Submission deadline: Sunday, 22 August 2021
Lead Instructor: Sabrina

Operation Orphan: Rubbish Science

Team: Chloé Koura, Irina Mateescu, Mayumi (Mayu) Hamaoka, Patricia Pedro, Oludare (Dáre) Soniran

INTRODUCTION

Operation Orphan is a charity that aims to increase scientific literacy by helping children develop problem-solving strategies using freely available resources. Their project, Rubbish Science, positively impacts underserved communities by increasing access to education, improving community health, raising environmental awareness and providing income for households.

Operation Orphan is in the process of writing a funding proposal to for a grant of £250,000 and the team is consulting on expanding Rubbish Science to vulnerable communities in Kenya in areas such as Mombasa. We analysed education and health data from the World Health Organisation, openAfrica, Kenya Open Data and the World Bank and produced visualisations presenting the level of education disaggregated by gender and the prevalence of malaria and cholera.

BACKGROUND

Our initial aim was to complete a comparison between the countries Operation Orphan currently operates in (Vietnam and Sierra Leone) with Kenya. However, due to data scarcity we thought it best to focus our efforts on a needs assessment for Kenya. We outlined multiple areas that could be used to measure the potential impact of Rubbish Science:

- Getting children back into education
- Self-built health resources
- Income-generation for households
- Knowledge-sharing for community wellness
- Environmental awareness

We chose to focus on increasing educational opportunities for children. A key area of Rubbish Science is teaching children to upcycle the resources around them into useful items, which in the process inspires and empowers them. These self-built items include mosquito prevention and clean water processes. Additionally, the funding will be used to provide scholarships up to university for participants with high potential.

To investigate the usefulness of expanding the Rubbish Science project to Kenya, we decided to focus on issues that may be highly impacted by the expansion:

1. What are the waste management statistics?
2. What are the education continuity statistics?
3. Does malaria affect different Kenyan counties at different rates? Are bed net usage and health spending correlated to malaria incidence?
4. What are the rates of cholera mortality and clean water availability per area?

The answers to these questions explain the need for systemic change now, as well as scalability within the near future. Throughout our analysis, we have aimed to disaggregate our data on different components such as gender (male/female), areas (county, or rural/urban) and educational level (primary, secondary...) depending on the nature of the analysis. These identifications of the target population are aimed at focusing resources and efforts where they are most needed.

STEPS SPECIFICATIONS

The project was initially framed as a way to make a decision on whether it was feasible to expand the Rubbish Science project into Mombasa, Kenya. Due to limited information on key indicators for Mombasa, the project was converted into an assessment of feasibility for Kenya as a whole. More granular data was used whenever available.
Many data sources were identified as potential candidates for data utilisation including World Data Bank, openAfrica, Kenya Open Data and the World Health Organisation Database via API. Each of these data sources had different pre-processing requirements which span from elimination of missing data, formatting data into different data types, re-ordering, and re-naming of columns. These data were fed into statistical methods

including calculation of mean and correlations and used to produce visualisations including line/bar charts, scatter plots and choropleth maps.

IMPLEMENTATION AND EXECUTION

- Development approach and team member roles

We followed the data analysis lifecycle framework, beginning with planning of our project. Several data sources were identified and data profiled; the team had discussions regarding the content, interrelationships and potential to present effective analysis that supported the subject matter. The team engaged in regular check-ins to develop the focus. Each team member gravitated towards their topic of interest. Chloé worked on qualitative research regarding Operation Orphan and recyclable material. Mayu and Patricia pulled health data from WHO and Kenya Open Data, processing and transformation and creating visualisations. Irina pulled education data from Kenya Open Data and Open Data Africa, cleaned the data and performed exploratory analysis. As Dáre joined the group in week 2, she took the responsibility of preparing the report. As a group, we interpreted the results and prepared the documentation.

- Tools and Libraries

Scientific computing libraries	● Pandas	● JSON/DataFrame/CSV applications
	● NumPy	● Statistical analysis
		● Arrays
	● Scikit-learn	● Preprocessing
Visualisations libraires	● Plotly	● Extension to other visualisation tools
	● Matplotlib	● Saving figures ● Bar and line charts
	● Seaborn	● Correlation heatmaps
	● Geopandas	● Choropleth maps
Modules	● requests	● API GET requests
	● json	● json object manipulation
	● pprint	● improved readability
	● datetime	● date class manipulation
	● string	● string manipulation

- Implementation process

All individual activities were uploaded onto GitHub (coding and documents) or Google Drive (homework 2 and draft of the project report). Communication was streamlined on Slack with biweekly meetings. We applied agile development methodology by implementing a weekly sprint with stand-ups at each meeting, sharing progress and limitations. Minutes were written allowing for an easy introduction for Dáre in week 2. Code was uploaded onto GitHub the night before each meeting, allowing for a code review before each stand-up.

Achievement - We were successful with teething our ideas and utilising everyone's skillset in areas that they were comfortable with. We worked well as a team, communicating effectively, sharing resources and knowledge.

Challenges - The main challenge was having broad ideas and attempting to prepare a project brief with specific hypotheses. This was overcome with regular meetings reviewing progress, allowing the team to be more realistic with expectations. The decision for Dáre to change groups was her desire to work on a topic with a more generalised understanding and feel possible to contribute more. The team welcomed her, thanks to the excellent minutes, she was able to come up to speed.

Throughout the project some data limitations were encountered, specifically lack of data granularity, scarcity of data on Kenya, missing data for sequential and non-sequential years.

Decision to change something - We had initially aimed to complete a comparison with the education and health data from countries where the project has been implemented, Vietnam and Sierra Leone. However, the data was either not reported or not openly available. Additionally, as Operation Orphan is a small charity it was difficult to distinguish their impact from the impact of large organisations operating in the same area such as the WHO. Hence, we decided to focus on Kenya exclusively.

SWOT ANALYSIS	
[S] Strengths **Team work:** Effective communication, good distribution of responsibilities, passion for the topic. **Data sources:** Multiple resources identified and used	**[W] Weaknesses** **Topic:** Broadness of ideas **Data:** Data security, outdated data and lack of data continuity.
[O] Opportunities **General python knowledge:** Growing knowledge with opportunity to apply skills, good use of multiple libraries and modules **Documentation:** Learning to read documentation	**[T] Threats** **Time constraints:** Lots to accomplish with limited time **General python knowledge:** Consolidation of continual learning

Table 2: SWOT analysis of the group project.

RESULT REPORTING

1. What are the waste management statistics in Kenya?

Mombasa generates approximately 2,000 tons of waste and only 65% is collected and 13% of households can access waste management services. It was found that household size has a negative correlation with proximity to the collection points, waste management expenditure, waste management behaviour (attitude) and exposure to (communicable) diseases. Thus, these relationship(s) depict that the larger the household size or the nearer the household is to the primary collection points, the more is spent on waste management, the higher the tendency to haphazardly dispose of waste and the higher the risk of exposure to communicable diseases.

The Kenyan government's Vision 2030 aims to have "fully functional and compliant waste management systems, by developing strategies towards achieving sustainable waste" and is underway in five cities including Mombasa. As Rubbish Science aims to solve local problems whilst recycling rubbish and increasing environmental awareness, this initiative would be welcomed.

2. What is the education continuity statistics?

From Kenya Open Data, Figure 1 presents school enrolment per education level and gender. Due to the lack of data continuity, 2016 was the most recent data available with a good level of detail.

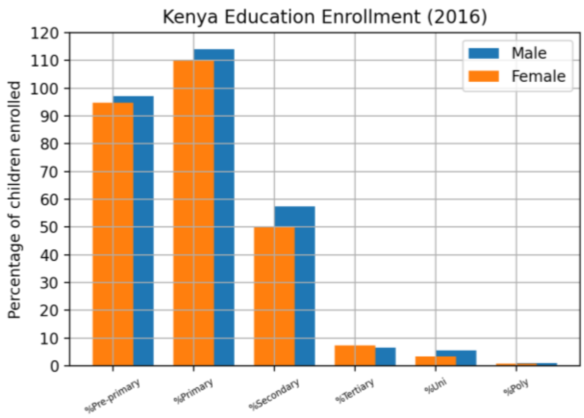

Figure 1: Education enrolment disaggregated by gender (2016)

The graph shows a sudden drop in the percentage of enrolment from primary to secondary school, indicating that 50% of children are not continuing their education. This could be due to primary school being the last level of free education. What is more, the secondary school fees (cheapest ~$400 per year) are well above what someone living under the poverty line can afford (35% of the Kenyan population lives on less than $1.90 per day). The school fees could also explain why primary education is above 100%, as one can assume that primary graduates are re-enrolled to keep them in education even if parents/caregivers cannot afford the transition to secondary education. The gender difference varies up to 5% meaning all children regardless of gender have their education impacted by financial hardship.

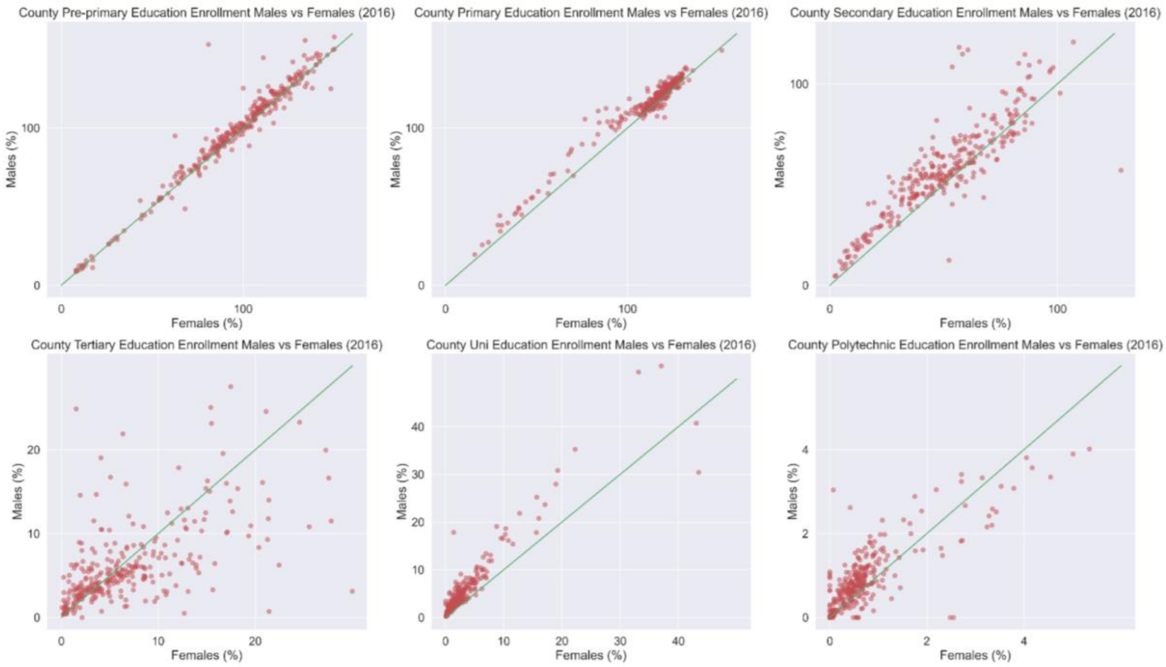

Figure 2: Subplots of gender distribution of education enrolment by county count (2016)

Figure 2 presents subplots exploring the gender distribution of enrolment at each education level per county. The data shows a skewness towards males in all areas except tertiary education. At the university enrolment, the difference is pronounced even though the overall enrolment percentage is small. The 2019 consensus revealed that only 3.5% of Kenya's population had a university degree with a surprising 59.4% being male. Overall, this presents the need for education. Rubbish Science will be welcomed especially by those without financial means and the scholarships will make a difference. However, Operation Orphan may not need to change their approach to focus on a particular gender as the needs assessment presents that all children require support alike.

Figure 3 is a set of choropleth maps of secondary school enrolment and population per county. This presents Nairobi and Nyeri as the counties with the highest rates of enrolment. Unsurprising, Nairobi, Kenya's capital, has the highest population. However Nyeri has a lower population but high rates of secondary school education. This could possibly be due to the income levels of the county population. Another possibility is the activities of the local government. Nyeri county has implemented a lot of educational initiatives. Operation Orphan could work in partnership with Nyeri county to understand how they increased their rates of enrolment.

Operation Orphan could also decide to focus on the counties in the west, southeast and northeast, which have a high population (between 60,000 and 120,000) and levels of secondary education of only 20-40%.

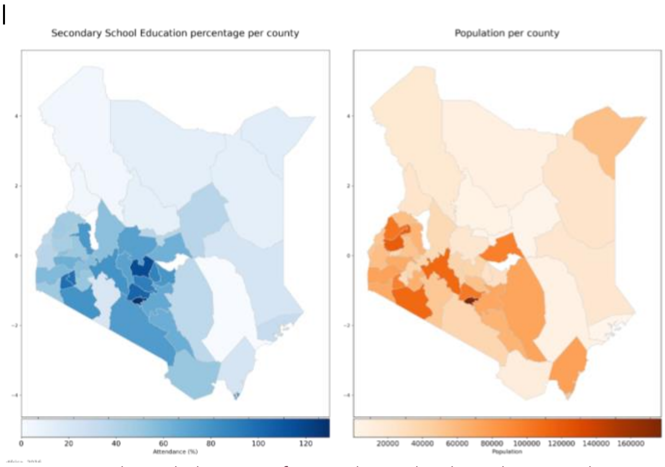

Figure 3: Choropleth maps of secondary school enrolment and population per county (2016)

3. **Does malaria affect different Kenyan counties at different rates? Are bed net usage and health spending correlated to malaria incidence?**

The analysis for the third question was based on data collected in 2016 and available through the Open Kenya API. For the aim of this project, the indicators observed are:

- % People Sleeping Under a Bed Net: The higher the percentage, the better is the county performance
- % People That Had a Fever or Malaria: The lower the percentage, the better is the county performance
- Health Spending Per Person: The higher the spend, the better is the county performance

Figure 4 presents the statistical analysis of the indicators. An outlier showed that the minimum percentage of population with a fever or malaria was 0% for one county. There is an inference of incorrect reporting and the dataset could be considered questionable. Another assumption is that there was no entry for that county and the data was defaulted to zero. The max percentage for each indicator is 62.4%, 63.3% and 50.0% respectively; this presents that there is still a need for health support.

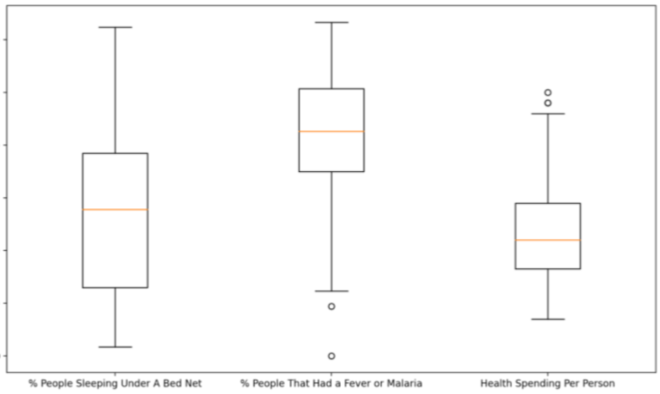

Figure 4: Box plot of three health indicators (% people sleeping under a bed net; % people that had a fever or malaria; and health spend per person) (2016)

Figure 5: Graph 1 presents the counties to use as benchmarks for the indicators. These counties were identified by their high percentage of people sleeping under a bed net in comparison to the Kenyan average (26.73%). One assumption is that there is a high encouragement of the use of bed nets as a method of protection from mosquitoes. However, the rates of those that had a fever or malaria is still relatively close to the mean (40.18%) for all counties excluding Kajiado. With additional time, we would consider investigating Kajiado to better understand their success. Figure 2 shows that these eight counties vary by their population size. Along with high temperatures, mosquitoes thrive in lower elevations and wet places. So another variable is the environment of the counties; if the topography is higher or they are further away from wet places.

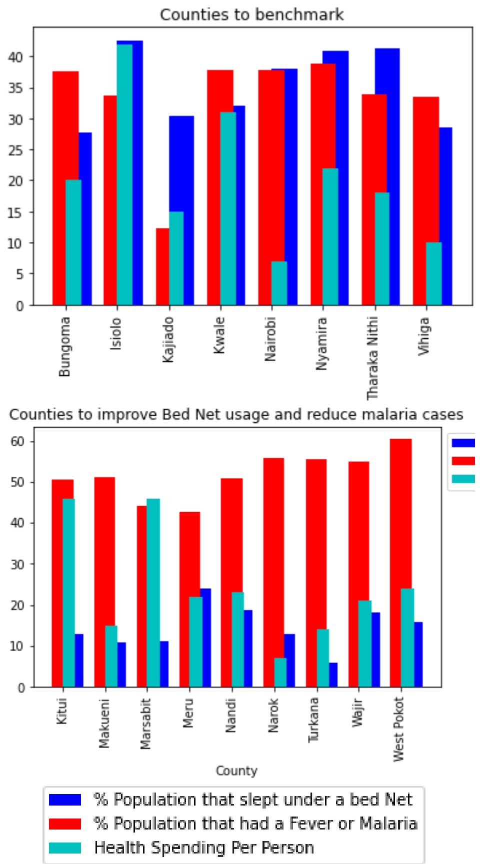

Figure 5: Graph 1: Counties to benchmark based on their high percentage in people sleeping under a bed net in comparison to the Kenyan average (26.73%) (2016). Graph 2: Counties with the lowest bed net usage in comparison to the Kenyan average (26.73%) (2016)

The Figure 5: Graph 2 shows the nine counties with the highest need to change social behaviours to reduce the high prevalence of fever or malaria. They have a lower percentage of people sleeping under a bed net than the country mean (26.73%) and a higher percentage of people with fever or malaria than the country mean (40.18%).

A surprising finding was that six out of the seven benchmark counties still had high rates of fever or malaria cases even though their rates of bed nets use were also high. As Operation Orphan teaches children to make bed nets from waste material, our expectation was that the higher the use of bed nets, the lower the rates of fever or malaria.

We created a scenario to analyse the two of the indicators per county. We measured the following:

1. The percentage of people that sleep under a bed net per county is higher than the median (27.8%).
2. The percentage of people sleeping under a bed net is higher than the percentage of people that had fever or malaria.

Figure 6 presents the two outcomes from these measurements. Expected behaviours can be seen from Isiolo, Kajiado, Nairobi, Nyamira and Tharaka Nithi. These five counties met both of the defined measurements for effective health incidence, suggesting that these counties have higher social awareness of mosquito-borne disease prevention. The health spend per person varies between 5-45%, this is likely due to the population differences per county. Thus cannot be used as a successful indicator for the social awareness of mosquito-borne disease prevention.

Unexpected behaviours can be seen from Kisumu and Mombasa; both counties failed on both defined measurements. Operation Orphan is interested in implementing Rubbish Science in Mombasa. With further in-county research, it would be interesting to understand why there is a high percentage of people that sleep under a bed net, yet the county still has high rates of fever and malaria. The health spend per person is also an interesting indicator as it is at the same level as Nairobi.

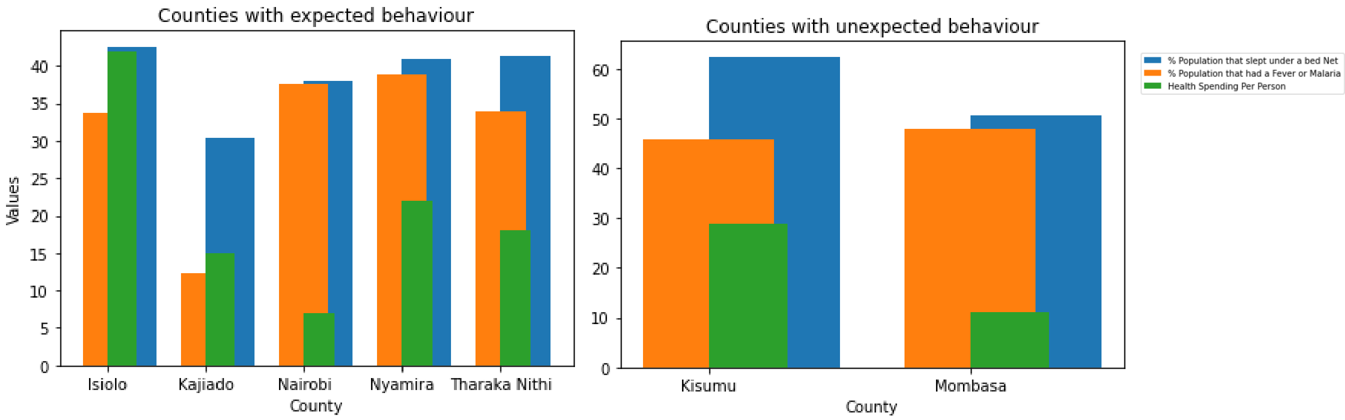

Figure 6: Exploratory analysis of the counties with expected and unexpected behaviours: (1) the % of people that sleep under a bed net is higher than the median (27.8%) and (2) the % of people that had a fever or malaria

To confirm that our assumption regarding correlation is accurate, Figure 7 presents a heatmap of the correlation of the three indicators. As believed, health spend has no correlation with people sleeping under a bed net and people that had a fever or malaria. There is a weak positive correlation (0.42) between people sleeping under a bed net and people that had malaria. This indicates that an increase in the percentage of people sleeping under a bed net also means the percentage of people that had fever or malaria. This is likely due to there being many variables that can influence the occurrence of a fever or malaria. It would be unwise to reject the assumption that sleeping under a bed net is an ineffective method of protection against malaria. We would advise Operation Orphan to continue to teach children how to make bed nets from waste materials as that would encourage at least one method of malaria prevention. With further research, it would be interesting to measure correlation of the rates of malaria against other features taught by Operation Orphan such as mosquito traps, malaria medication and other methods of malaria prevention.

Figure 7: Heatmap of the correlation between the three health indicators

4. What are the rates of cholera mortality and clean water availability per area?

[Solar water disinfection](#) (SODIS) is a technique that consists of placing water into transparent plastic or glass containers (normally [PET beverage bottles](#)) which are then exposed to the sun. The exposure times vary from 6 to 48 hours depending on the intensity of sunlight and sensitivity of pathogens. It has been repeatedly shown to be effective for eliminating microbial pathogens and reducing diarrhoeal morbidity including cholera. It is simple and inexpensive. Part of the Rubbish Science project is teaching children about SODIS with the use of waste bottles. We reviewed the cases of diarrhoea deaths and access to safe drinking water services.

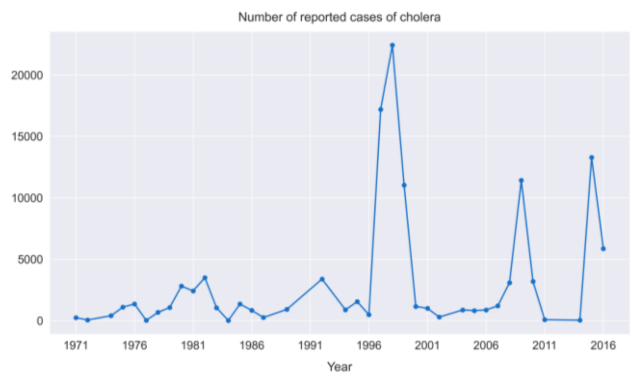

Figure 8: Number of reported cholera cases (2016)

Figure 8 is a time series analysis of the number of cholera cases in Kenya per year, from 1971 until 2016. The highest number of cases of cholera occurred 1998. However, outbreaks of cholera seem to happen every few years (2008 and 2015), indicating that transmission of this disease is not well controlled and that further intervention may still be needed.

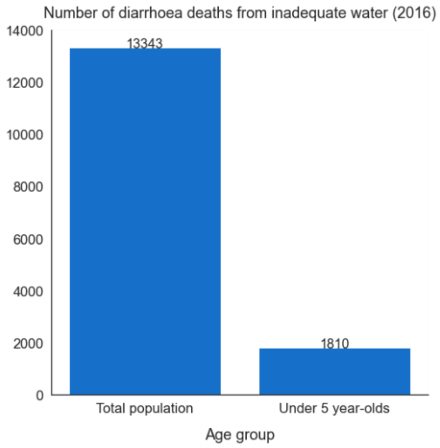

Figure 9: Number of diarrhoea deaths from inadequate water (2016)

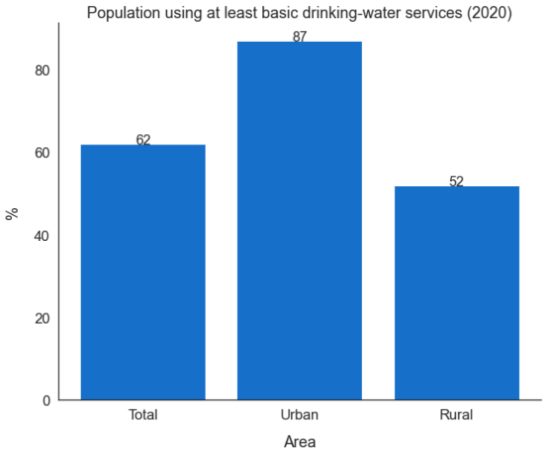

Figure 10: Population using at least basic drinking water services (2020)

Exposure to unsafe water is directly linked to diarrhoeal diseases. Figure 9 shows the number of deaths from diarrhoea due to inadequate water access in Kenya in 2016 (latest available data). This data does not include water used for sanitation or hygiene. More than 13,000 people died due to use of unsafe water, of which 1810 were children under 5 years old. The transmission of waterborne diseases contributes to significant disease outbreaks such as cholera. Thus, access to improved water sources such as piped water, protected dug wells, protected springs, rainwater, and packaged or delivered water, should hinder the transmission of waterborne diseases.

Figure 9: Graph 2 shows the percentage of the population that had access to improved water sources in 2020. The reach of improved water sources in urban areas is 87%, whilst these only reach half the population living in rural areas of Kenya. Overall, almost 40% of the population in Kenya does not have access to safe water sources, making them more vulnerable to waterborne diseases.

CONCLUSION

Operation Orphan should expand to Kenya with their Rubbish Science project as:

- Rubbish is readily available. In collaboration with the government's initiative, Rubbish Science would be a great project to aid the upcycling of recyclable waste material.
- Education drops vastly after primary school, most likely due to the introduction of school fees. Rubbish Science's educational project is likely to be beneficial for communities.
- There are high rates of malaria which have a weak positive correlation with the use of bed nets. This should not be a deterrent to the usage of bed nets in malaria prevention. Rubbish Science should share their focus of teaching various methods of mosquito-borne disease prevention such as mosquito traps.
- The inexpensiveness and ease of SODIS would be an excellent lesson in Rubbish Science. It would teach children methods of creating safe drinking water to care for themselves and their families and hopefully reduce the cases of cholera.

Operation Orphan was initially considering implementing the project in Mombasa. We recommend that the counties with high population and low education levels should be the initial areas considered for Rubbish Science such as Nyeri and the counties in the west, southeast and northeast. The target audience should be all children from underserved communities, regardless of their gender. In addition, we recommend the emphasis on practical DIY experiments regarding mosquito nets alongside traps and water cleansing.