Jacques Booysen

✉ booysenjacques@gmail.com 📱 +278-35085567 📍 Johannesburg, South Africa in Jacques Booysen ⬡ booyse 🌐 J.Booysen

PROFESSIONAL SUMMARY

Results-driven data professional with 20+ years of experience in Software Engineering and Data Science within the electricity industry. Expert in renewable energy projects, offering strategic insights and analytics to utility sector stakeholders using advanced machine learning, statistics, interactive data visualizations and software. Adept in crafting SaaS solutions with a focus on serverless computing, K8S, microservices architecture, and data orchestration. Dedicated to driving innovation and fostering team success through collaboration.

SKILLS & CERTIFICATIONS

Domain Knowledge: Electricity Industry, Timeseries Forecasting, Machine Learning, Renewable Energy, Statistics, Data Analytics, Interactive Visualizations, DevOps/MLOps, Agile Practices

Technical Skills: Expert in Python, R, Linux and R Shiny; Proficient with Docker, K8S; Argo CD, SQL, NoSQL, GIS, Tableau. Experienced in cloud computing, BigQuery, AWS, Azure DevOps, infrastructure as code, CI/CD, Apache Airflow. Competent with Git, VS Code, Jupyter

Language Skills: Fluent in English, Native in Afrikaans, Basic understanding of Dutch

Certifications: AWS Certified Solutions Architect, Associate (Amazon); Data Scientist's Toolbox (Coursera)

WORK EXPERIENCES

Senior Data Scientist – Enerweb, South Africa   Jan 2019 – Present

- **Renewable Energy Forecasting**: Developed day-ahead renewable energy forecasting models for utility-scale PV/Wind using global circulation models and gradient boosting machines.
- **Innovative Load Shedding Detection**: Developed the 'Load Shedding Watchdog' for Eskom Transmission TSO. Utilized Python/R Shiny to implement unsupervised learning algorithms for detecting load shedding events in municipal bulk load data.
- **Performance Monitoring Tool**: Created a performance monitoring tool for large customer load curtailment at Eskom. This tool supports embedded renewable energy and generates dynamic baselines using a sophisticated Prophet model on half-hourly data.
- **Leadership in Data Science**: Guided a team in developing key projects, including the optimization of wheeling allocation and conducting Monte Carlo simulations for virtual wheeling in both project and product development contexts.
- **Hydropower Modeling and Price Forecasting**: Conducted advanced modeling of hydro power using WEAP/R, simulating the Zambezi basin flow with ERA5 data. Also developed machine learning-based models for day-ahead hourly pool price forecasting.
- **Data Science Workflow Enhancement**: Engineered the 'WolfStack' analytics tech stack, leveraging Kubernetes and AWS to optimize data science workflows in production environments.

Data Scientist – Enerweb, South Africa   Jan 2010 – Dec 2018

- **Research and Development**: Contributed to Eskom research projects, focusing on areas like optimal maintenance, spatial temperature interpolation, energy density estimation, and market intelligence and DSM impact tools. Interactive tools for all.
- **Energy Modelling Initiatives**: Modelled and assisted in diverse energy modeling projects. Applied optimization techniques to analyze electricity market flows, linking Southern and East African power pools.
- **Full Stack Data Scientist**: Energy modelling with user interfaces for scenario analysis for significant projects like ANNA (Interconnector feasibility study) and CRIDF (climate-resilient infrastructure development simulations) climate change models.
- **Regional Load Forecasting**: Developed geo-based models for optimal regional load forecasting (GLF) paper, also domestic building density estimates leveraging clustering and spatial buffering algorithms to estimate domestic building density, further resulted in IEEE article findings

Software Engineer – Enerweb, South Africa   Jan 2002 – Dec 2009

- **Virtual Power Station Development**: Led the backend development of the Virtual Power Station (VPS), a SCADA-based solution for Eskom. Functioned as both Solutions Architect and Lead Developer, implementing the system using Linux, Perl, PLC programming, and custom Modbus/Spabus protocols in C.
- **Data Warehouse Solution**: Created a versatile data warehouse solution featuring a generic recursive roll-up engine, deployed across multiple Eskom departments.
- **Secure Infrastructure Solutions:** Designed and implemented secure infrastructure systems, including X.25, VPN, and customized firewall authentication, to enhance fault detection capabilities on Eskom's electricity grid.

EDUCATION

Bachelor of Science Honors (Computer Science & Applied Mathematics)   University of Johannesburg,1998-2001

PERSONAL INTERESTS

Adventure Racing; Ultra Trail Running; Mountain Sports

REFERENCES

Schalk Heunis - Managing Executive: Big Data, AI & RPA (Vodacom)
schalk.heunis@vodacom.co.za
PhD (Stell) on Probabilistic Methods Applied to Power Systems

Marcus Dekenah – Load Research Specialist (MD Consulting)
marcus@mdekenah.co.za
NHD Elec. Eng (HC), M Dip. Tech. Eng, Bsc Elec Eng (Cum Laude), MBLII

COURSES AND PROFESSIONAL MEMBERSHIPS

Water Evaluation and Planning System (WEAP) Intro, SEI — Stockholm Environment Institute
October 2020 — October 2020

The South African Institute of Electrical Engineers, Johannesburg
April 2019 — Present
Member and presenter at the load research chapter,

Presenter at SatRday 2018: Practical applications using R for spatial data visualization, creation, and manipulation.
Member and presenter at the R User Group Johannesburg 2023, Ploomber: Elevating Your R-Based Data Science Workflow

Key:	AWS/Cloud	R / Machine Learning	Python	ML/Data/Dev-Ops
Years Experience	5 years +	13 years	5 years +	2 years
Notes	Training at AWS Popup Loft Johannesburg. Been using AWS at various levels, for HPC for the last 5 years+. I am an AWS certified solutions architect since 2022, use AWS for our analytics production pipelines.	Implement a large quantity and range of operational research projects using statistics machine learning in R for Eskom Research. Bottom-up models, exploratory work, clustering, MLR, Random Forests, Decision Trees, Simulated annealing, gradient boosting machines.	Time series forecasting, anomaly detection, clustering	Apache Airflow Ploomber Azure (DevOps) Argo CD K8S
Some Tools/Libraries	Google BigQuery, K8S, Airflow, AWS Lambda, SQS, Textract, RDS, EC2, VPC, CloudFormation, MWAA, S3, EFS, AIM, etc.	dplyr, raster, ggplot2, shiny, rhino, lm, random forests, rgdal, clustering, elm, etc.	prophet, keras, torch, pvlib, pandas, numpy, scikit-learn, pylint, pytest, boto3, re, xgboost, papermill statsmodels	airflow, papermill, ploomber