**Essential criteria**

**Data scientist experience: •Proven experience as a Data Scientist or Senior Data Analyst. Individuals with more experience may assume additional responsibilities. •Analytical mind and business acumen. •Strong math skills (e.g. statistics, algebra). •Problem-solving aptitude. •Developing and maintaining a product backlog, including user stories and acceptance criteria.**

I am a Data Scientist and Data Engineer at the Department of Agriculture. I work across teams and have developed detailed custom scripts (SQL – SSMS) extracting data from ONPREM enabling program APIs to harness that data hosted in Azure-based services. I note that I played a key role in the architecture and enabling to program’s use of Azure Data Factory. This involved engagement with cyber security, developing schemas, infrastructure, developing data connections and ETLs, pipeline and data debugging, Integration Runtimes and Linked Services. I developed documentation and technical and strategic presentations. ETLs for three program line areas are in operation.

I was a Consulting Data Scientist for IP Australia modelling, using primarily Python (AWS), the impact of IP applications during Covid-19 pandemic. I developed a range of statistically-based machine learning processes that revealed the pandemic positively impacted applications in Australia at a total level and also revealed by industry and region where change occurred. These models enabled IP Australia to consider future funding and staffing requirements.

At PwC I developed a bespoke, or lite, version of the FPA that employs all major compliance and data quality tests (usually in a SSAS/Power BI-based framework) and instead uses a SQL/Python data model (results and compliance tests visualised using Bokeh framework). The primary reason for the lite version was client budget and their data security requirements (where all data was required to be local rather than the usual AWS-hosted environment).

I was the senior developer for PwC’s Financial Processes Analyser (FPA) Tool.  This tool was a well-developed, flexible, yet robust, cloud-hosted SSAS/Power BI data model providing for low-cost updates of the tool.  My responsibilities are i) SME for development, testing, deployment, maintenance; ii) rhythm – budget, staffing, technology; iv) client engagement and updates; and v) client satisfaction. The tool has four major agencies currently subscribed to monthly or quarterly refreshes with most updates operating in parallel. To manage this load on the business an automated ETL process is structured using virtual servers, SSAS Tabular connected to Power BI and accessed by the client via a secure AWS-hosted interface.

As technical SME I was part of a team that evaluated the value-for-money (and compliance) of maintenance services for a major national rental housing firm. The review of source data showed a variety of formats (SAP, Excel, pdf), conventions, key types, and mixed internal ETL processes. Attending to some these risks, a model was designed and ETL process was developed to transfer the data to Power BI (for additional transformation and analysis including visualizations). The project provides the client with a framework that statistically assessed (ie cluster, outliers, text) property maintenance costs from data that is uploaded. Issues in their data has also commenced a systems remedy project for the way the data is handled on their side.

At Austrade I developed and implemented a data model for a business grants assessment process to score applicants’ risk profiles using Azure ML. A decision tree model allowed the assessment team to focus primarily on high-risk applicants. With the PoC developed it was redeveloped in SQL Server, R due to cost and security.

**Technical Specialisation: •Experience using data visualisation and reporting tools (e.g. Microsoft Power BI, Qlik, SAS, Tableau). •Technical expertise regarding data models, database design development, data mining and segmentation techniques. •Understanding of machine-learning and operations research. •Knowledge of R, SQL and/or Python.**

Strong knowledge of database structure systems, data analytics, data mining and data modelling techniques.

Documented minimum of 5 years’ experience in data analysis and management, with excellent analytical problem-solving abilities.  I have over 15 years in data analysis and data management, both government and consulting. I have qualifications in economics, finance, and statistics and have analysed the economy with data and policy analysis of education and employment. I have led technical data and IT meetings, developed project costings, and provided analytical briefings Ministers and Executive Teams.

As Principle Data Scientist in Austrade’s ITS Division I was part of a team that delivered a business intelligence tool for the agency’s Executive to analyse agency resources. My role was both technical project manager and developer. Primary project requirements were self-service visualisation (via Power BI), internal and external data sources, restricted access (ie Executive), and low-cost operation (ie automation processes). The tool was required to be compliant with the agency enterprise technology (MS) stack.

I was SME for data modelling and the data was from a range of internal and external sources - SharePoint, TM1, SAP, and the agency CRM. A data extraction tool (in Python) was developed to automate update processes for DET, ABS, UN, and IMF data that was integrated with agency data. An authentication layer provided security, namely, access.

I led the development of the data modelling: cluster analysis, text analysis, and a prescriptive model (ie three predictive options – trend, +5%, +10%) for the Executive to base strategic decisions (R plug-in within Power BI). The tool was taken deployed for use on time and to budget. Additionally, the Department of Agriculture’s request for Austrade to develop a PoC for them (using agriculture data).

Also refer Selection Criteria 1 and (where I initiated the transformation of user interaction with Austrade’s Australia’s International Business Survey - SQL, R, Tableau).

**E3. Report development experience.**

My government and consulting experience requires me to work with, and develop, trusting relationships with clients, data and IT teams, and project teams. I have broad experience working with data and IT teams ranging from access and delivery, procurement, infrastructure, design and UI, local and cloud hosting, through to data risk and cyber security.

I have developed a range of visualisation products, notably Power BI, Tableau, Qlik, Python, and Excel/VBA. At Austrade I led the introduction of Power BI (initial use was Tableau) through a cost-benefit analysis and recommendation paper to the Executive.

For a (confidential) project I was co-SME with a team of analysts (based in Canberra, Adelaide and Sydney) to develop of a Qlik-based tool that analysed payroll operations and any liabilities. Important steps in the process were:

1. kick-off meeting including agreement on scope
2. user-centered workshop(s) focusing on user, design, functionality, and operation: 1) COO, CFO, CHR, 2) select EL2s and EL1s, 3) data teams (payroll)
3. deep-dive with the client’s legal and IT team (data security and systems)
4. setting security requirements
5. developing a data request and building a project-specific VPN with the client’s IT team
6. Alteryx ETL and Qlik testing (on client site)
7. cloud-hosted transfer of data (with sample testing for any error)
8. building the data model (Alteryx) with analysis in Qlik
9. iterative development data visualisations and tests through client workshops
10. analytical report (using common-sense language and charts)
11. delivery and workshops.

The project’s analysis, and Qlik product, provided the client with the analysis they were seeking however this was only possible through an actively promoted and led design process. Open and transparent engagement with their IT team as well as PwC’s IT team as facilitated two-way flow of data and analysis).

I am highly skilled in developing solutions using Azure Data Factory – ETLs to extract from ONPREM to ADF (using basic extracts through to custom queries with triggers). I have built Data Flow solutions to model data. These solutions have included detailed documentation posted on DevOps and Git Hub.

In my current role I have developed ONPREM and cloud solutions (Azure). I have experience in developing ETL solutions that have required base infrastructure and service accounts. I have developed and proposed solutions where administrators and security can confirm tokens, map devices, containers – both internal and external to Commonwealth guidelines.

At PwC I deliver tools for clients to dynamically visualise their financial and resource data. One service is the Financial Processes Analyser (FPA) which relies on a multi-technology stack (loading client data from AWS to Alteryx to SQL Server through to Power BI). I both lead and am part of specialist team – made up of consultants through to technology SMEs - to customised pipelines (for SAP, TechOne, Expense8, TM1), SQL Server, set up cloud-based VPNs, deliver workshops, and manage project delivery. For a recent FPA project I was involved in all stages:

1. I led the required Commonwealth Data Protection Plan. This was a PwC-first I met with the client’s IT team, included advice from ANAO (ie cyber security) and the Australian Government’s PSPF and ISM, and managed compliance with PwC settings. The Plan provides assurance ensuring risks within the data pipeline are minimised (collection, processing, privacy, delivery, and storage), responsibilities are understood, and a multi-stack data model delivers compliance testing capability of large data sets in real time.
2. Customised a set of statistical tests to comply with client data. This included outlier identification and cluster analysis of expenditure, expenditure patterns (division, branch, section), text analysis, and bottom-up prediction of expenditure (based on last four quarters of data).
3. Developed a data model from scanned pdf invoices (UiPath). Data was matched to data originating in TechOne for compliance testing, namely approvers of contracts and approvers of the invoices. Due to various factors not always matching primary factors were amounts, and supplier, and ABN with matching techniques applied to the name of the approver of contract and invoice (for compliance checking).
4. Refined the monthly refresh process - primarily steps in the ETL and reporting to identify data issues (this halved the update process).

I have additional coding experience across a range of sectors including, but not limited to:

* visualisation tool for SAP/Aurion (SQL, Python)
* activity-based costing of internal resources (TM1/SAP in Python)
* global model of agency resource allocation including forecasts (SQL, Python, Power BI)
* simulation model of the education sector (SQL, Python, Excel/VBA)
* decision tree analysis of the Export Management Development Grants program (SQL R, Excel/VBA)
* activity-based costing of labour and skills policies (R, Power BI)
* web-page data extraction and text-based analysis of shared economy (Python - Pandas, Numpy, Beautiful Soup, SciKit, NLTK/SpaCy, Seaborn; Excel, HTML)
* per-capita income analysis of Australia’s international education sector (Excel/Tableau)
* the annual satisfaction survey of Austrade clients (Excel, R, Power BI).
* regional expenditure estimates of tourists (SAS)
* TRA Tourism Forecasts (SAS, EVIEWS, Excel/VBA)
* Tourism Investment Monitor (SAS, R)

**Well-developed team collaboration and communication skills: •Demonstrated ability to work under limited direction and accountable for completion of work within timeframes. •Demonstrated ability to work closely and effectively within an Agile multi-disciplinary team environment including liaising with diverse stakeholders.**

Demonstrated ability to work under limited direction and be accountable for completion of work.

Demonstrated ability to work closely and effectively within an Agile multi-disciplinary team environment including liaising with diverse stakeholders with competing priorities. Demonstrated ability to adapt and work in a constantly changing environment.

In my current role I have been required to present technical cloud-based solutions to technical administrators. I regularly met with my team through daily stand-ups and individually to discuss issues such as progress, methodology. I regularly met with the project manager and the project sponsor on progress and analysis. I met with my Deloitte partner and director on advice and issues. I engaged with external stakeholders on data and systems. Prior to the inter-departmental committee (IDC) meeting I advised on additional analysis on three risk areas found in our data analysis; I advised the client that further examination of these risks was required to provide documented assurance and independence. Using the analysis I led a successful argument at the IDC meeting for a shift from volumes-based reporting to volume and value-based reporting (which has been implemented by the client). Separate advice on data governance issues within the client’s Power BI processes was developed for the client and implemented. Refer: <https://www.niaa.gov.au/sites/default/files/publications/third-year-evaluation-indigenous-procurement-policy.pdf>

**Demonstrated ability to adapt and work in a constantly changing environment. •Previous experience with training and/or mentoring staff.**

I have over 15 years in data analysis and data management. I have strong team experience both as project manager and SME. I regularly lead technical meetings on IT and data issues, develop project costings, and provided analytical briefings to Executive and business leaders. I ask questions and address issues that either are, on may become, roadblocks.

I led the ‘Third Year Review of the Australian Government’s Indigenous Procurement Policy’ (IPP) at Deloitte and was responsible as delivery manager and principle data scientist. I managed a team of economists, accountants, and data scientists using a range of program and assessing agency performance against policy objectives (using SQL, R, Python, Azure, AWS, Power BI). Engagement with PM&C IT, Finance, and Supply Nation was key to obtaining industry and grants data to be integrated for analysis. I mentored data analysts on analytical processes, develop and analyse a survey of businesses (analysis in R and Power BI), and develop an alternative measure for compliance reporting. Government procurement data (AusTender) was extracted using an adapted Python-scripted process that joined national procurement data with state-based procurement data (as well as data from ASIC and other public sources). Analysis was undertaken in R and Power BI.

**Well-developed communication skills: •Excellent written, verbal and interpersonal communication skills. •Demonstrated experience in conducting workshops with business users and working effectively in a team environment. •Demonstrated presentation and training skills to stakeholders at all levels. •Negotiation and liaison skills.**

In my current role I meet with my team and other program stakeholders on a daily basis. As Data Scientist I engagement with executive on project aspirations and break down strategic goals into areas of operation and technical requirements. I regular engagement with coders of best practice and have reviewed other scripts, data warehouse/lake architecture (especially relating to Azure infrastructure). I am seen as a go-to on advice relating to databases, security and clearances.

As Senior Data Scientist and Project Manager (Deloitte) I regularly met with my team through daily stand-ups and individually to discuss issues such as progress, methodology. I regularly met with the project manager and the project sponsor on progress and analysis. I met with my Deloitte partner and director on advice and issues. I engaged with external stakeholders on data and systems. Prior to the inter-departmental committee (IDC) meeting I advised on additional analysis on three risk areas found in our data analysis; I advised the client that further examination of these risks was required to provide documented assurance and independence. Using the analysis I led a successful argument at the IDC meeting for a shift from volumes-based reporting to volume and value-based reporting (which has been implemented by the client). Separate advice on data governance issues within the client’s Power BI processes was developed for the client and implemented

**Desirable criteria**

**BSc/BA in Computer Science, Engineering or relevant field; graduate degree in Data Science or other quantitative field.**

|  |
| --- |
| * Master of Applied Statistics, Australian National University |
| * Bachelor of Commerce (Finance, Economics), University of Canberra |
| * Leadership Program, Melbourne Business School (Mt Eliza) |

**Experience working in working in an Agile, multi-disciplinary team.**

Demonstrated ability to work closely and effectively within an Agile multi-disciplinary team environment including liaising with diverse stakeholders with competing priorities. Demonstrated ability to adapt and work in a constantly changing environment