



Luis de Sousa

DATA, ANALYTICS AND AI CONSULTANT

Johannesburg, South Africa

☎ tel: +27835908454 | ✉ luisd@syep.co.za | 📱 luisdza | 🌐 luisdza | 🐦 luis_de_sousa

Profile

With over 20 years of experience in data and analytics consulting, I specialise in designing and building tailored data solutions for large organisations. My expertise spans data warehousing, integration, analysis, visualisation, and deployment in production environments, with recent experience in developing generative AI solutions. Additionally, I excel in leading cross-functional teams in dynamic settings. I firmly believe in the power of technology to drive sustainable change and am dedicated to problem-solving and fostering collaboration and transparency. I take pride in finding unique and practical solutions to challenges. Outside work, my interests include reading, exercising (CrossFit and running), dancing, exploring note-taking systems, and co-organising user groups.

Professional Experience

Syep

Johannesburg, South Africa

PRINCIPAL CONSULTANT AND DIRECTOR

May. 2007 - Present

- I lead high-impact data and analytics initiatives across diverse sectors such as real estate, food and beverage, education and banking. My managerial responsibilities encompass prospecting, negotiating, and overseeing business and project management. Additionally, I contribute to team development by recruiting, training, and mentoring staff. My technical expertise includes market analysis, consolidating financial data on cloud platforms, forecasting, and developing Management Information Systems (MIS) architectures.

Skills

Data Engineering and Analytics Expertise

ARCHITECTING END-TO-END DATA SYSTEMS

- Proficient in designing and implementing robust data solutions, including data warehousing, integration, and analytics.
- Skilled in using SQL, R, Python, Microsoft Data Platform, Azure DevOps, DuckDB, MySQL, Google BigQuery, Databricks, Github Actions and FiveTran.
- Experienced in consolidating financial data, conducting market analysis and forecasting, performing spatial data analysis, and integrating GIS data into comprehensive data systems.

Project Leadership and Management

LEADING CROSS-FUNCTIONAL TEAMS IN DYNAMIC SETTINGS

- Demonstrated ability to lead high-impact data and analytics projects across diverse sectors like real estate, food and beverage, and banking.
- Expert in overseeing project scopes, timelines, and deliverables while mentoring and training cross-functional teams.
- Effective in stakeholder management, negotiating contracts, and driving team performance.

Community Engagement and Thought Leadership

LEADING CROSS-FUNCTIONAL TEAMS IN DYNAMIC SETTINGS

- Actively involved in co-organising local data science communities, including Johannesburg R User Group and PyData Johannesburg.
- Regular presenter at conferences, sharing insights on data analytics, machine learning, and cloud technologies.
- Contributor to open-source projects, enhancing knowledge sharing and collaboration within the tech community.

Education

Wits

Johannesburg, South Africa

MASTER OF ENGINEERING (MENG) IN SOFTWARE AND INFORMATION ENGINEERING

2015

- Investigation Report: Using agile principles during software procurement in build versus buy decision-making.

User Groups

Johannesburg R User Group	<i>Johannesburg, South Africa</i>
CO ORGANISER	2018 - present
<ul style="list-style-type: none">This group serves as a platform for sharing knowledge, networking, and promoting the use of R in various statistical, data analysis, and programming endeavors within the Johannesburg area. https://www.rusergroup.co.za	
PyData Johannesburg	<i>Johannesburg, South Africa</i>
CO ORGANISER	2023 - present
<ul style="list-style-type: none">This local chapter of an international initiative focuses on nurturing community growth and engagement through Python and data science topics tailored to Johannesburg. https://www.pydata.co.za	

Presentations

A full list of my technical presentations can be found on Sessionize.

Oxcon	<i>Johannesburg</i>
COMPRESSING TRUST - UNPACKING THE XZ-UTILS BACKDOOR	2024
PyConZA 2024	<i>Cape Town</i>
DUCK, DUCK, PYTHON - OLAP DATA WITH DUCKDB	2024
Cape Town MS Developer User Group	<i>Cape Town</i>
RAG AND GRAPH: BETTER TOGETHER FOR NEXT-LEVEL GENAI WITH GRAPHRAG	2024
Microsoft Season of AI	<i>Johannesburg</i>
ENHANCING RETRIEVAL AUGMENTED GENERATION WITH GRAPHRAG	2024
Data and AI Community Day	<i>Johannesburg</i>
BUILDING A RAG APPLICATION WITH AZURE OPENAI AND MICROSOFT FABRIC	2024
AI/ML and Data Infusion Roadshow	<i>Durban</i>
DEPLOYING A LLM WITH VECTOR SEARCH IN DATABRICKS	2023

Consulting Projects

Financial Services	
DATA CAPACITY SERVICES	2024
<ul style="list-style-type: none">As a senior data engineer and architect, I contributed to designing a data product capability. I played a key role in optimising and expanding the 'next best action' as a service capability.As a senior data consultant, I designed and implemented a credit risk analysis process for a loan book. This project involved using R for data management and advanced analytics, yielding critical insights into the loan portfolio's performance and risk metrics.	
Education	
AI AND DATA ENGINEER	2024
<ul style="list-style-type: none">As a senior data consultant, I designed and implemented data solutions on Azure Synapse Analytics and Microsoft Fabric. I built a Retrieval Augmented Generation (RAG) capability using a vector database, Azure OpenAI and Python to enhance student support through data-driven insights.	
Policy and Advocacy	
DATA CAPACITY SERVICES	2024
<ul style="list-style-type: none">As a senior data consultant, I created data structures that allowed researchers to analyse South Africa's property markets. I used DuckDB for data management and R for spatial analysis to generate insights across national and metropolitan areas.	
Real Estate	
CONSOLIDATED FINANCIAL WAREHOUSE	2023
<ul style="list-style-type: none">As part of a two-person senior consulting team, we created and implemented a unified overview of the company's financial performance across all its entities. We used financial systems like Xero, JDE, and LOB applications. For this project, we employed FiveTran for data integration, BigQuery for data storage and analytics, Looker Studio for data visualisation, and Google Sheets for data sharing and collaboration.	

Health and Fitness

DATABRICKS PROOF OF VALUE

2023

- As a senior consultant in a two-person team, I co-engineered and implemented a Databricks solution in the cloud for both development and production environments. This project involved using technologies such as Azure Databricks, Azure Data Lake Storage Gen2, Azure DevOps, KeyVault, Data Factory, and PowerBI.

Food and Beverage

FORECASTING CAPABILITY BUILDING

2021-2022

- As the project lead and senior consultant on a two-person team, we developed and implemented an automated statistical forecasting system for SKUs within a valuable supply chain. For this initiative, we used R and Microsoft SQL Server.

Banking Innovation

MIS SOLUTIONS ARCHITECTURE

2010-2020

- As the lead and senior consultant for a five-person team in the innovation business unit, I managed all Management Information. My responsibilities included reporting and analytics for various payment systems like domestic and cross-border remittances, e-wallets, mobile payments, and banking transactions conducted via the Internet, apps, ATMs, and USSD. I also oversaw the integration of third-party systems, point-of-sale devices, prepaid services, and transit products. Analysed geospatial data to enhance insights into customer behavior and transaction patterns. Our efforts extended to regulatory reporting, fraud analytics, sanctions screening, and performance monitoring. We developed, maintained, and updated a data warehouse utilising the Microsoft Data Platform, employing technologies such as Microsoft SQL Server, Integration Services, Analysis Services, Reporting Services, and R.

Entertainment

SOFTWARE PROJECT LEAD

2007-2010

- I played a crucial role in supporting and revamping a royalty calculation engine. This pivotal project involved developing a C# front-end to manage complex financial processes seamlessly integrated with the SAP ERP system. We utilised the Microsoft Data Platform and leveraged Microsoft BizTalk for seamless integration.

Accreditation Courses

2024	Talend Data Health Explorer	Talend
2024	Solutions Architect Essentials	Databricks
2023	Generative AI Fundamentals	Databricks
2023	Azure Databricks Platform Architect	Databricks
2023	AWS Partner: Cloud Economics Accreditation	AWS
2023	Databricks Platform Administrator	Databricks
2019	RStudio Professional Administrator Certification	R Studio
2018	R Software Fundamentals with Applications to Statistical Analyses and Modelling	Wits
2009	CMMI-DEV for Systems Acquirers Workshop	ISACA
2009	Accelerating Process Improvement	ISACA
2009	Acquisition Supplement for CMMI V1.2	ISACA
2009	Introduction to CMMI V1.2	ISACA
2009	TOGAF 9 Foundation and Certified (Level 1 and 2)	The Open Group
2004	Implementing Data Protection on Dell Storage - IDPDS	Dell
2004	Implementing Dell Enterprise Storage Solutions - IDESS	Dell

Certifications

A full list of my certifications can be found on Accredible, Credly and MSLearn.

2024	Azure AI Engineer Associate	Microsoft
2024	Fabric Analytics Engineer Associate	Microsoft
2023	Databricks Certified Data Engineer Associate	Databricks
2023	Databricks Certified Data Analyst Associate	Databricks
2022	Azure Data Engineer Associate	Microsoft
2022	Power BI Data Analyst Associate	Microsoft
2022	All Microsoft Fundamentals Certifications, 8 exams	Microsoft
2021	Microsoft Office Specialist: Expert (Office 2019), 5 exams	Microsoft
2017	SAFe Agilist (SA)	Scaled Agile
2011	IASA Architect CITA-F (Foundation) Core	IASA Global
2006	CompTIA CTT+ Classroom Trainer	CompTIA
2006	Microsoft Certified Trainer (MCT)	Microsoft
2005	LPI (Linux Professional Institute) 101	LPI
2002	Server+	CompTIA
2002	Network+	CompTIA
2002	A+	CompTIA

Retired Certifications

2009	Microsoft Certified Technology Specialist	Microsoft
2009	Microsoft Certified Technology Specialist - Microsoft Forefront Client and Server, Configuration	Microsoft
2008	Microsoft Certified Systems Administrator (MCSA) Windows 2003	Microsoft
2007	Certified Wireless Network Administrator (CWNA)	CWNP
2006	Microsoft Certified Desktop Support Technician (MCDST) Windows XP	Microsoft
2005	Microsoft Certified Systems Engineer (MCSE) Windows 2000	Microsoft
2005	Microsoft Certified Database Administrator (MCDBA) SQL Server 2000	Microsoft
2003	Microsoft Certified Professional	Microsoft
2003	Microsoft Certified Systems Administrator (MCSA) Windows 2000	Microsoft
2003	Cisco Certified Network Associate (CCNA)	Cisco