

JOSEPH MWANGI

DATA ANALYST



+254723500219



github.com/Josetedd



josetedd@gmail.com



JOSEPH MWANGI

DATA ANALYST

ABOUT ME

I am a data Analyst and IT professional with over 8 years of Experience in Systems Development, IT Support, and 5 years of experience in Developing tools for Analyzing and interpreting data for data driven solutions. I have an excellent understanding of data management & analytics tools including Python, MS Excel, Power Bi, Tableau, SAS & SQL.

SKILLS



Data Analysis and visualization
(Python, Excel, Power BI, Tableau,
Google data studio & SAP Lumira)



Cloud Administration
(Azure administration)



**Database Design, Development
and Administration**
(Ms. Access, SQL, MySQL)





4Life + UNHCR Uganda Emergency response to deforestation (2021)

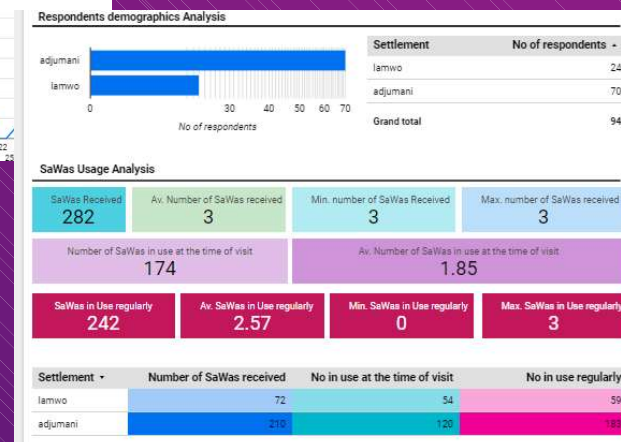
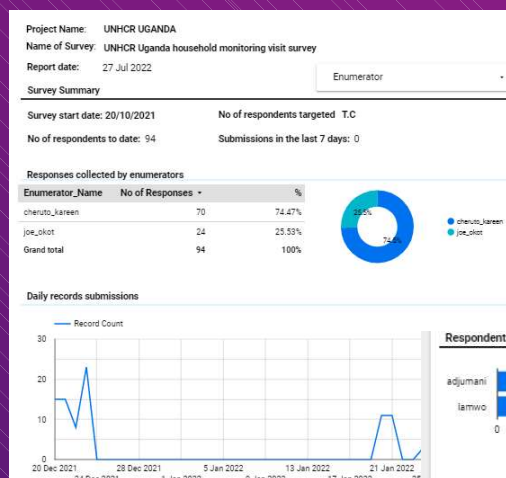
- ✓ Developing Data collection tools using Kobo Collect
- ✓ Training and Supporting data collection
- ✓ Performing data analysis and reporting using python and Google data studio

<https://4lifesolutions.com/unhcr-uganda/>

Project contact persons

Ms. Joan Kitika
Programs director-Africa
jk@4lifesolutions.com

Ms. Barbra Uwera
Projects Manager- Uganda
bu@4lifesolutions.com





4Life + Pharm Access Tanzania MomCare project (2020- ongoing)

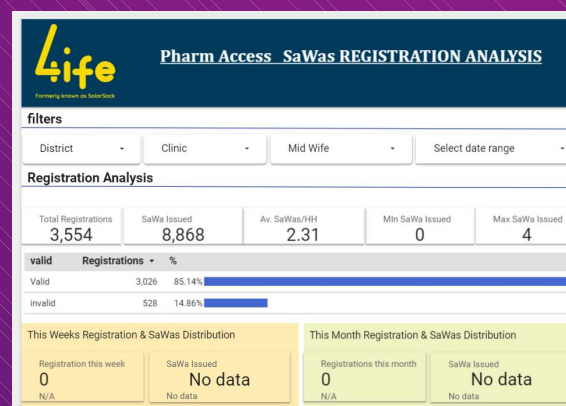
- ✓ Developing Data collection tools using Kobo Collect
- ✓ Training and Supporting data collection
- ✓ Data cleaning & error correction
- ✓ Performing data analysis and reporting using python and Google data studio

<https://4lifesolutions.com/momcare-with-pharmaccess/>

Project contact persons

Ms. Joan Kitika
Programs director-Africa
jk@4lifesolutions.com

Ms. Happiness Ursa
Projects Manager- Tanzania
hu@4lifesolutions.com



```
import warnings
warnings.filterwarnings('ignore')

locate_errors = (all_data["serial1_check"] != "Valid") | (all_data["serial2_check"] != "Valid") | (all_data["serial3_check"] != "Valid")
errors_df = all_data[locate_errors]
available_serial_n = [int(45000 + x) for x in range(0, 5000)]
all_clinics = errors_df["Health_Clinic"].unique()
counter = 0
for clinic in all_clinics:
    errors_df_c = errors_df[errors_df["Health_Clinic"] == clinic, :]
    s1, s2, s3 = [errors_df_c[errors_df_c[x] != "Valid"].index.to_list() for x in ["serial1_check", "serial2_check", "serial3_check"]]
    s1, counter = get_serial_number(s1, counter)
    s2, counter = get_serial_number(s2, counter)
    s3, counter = get_serial_number(s3, counter)

series = [s1, s2, s3]
for (i, s) in zip(range(1, 4), series):
    all_data["Old_serial_no" + str(i)] = \
        replace_serial_number(s, all_data["Old_serial_no" + str(i)], all_data["serial_no" + str(i)])
```



4Life + CRS Kenya & South Sudan (2020- ongoing)

- ✓ Developing Data collection tools using Kobo Collect
- ✓ Training and Supporting data collection
- ✓ Data cleaning & error correction
- ✓ Performing data analysis and reporting using python and Google data studio

<https://4lifesolutions.com/crskkenya/>

Project contact persons

Ms. Joan Kitika
Programs director-Africa
jk@4lifesolutions.com

Mr. Riek Lam
Projects Manager- South Sudan
gl@4lifesolutions.com



```
def serial_check(sn, number):  
    #serial check  
    if pd.isnull(sn):  
        #check= 'Serial number missing'  
        checks='Valid'  
    else:  
        if serials_list.count(sn) > 1:  
            check = 'serial' + str(number) + ' duplicated'  
        else:  
            checks='Valid'  
    return check  
all_data["serial1_check"] = all_data["Serial_number_of_SaWa_1(TYPE_the_LAST_6_DIGITS_only)"].apply(lambda x: serial_check(x,1))  
all_data["serial2_check"] = all_data["Serial_number_of_SaWa_2(TYPE_the_LAST_6_DIGITS_only)"].apply(lambda x: serial_check(x,2))  
all_data["serial3_check"] = all_data["Serial_number_of_SaWa_3(TYPE_the_LAST_6_DIGITS_only)"].apply(lambda x: serial_check(x,3))  
all_data["serial4_check"] = all_data["Serial_number_of_SaWa_4(TYPE_the_LAST_6_DIGITS_only)"].apply(lambda x: serial_check(x,4))  
def valid_check(serial1_check,serial2_check,serial3_check,serial4_check):  
    if serial1_check != 'Valid' or serial2_check != 'Valid' or serial3_check != 'Valid' or serial4_check != 'Valid':  
        return 'Invalid'  
    else:  
        return 'Valid'  
all_data['valid'] = all_data.apply(lambda x: valid_check(x['serial1_check'],  
                                                         x['serial2_check'],  
                                                         x['serial3_check'],  
                                                         x['serial4_check']),  
                                  axis=1)
```



Save The Children Somalia Short term regional emergency response project 2018

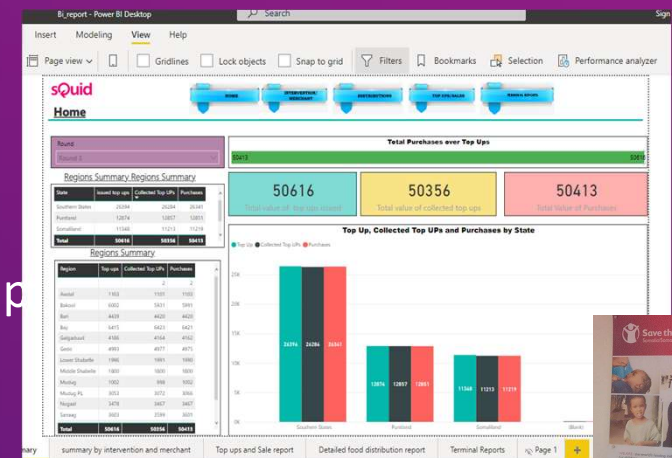
- ✓ Training and providing technical support to Save the Children Somalia staff and merchants.
- ✓ Application Testing, System deployment
- ✓ Worked with the development team to build Android data collection tool for the project
- ✓ Designed and implemented from the ground up data visualization tool using Power Bi

<https://squidcard.lpages.co/save-the-children/>

Project contact persons

Mr. Simon Kaniu
General Manager-sQuid
skaniu@squidcard.com
+254721586966

Mr. David Loki
Save the children international
David.loki@savethechildren.org
+254733753337





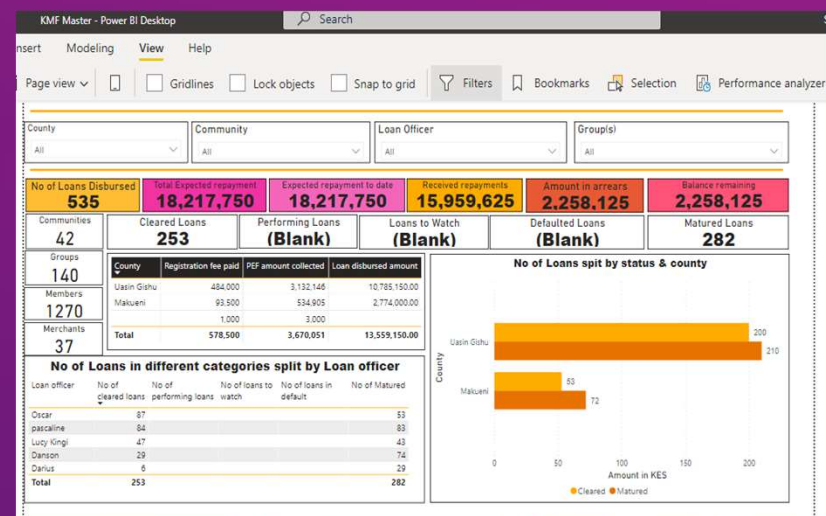
iMlango Project (DFID Funded)-Kenya (December 2016 to 2020)

- ✓ Providing Technical Support to field officers on the ground,
- ✓ Developing Data collection tools using data collection tools such as ODK and Kobo Collect
- ✓ Repairing Tablets and computers. cleaning and analyzing data collected from the field

<https://www.imlango.com/>

Project contact persons

Mr. Simon Kaniu
General Manager-sQuid
skaniu@squidcard.com
+254721586966





Tear Fund Cash Transfer Project- South Sudan (Oct 2018 to Jan 2019)

- ✓ Training and providing technical support to Tear Fund staff and merchants.
- ✓ Application Testing, System deployment
- ✓ Monitoring the projects progress
- ✓ Designed and implemented from the ground up data visualization tool using Power Bi



Project contact persons

Mr. Simon Kaniu
skaniu@squidcard.com
+254721586966

Ms Emiliane
Project manager – Tear Fund SS
cyuzemlly@gmail.com
+250783648039

