Jacques Booysen, Senior Data Scientist

City of Johannesburg, South Africa, +27835085567, booysenjacques@gmail.com

---

LINKS

Linkedin, Portfolio, GitHub

PROFILE

I am a scientist at heart, passionate about data science, problem solving and software engineering. I am responsible for helping utilities and other companies interpret and manage data and solve complex problems using expertise in a variety of data niches often employing computer science, server-less cloud computing , machine learning, optimisation, modelling, statistics and advanced visualisations.

I have been involved in the development of Geo-based load sub classes , a market intelligence system with sentiment analysis using machine learning and integrated resource least-cost planning for energy trading market analysis on the Southern African Power Pool, including the impact of climate change models on this. I have produced day ahead hourly renewable forecasting models for IPP's. Solutions were developed using R, Python, specialised software and tableau.

Currently I am the technical lead in the disruptive innovation business unit at enerweb, I provide Data Science as a service on the Python based in house SaaS technology stack enerflow. We use AWS micro services architecture and a custom framework i developed that runs on top of Apache Airflow making use of AWS Lambda and allowing you to run Jupiter notebooks in the cloud from MS Excel. My strength lies in the ability to do rapid prototyping because of my strong software engineering background, and my passion for machine learning and constantly improving doing things a better way.

EMPLOYMENT HISTORY

Jan 2020 — Present

Senior Data Scientist, Enerweb
South Africa

Research and Development of forecasting models and visualisations in Python/R for the day ahead power pool electricity markets in Southern Africa - Renewable energy forecasting models and interactive tools for forecasting Solar PV, Wind and Hydro power plant outputs hourly using global circulation models and historic satellite data.

Developed AWS micro services python framework for Data Science stack, CI/CD using Amazon Apache Airflow (MWAA), Lambda, S3, SQS, TexTract (OCR), Cloud-formation (AWS SAM) for infrastructure as code. Developed Data extraction and anomaly detection tool set for energy utility using k-means clustering and statistics.

Skills: Apache Airflow · Ploomber · Python · R · AWS · Jupiter

Jan 2012 — Present

Data Scientist, Enerweb

- I developed an Market intelligence system in R Shiny for network planning support and sentiment analysis using MongoDB and machine learning. Used full text and spatial search functionality. - Optimal Regional Geo Based Load Forecast Subclass development finding sweet spot between error and model complexity. - Spatial Interpolation of Temperature using digital elevation model and comparing and evaluating various modelling techniques. - Developed spatial buffering algorithm to calculate domestic building density estimates per spatial LSM for Eskom Geo Based Load Forecasting (Network Planning) - SAPP-EAPP: Electricity Market Model in MESSAGE used to estimate least cost energy flows within and between the interconnected SAPP-EAPP (Southern/Eastern African) power pools, given a range of scenario assumptions. - Energy modelling of ANNA 400kV interconnector between Angola and Namibia (Interconnector Impact Study) by Utilisation of a least-cost energy flow Electricity Market Model in MESSAGE, to estimate the energy flows and changes in energy flows within and between the interconnected SAPP power pool countries, modelling probable variant scenarios. (Tableau Front-end) - R Shiny App and assisted in Monte Carlo Simulations using power pool least-cost energy MESSAGE model for the climate resilient infrastructure Development Facility (CRIDF) via a rationalised long term energy planning and scenario tool and climate change models to determine climatic and economic impacts in the region on the electricity grid in the Southern African power pool over the coming 30 years (2015-2045)

Skills: MongoDB · MESSAGE · Energy Modeling · Optimization · R Shiny · Tableau · Python (Programming Language) · R (Programming Language) · SQL

Jan 2002 — Dec 2009

Software Engineer, Enerweb

Started as a Linux programmer and Back-end Specialist, doing various research projects in C and Perl.

I helped pioneer the Virtual Power Station SCADA System over Thin networks / GPRS and corporate APN's, including Modbus (PLC) and SpaBus (Frequency Relay), protocol implementation in objective C. I also developed an asterisk IVR voice system running in an J2EE server. After the VPS proof of concept a new dev team wrote a C# product named DR Pro after the initial work that i did.

I worked on the Eskom Data Hub and was responsible for developing a generic Rollup Engine (Java Recursive Oracle Stored Proc) for a custom OLAP Data Warehousing solution able to have a single value of the truth.

Wrote a X.25 Protocol gateway using Linux C/Visual Basic and Open VPN, assisting Eskom Engineers to connect to substation equipment from the Eskom WAN.

I also wrote an authenticating firewall custom system for Eskom Telecoms, enabling control and secure access to the operational telecoms network.

Skills: Java · Perl · C (Programming Language) · J2EE Application Development · PHP · PL/SQL · Oracle Database · Modbus · SCADA · Programmable Logic Controller (PLC) · Linux

---

EDUCATION

Jan 1998 — Dec 2001 **B.Sc (Hons) Computer Science, University of Johannesburg**

Jan 1998 — Dec 2001 **Bachelor of Science (BSc) Applied Mathematics, University of Johannesburg**

LANGUAGES

Afrikaans	Native speaker	**English**	Highly proficient

PUBLICATIONS

Apr 2014 **Domestic building density estimates for network planning**

Domestic Use of Energy Conference (DUE) - see LinkedIn

Aug 2013 **Regional electricity load profile subclasses for network planning**

Industrial and Commercial Use of Energy Conference (ICUE) - see LinkedIn

---

CERTIFICATIONS

Aug 2022 — Aug 2025 **AWS Certified Solutions Architect – Associate**

Feb 2015 — Present **R Programming**

Coursera Verified Certificates X9HT2EW4WW www.coursera.org

Dec 2014 — Present **The Data Scientist's Toolbox**

Coursera Verified Certificates 2NZPCAYE53 www.coursera.org

SKILLS

Critical thinking and problem solving	Ability to Work Under Pressure
R Programming	Self Starter
Python Programming	Rapid Prototyping
Geospatial Data	

HOBBIES

Adventure Racing,
Ultra Trail Running,
Mountain Sports

COURSES

Oct 2020 — Oct 2020 **Water Evaluation And Planning System (WEAP) Intro, SEI — Stockholm Environment Institute**

REFERENCES

Marcus Dekenah from MD Consulting

marcus@mdekenah.co.za

Schalk Heunis from Vodacom

schalk.heunis@gmail.com

PROFESSIONAL MEMBERSHIPS

Apr 2019 — Present   **The South African Institute of Electrical Engineers**   Johannesburg

Member and presenter at the load research chapter, Presenter at SatRday 2018: Practical applications using R for spatial data visualisation,creation and manipulation