Wycliffe Mwebi

SENIOR DATA SCIENTIST

+254712514020 | linkedin.com/in/wycliffe-mwebi-aa946714/

Wikie7@gmail.com | github.com/Wycliffe-Mwebi

Nairobi-Kenya | wycliffe-mwebi.github.io/portfolio/

1. PROFILE SUMMARY

Senior Data Scientist with 10+ years of broad based experience in building data intensive applications, overcoming complex architecture, and scalability issues in diverse industries. Proficient in Predictive modeling, data processing and data mining algorithms, as well as scripting languages, including Python and Java. Capable of creating, developing, testing and deploying highly adaptive diverse services to translate business and functional specifications into substantial deliverables.

2. TECHNICAL SKILLS/TOOLS

MACHINE LEARNING & STATISTICAL TECHNIQUES *Generalized Linear Models, Tree Based Models, SVM, Naive Bayes, Neuro Networks, Clustering, Feature Engineering, NLP, Hypothesis & A/B Testing, Experimentation*	**DATA WAREHOUSING AND ETL TOOLS** *OLAP, OLTP, MS SQL (SSAS, SSIS, SSRS).* *ETL tools (DataStage, Informatica, Pentaho, Talend, Tableau Prep Builder)*
DATA ENGINEERING FRAMEWORKS *Apache Hadoop, Spark, Kafka, Cassandra, Hive*	**WORKFLOW AUTOMATION AND SCHEDULING** *AirFlow, Luigi, NiFi, Jenkins*
DATA EXTRACTION/CLEANING/ VISUALIZATION *MySQL, Spark SQL, Tableau, Python (matplotlib, seaborn), R (ggplot), MS Excel, Geospatial (QGIS, ArcGIS)*	**CLOUD WAREHOUSING** *Amazon Redshift, Azure Synapse Analytics, Google BigQuery, Azure SQL Database.*
PROGRAMMING *Python, Scala, R, Java, C++*	**DATABASES: (SQL and NoSQL)** *MySQL, SQL Server, PostgreSQL, MongoDB, NoSQL.*

3. WORK HISTORY

Aug 2016- to date	**Senior Data Scientist**	**Atos (www.Atos.net)**

Responsibilities

- Design and implement adaptable, scalable, and reliable ETL data pipelines and algorithms to process unstructured and structured data from wide variety of data sources using Amazon cloud technologies.
- Day-to-day activities include (1) Developing an integration process to move and transform data, (2) Talking directly with stakeholders to better understand their requirements, (3) identification of strategic data insights and growth opportunities, (4) Test and validate work that has been completed.
- Managing and Organizing data - develop algorithms to help make raw data more useful to the enterprise, keep an eye out for trends or inconsistencies that will impact business goals, communicate data trends to others in the organization and to help the business make use of the data it collects.

- Data Loads - one-time loads for adhoc projects, or recurring extractions for regular reports. SQL, CRON jobs, ETL tools to take data from somewhere or some other system and land them onto the sandbox for the data analysis.

- Data Preparation - Side-by-side wrangling, coding, labeling, creating synthetic data, creating metadata, to transform data into a usable form. - SQL, Python.

- Data Pipelines / Exposing Data - creating dimensional tables to support BI cubes, reports, or analytical base tables for machine learning. Creating database views or regular materialization to update data that is used by analysis. - SQL.

- Optimizing Queries - monitor all data jobs hitting the servers and data loads. Kill long-running queries or jobs that balloon data volume. Find ways to shorten regular loads and transformations for efficiency. Linux, Ganglia, SQL.

- Automating and Publishing Analysis - take results of adhoc analysis and expose them as simple web apps and visualizations for easier interaction with clients. Take prototype scripts of machine learning and analytics jobs and productionize them as in-database processes for regular runs. SQL, Python, D3.

- Own all phases of the data Engineering lifecycle (architecting data platforms, designing data stores, and gathering, importing, wrangling, querying, and analyzing data. Performance monitoring and fine tuning to ensure systems are performing at optimal levels.)

Jan 2015-Jul 2016	**Business Intelligence Specialist**	**SICPA -** (www.sicpa.com)

Responsibilities

- Design of the Microsoft SQL data warehouse, including all information structures (staging area, data warehouse, data marts, and operational data stores).

- Design, develop, test, monitor, manage, and validate data warehouse activity, including data extraction, transformation, movement, loading, cleansing and updating processes.

- Work with end users to translate business questions and requirements into applications that employ the appropriate reporting tools.

- Own end-to-end reporting and MIS delivery for multiple processes. Responsible for ensuring the timeliness and accuracy of the all reports.

- Create dashboards and visual interactive reports using Tableau.

- Translate business propositions into quantitative queries and collect/clean the necessary data.

- BI tools and BI systems, Tableau, SAP, creating data-rich dashboards, implementing Row-level Security (RLS) in Tableau, writing DAX expressions, developing custom BI products with scripting and programming languages. - R, Python

2009-2014	**Project Officer -Information Management**	**International Monetary Fund (IMF),** Washington DC, (*www.imf.org*)

Responsibilities

- Perform data analysis, data profiling, on demand reporting, presentation of data for senior stakeholders, user training and mentoring team and department members.
- Developing Data Warehousing solutions in SQL Server, Designing ETL solutions using SSIS and OLAP cubes using SSAS (tabular and multi-dimensional).
- Participate in all aspects of the BI life cycle, including analysis, design, development, testing, production deployment and support, including creating formal written deliverables and other documentation.
- Ensure that the data loaded from the ERP systems into the data warehouse is accurate and meets the informational requirements of the organization.
- Creating and maintaining reports using SSRS; deploying the BI reports to SharePoint and managing BI report change requests.
- Interpreting and designing bespoke reporting specifications from end-users (Finance, IT, PM, Operations, clients and Third Party Suppliers).
- Deriving insights from non-structured non-relational data, and providing business value from content through improved interactive reporting dashboards.
- Work with other BI team members to analyze and recommend BI solutions that meet business requirements that are aligned with industry trends.

4. EDUCATION

Master degree Innovation, Technology & Law	*-University Of Edinburgh, UK* ***-2010***
Bachelor of Science (BSc Hons 2:1), Computer Studies	*-University of Sunderland, UK* ***-2006***
Higher Diploma, Management of Information Systems (MIS)	***-****Institute for the Management of Information Systems, UK* ***-2002***

5. REFERENCES

- **Mr. Fredrick Mwangi**
  Projects Manager, Atos
  Tel; +254 101 514020
  freduke70**@gmail.com**
- **John Kerina**
  Senior Lecturer &Head of Department, JKUAT University.
  Tel; +254 722 633980
  jmkerina@gmail.com