLAP (Snowflake), Geospatial (PostGIS)
- **Cloud:** AWS (ECR, S3, EC2, Aurora, API Gateway, Lambda, SageMaker, Step Functions, CodePipeline)
- **Orchestration:** Airflow
- **ETL:** dbt, SQL, Alteryx, Knime
- **CICD:** Jenkins, Azure DevOps
- **Source Control:** GitHub, GitLab, BitBucket
- **Data Visualisation:** Tableau, Power BI, Metabase
- **Virtualisation:** Docker
- **Machine Learning:** TensorFlow, PyTorch, Keras, Hugging Face, scikit-learn
- **RPA:** Selenium, Blue Prism, UiPath

Industries and Summaries

- **Financial Services:**
 - Intelligent automation, electronic anti-money laundering, regulatory intelligence
 - ML feature engineering/pipeline development (NLP, topic modelling, sentiment analysis and classification models)
- **Government:**
 - AWS cloud solution design and implementation with DevOps and CICD
 - Developing EO geospatial processing/ML solution utilising satellite imagery
 - Productionised business intelligence solutions (reporting, analytics, self-service)
- **Transport/Infrastructure:**
 - Statistical modelling and analytics
 - Asset risk management, identification and predictive modelling
- **Healthcare:**
 - End-to-end data engineering/model refinement and visualisation
 - Simulation/agent-based modelling
- **Mining:**
 - Inventory analytics and reporting solution design and implementation leveraging AWS, Tableau, Snowflake, Python, Airflow and dbt in agile CICD development framework using Git

Education

The University of Queensland

Bachelor of Mathematics/Commerce

January 2015 – June 2019

Majors: Data Analytics & Operations Research/Finance

Professional Experience

Deloitte

Consultant – Data & AI – Advanced Modelling & ML

January 2020 – Present

- Designing, implementing, and managing AWS cloud solutions (ECR, S3, EC2, Aurora, ELB)
- Implementing geospatial solutions utilising Docker, PostgreSQL, TileDB, GDAL, PostGIS and Python for processing
- Implementation of production grade, end-to-end ELT pipelines and data models using dbt, Snowflake, Airflow and Jenkins
- Designing, implementing business intelligence solutions in Tableau/Power BI for various solutions

EY

Consultant – Data & Analytics (EY DnA)

November 2018 – January 2020

- Intelligent automation, electronic anti-money laundering, regulatory intelligence ML feature engineering/pipeline development
- Delivered productionised, business intelligence solutions to streamline reporting, provide actionable insights and provide self-service analytics

Australian Government Department

Data Engineer - Satellite Imagery Analysis and Insight Tool (DR SAT)

Current

Developing a digital tool to help farmers self-assess their resilience against environmental, economic, and social indicators.

- Managing AWS platform resources (ECR, S3, EC2, Aurora, API Gateway, Lambda)
- Developing serverless (Lambda) Python based geospatial API leveraging geospatial libraries (GDAL, Rasterio, TileDB)
- Building end-to-end ELT pipeline and data model using dbt
- CI/CD using Azure Pipelines

Large Australian Bank

Geospatial Developer & Test Analyst – Regional Market Research Assessment

1 Month

Explored Testing performance of various architectures to deliver cloud geospatial capabilities and prove performance of using raster's in a hybrid serverless environment.

- Developed architecture utilising PostGIS and S3 to store, index and access raster data
- Developed Java based REST API using GDAL Java bindings to query point values
- Tested the suitability of different combinations of cloud storage options (S3, EBS), raster formats (GeoTIFF, COG, VRT, MRF), compressions and tile sizes to deliver required performance using JMeter

NSW Government Department

Data Engineer & Business Intelligence Developer – Regional Market Research Assessment

3 Months

Developed a solution to profile regional areas and provide insights on optimising job advertisements based on various attributes/demographics and supply/demand of an area's workforce.

- Translated high-level requirements to wireframe designs for design workshops
- Developed geographical data model and forecasting methodology for regional profiling
- Developed data model and ETL pipeline in SQL to integrate data sources
- Designed, implemented, and presented suite of Tableau dashboards to provide insight on job market supply/demand in regional areas

Global Mining Company

Data Engineer – Work Management & Inventory Obsolescence

4 Months

Developed an inventory advisory system leveraging analytics to alert inventory controllers to excessive or at-risk stock.

- Developed tool leveraging AWS, Tableau, Snowflake, Python, Airflow and dbt in an agile development framework using Git
- Built production, end-to-end ELT pipeline and data model using dbt, Snowflake, SQL using Jinja framework
- Developed CI/CD with Jenkins (Dev to Production)

Health/Pharmaceutical Company

Data Scientist – Customer & Workforce Segmentation

1 Month

Built unsupervised learning pipeline to cluster employees based on survey question responses to verify employee archetypes. Conducted principal component analysis for feature importance and to clustering based on question importance.

- Built machine learning pipeline with SciPy/Scikit, Pandas, Numpy, Postgres
- Explored different unsupervised training models assessing correlation between questions and groups
- Explored emergent cluster types, validated/presented accuracy, and logic with stakeholders
- Performed demographic analysis resulting from hierarchical clustering and Self Organising Maps (SOMs)
- Developed suite of Tableau dashboards providing insight into interactions between employee clusters, survey questions and defining clusters features

Global Health Advisory Body

Data Engineer & Business Intelligence Developer - COVID-19 Response Strategy

3 Months

Developed suite of Tableau dashboards enabling government and business leaders to understand the intervention outcomes on both pandemic management and on their own country's economy or business.

- Delivered components on technical requirements gathering, data engineering, model refinement and dashboard build (Tableau)
- Implemented data models using SQL and agent-based/scenario modelling in R and orchestration/ETL using Kame

EY - Key Projects

Large Australian Bank

Machine Learning Engineer – Anti-Money Laundering, Politically Exposed Persons (PEPs)

2 Months

- Developed Selenium web scraper in Python for structured/unstructured data sources
- Performed natural language processing for EAML purposed to identify named entities, perform sentiment analysis, and produce recommendations for unstructured texts
- Developed machine learning NLP transformer model/pipeline in TensorFlow utilising project specific data to augment pre-trained models for sentiment analysis and named entity recognition
- Fine-tuned pretrained Hugging Face models based on variations of BERT/XLM/GPT model architectures

Large Australian Bank

Intelligent Automation Consultant - Business Banking Process Optimisation

3 Months

- Developed Robotic Process Automation (RPA) solutions in Blue Prism for several use cases throughout the client's asset, compliance, customer, and finance divisions
- Developed Power BI dashboards to track value/benefits of RPA projects
- Liaised with multiple vendors (Automation Anywhere) regarding new automation products, introduction of an online form and API product integration within the bank's current RPA environment

Large Australian Council

Business Intelligence Developer – Internal Audit Reporting Analytics

6 Months

- Gathered business requirements and key KPIs from main stakeholders for different audit tests for risk prioritisation
- Current state process flow mapping, condensed and migrated multiple source systems into a single database
- Developed data models for reporting including key KPIs and measures ingested from Oracle databases
- ETL multiple data sources across the organization using SQL
- Developed, tested and prioritised test hypotheses aimed at delivering business value and mitigating the highest risks
- Designed and implemented 10 interactive Power BI dashboards to visualize different risks, identify and prioritise exceptions and achieve the objectives of continuous assurance using data analytics

Global Investment Bank

Data Analyst – Indirect Tax/Process Automation, Data Analyst

2 Months

- Developed RPA solution in UiPath to scrape and download financial statements from online business intelligence tools storing within Excel workbooks and developed a solution in Python to analyze the statements and find 'comparable' companies for further analysis
- Developed Alteryx workflow that condensed disparate books of financial derivatives from multiple sources summarizing payment dates and accrued amounts for each source, division and sub-division
- Developed Tableau dashboard and stories which integrated workflow outputs to provide insight for future and compliance purposes for each hierarchical layer

Other - Key Projects

Australian Road Research Board

Data Analyst Intern – Infrastructure Research, Modelling & Analytics

6 Months

- Conducted research on the viability of new types of structural information coming from under-utilised data sources such as pavement data
- Developed explanatory models allowing structures at risk to be identified quickly and efficiently from statistical analysis using Python
- Developed predictive models used to validate cost benefit analyses across asset life cycles using Python
- Used Power BI and ArcGIS to assist reporting and provide results into an interactive dashboard for internal/external client use