CONNECTING LAST-MILE WOMEN ENTREPRENEURS WITH DATA: A CASE STUDY ABOUT SOLAR SISTER'S GENDER DATA JOURNEY

March 2023

Written By:

Kathryn Farley (Solar Sister), Natasha Wheatley (Solar Sister), Sinead Maharrey (Solar Sister), Jackie Mwaniki (Fraym), Ilse Paniagua (Fraym), Yulia Korableva (Fraym), Kevin Lustig (Data Science Consultant), and Catarina Morais (Solar Sister)

ABOUT US

Solar Sister invests in women's enterprise in off-grid communities by recruiting, training, and supporting Solar Sister Entrepreneurs to run successful and sustainable clean energy businesses in their own communities. Solar Sister is always seeking to identify and pilot new and innovative ways to help female Solar Sister Entrepreneurs sustain and grow their business and ultimately increase clean energy access in underserved, last-mile communities.

INTRODUCTION

In 2021, with support from data.org, in partnership with the Mastercard Center for Inclusive Growth and The Rockefeller Foundation, Solar Sister received funding from the *Inclusive Growth and Recovery Challenge* to embark on a data exploration project, using the power of data science to put data directly into the hands of female clean energy entrepreneurs in last-mile communities in Sub-Saharan Africa. To strengthen Solar Sister's goal of bringing light, hope, and opportunity to women and their communities in Sub-Saharan Africa, Solar Sister partnered with Fraym, a global provider of geospatial data for understanding population dynamics, to produce easily accessible, high-resolution data for female clean energy entrepreneurs in Tanzania, Nigeria, and Kenya, allowing them to use data to identify where target customers may live.

Through this project, Fraym and Solar Sister also developed data-driven profiles that allow Solar Sister field staff to identify where potential Solar Sister Entrepreneurs are located, thereby creating operational efficiencies in expansion and recruitment activities. Fraym created valuable data visualization outputs for this project using its interactive dashboard, DATAFraym, which Solar Sister has now integrated into its core business model and uses to make data-backed decisions at all levels of the organization.



ABOUT THIS REPORT

This report is a summary of Solar Sister and Fraym's journey through this collaborative, data-driven process. It documents their challenges, lessons learned, and recommendations for those embarking on a similar data-led journey. This report is for organizations that are interested in and/or thinking about how to use data in new ways to create social impact, enabling people and communities around the world to not only survive but thrive. Both partners hope that this will be a useful tool for like-minded social entrepreneurs and organizations who have yet to embark on or are at any point in their own data journeys. Solar Sister has learned so much through this project and hopes to share it with other organizations through this report.

PROJECT ACTIVITIES & METHODS

This project sought to use data science to develop a methodology for providing data-backed insights for Solar Sister Entrepreneurs, field staff, and Solar Sister's Leadership Team to bolster strategic decision-making, improve entrepreneur recruitment and retention, and support female entrepreneurs in sustaining and growing their clean energy businesses. Through this partnership, Solar Sister strengthened data-based decision-making at all levels of the organization. Of particular note, Solar Sister was able to provide African women entrepreneurs in underserved communities access to data for the first time, which allowed them to understand where target customers live, thereby amplifying their impact.

KEY TAKEAWAYS

1. Take stock of the data capabilities, needs, and desires of the population(s) you are looking to serve.
2. Anticipate possible data limitations from the outset and be ready to pivot.
3. Budget time and resources for identifying and unpacking biases in datasets and leave room for iteration.
4. Remember that you don't know what you don't know and be prepared to evaluate and pivot as you go along.
5. Meet end-users where they are and tailor data dissemination and visualization tools accordingly.
6. Think about the quality and validity of the data you are producing and consider how you will generate it in a feasible and streamlined manner.

The project was executed in the following phases, as illustrated in Figure 1 below.

DESIGN AND CONDUCT AN INITIAL NEEDS ASSESSMENT

As a first step in this data-driven process, Solar Sister designed an assessment to understand the data needs and gaps within the organization. With the help of a consultant, Solar Sister administered surveys to team members at various levels of the organization. These surveys sought to understand he current use of data and level of data literacy across different Solar Sister roles, with a particular focus on seeking to understand if and how Solar Sister staff and Solar Sister Entrepreneurs were using data to make business decisions.

The survey results confirmed the impetus of this project; that the selection of sales and recruitment areas was not currently data-driven. Solar Sister learned that, across the organization, there was limited knowledge of what data could be used and how to use it for decision-making. Respondents also highlighted a significant interest in learning and bolstering their data literacy.

Taking the time to develop an organizational needs assessment allowed Solar Sister to validate the need for the project and helped both partners to understand what kinds of data literacy gaps existed for end-users, as well as how best to address them. The needs assessment survey indicated a desire for improved tools and skills that would allow Solar Sister Entrepreneurs to make informed decisions about where and to whom to sell products in order to maximize their profits and community impact. The project was developed with this, and the understanding that Solar Sister Entrepreneurs currently reported a low level of data literacy, in mind.

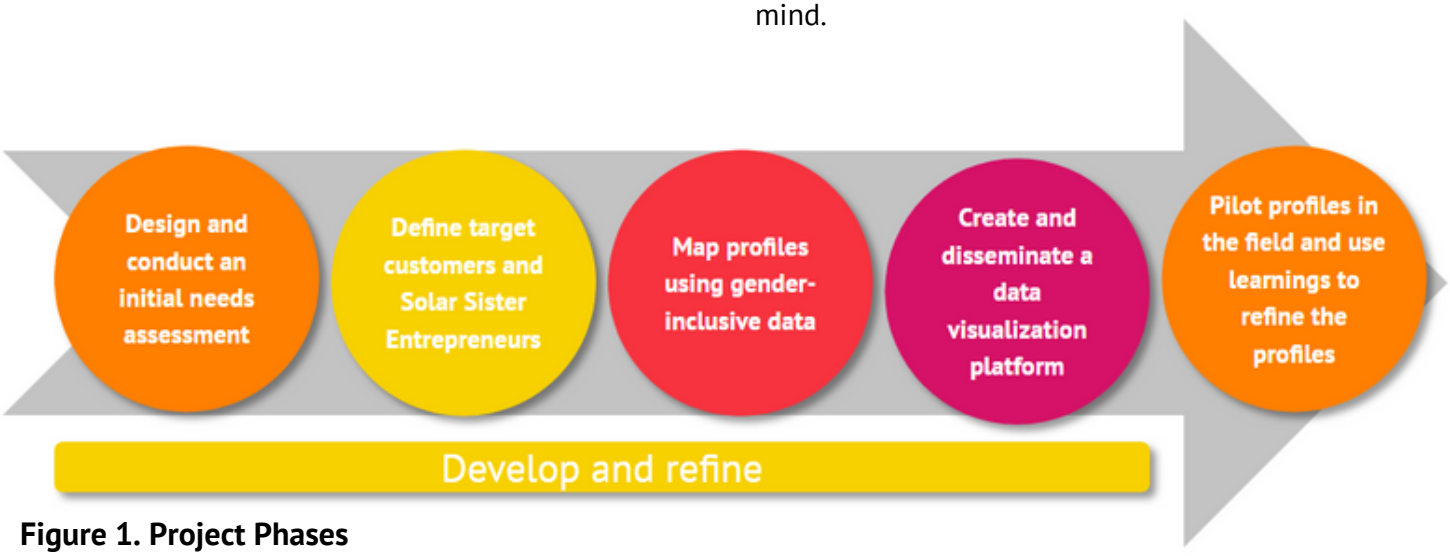


Figure 1. Project Phases

DEFINE TARGET CUSTOMERS AND SOLAR SISTER ENTREPRENEURS

The next step in this process was to map and create profiles of Solar Sister customers. Fraym and Solar Sister began with an analysis of target customer characteristics using existing sales data. Fraym extracted sociodemographic characteristics for each geographic sales area using country-level survey data and analyzed the relationship between sales and the geographic population demographics, such as asset ownership, employment, and access to clean water. They then analyzed the demographics of those communities where the highest-performing Solar Sister Entrepreneurs work in order to understand what types of communities Solar Sister Entrepreneur businesses are most successful in.

KEY TAKEAWAY

Anticipate and actively seek to identify any possible data limitations from the outset and be ready to pivot and try different strategies to test out assumptions and to ensure data validity and usefulness.

Through this initial analysis, the partners determined that Solar Sister Entrepreneurs with the top sales are typically located in medium-sized urban towns with high levels of quality housing, access to clean drinking water, and women who have completed primary school. The characteristics of high-sales communities pointed primarily to urban environments, home to wealthier customers. This posed an additional challenge, which was a theme throughout this activity, that although a goal of the project was to highlight communities in which Solar Sister Entrepreneurs might scale successful and profitable businesses, Solar Sister's mission is also focused on serving communities that can most benefit from solar energy products. Fraym and Solar Sister had to be cognizant throughout this process of the polarity of those two goals.

KEY TAKEAWAY

Before getting too far along in your data journey, make sure to take stock of the data capabilities, needs, and desires of the population(s) you are looking to serve.

Additionally, since there are so many different factors that go into a successful Solar Sister Entrepreneur business beyond just geographic and customer characteristics, this exercise could not say with confidence that a Solar Sister Entrepreneur's success was attributed to the type of area she operates in. For example, the differences between where Solar Sister Entrepreneurs had high levels of success could be caused by the characteristics of Solar Sister Entrepreneurs themselves rather than the customers in those areas.

Consequently, since the initial data proved less useful than originally hoped, target customers and entrepreneurs had to be defined manually based on the anecdotal knowledge of Solar Sister staff, rather than by using data science or Artificial Intelligence. There are potential gaps and limitations in this data as a result, including the possibility of unintentional human bias in the creation of the profiles. To evaluate and

address these potential biases, Solar Sister and Fraym amended the scope of work and project timeline to include a pilot period to evaluate the validity and usefulness of the profiles.

MAP PROFILES USING GENDER-INCLUSIVE DATA

To further define target customers, Solar Sister and Fraym categorized customers into three profiles: 1) populations with the ability to acquire products with cash; 2) populations with the ability to acquire products on credit; and 3) populations with the ability to benefit from Productive Use of Energy (PUE) products. Populations without the need for solar energy products and populations without the ability to acquire solar products on cash or credit were excluded from the profiles based on the understanding that people without the need for solar would not want to buy products and that those that could not afford to buy products cannot be customers by default. This decision was also based on Solar Sister's preexisting knowledge of the types of customers who tend to be able to buy products.

Solar Sister and Fraym engaged in several discussions to try to identify communities where the largest "mission" impact could be made in improving energy access while still maximizing profit for Solar Sister Entrepreneurs, but the data and analysis ultimately did not support the ability to do this with confidence. For this reason, the profiles focused on customers with both the need for energy access and the ability to pay for products.

Starting from this base population of customers with energy demand and some ability to pay,

three customer profiles were built out using a variety of Fraym indicators generated by a cutting-edge process that turns household survey data into census-like data across entire countries. Similarly, together with Solar Sister Africa field teams, the partners identified the key characteristics of an ideal Solar Sister Entrepreneur, and from there developed the Solar Sister Entrepreneur recruitment profile focused on the idea that Solar Sister is aiming to recruit women living in communities with a need for solar products who have a modest level of purchasing power, entrepreneurship experience, and at least two children or young adults in the household.

Defining the profiles in a gender-inclusive way was critical to this process, as Solar Sister recognized even before the project began that women are often unintentionally left out of traditional datasets and country-level data indicators. Several versions of each profile were considered and analyzed to create profiles that aimed to capture all women within each defined customer category and in the Solar Sister Entrepreneur recruitment profile. It is important to note that even though both Fraym and Solar

Sister identified this gender-inclusive approach as a key part of the data science journey from the outset and as a priority for the project, the partners underestimated the amount of time and effort that would be needed to identify, unpack, and counteract these gaps to ensure a more robust and gender-inclusive dataset. As such, project partners had to create new and innovative feedback loops to ensure the quality of outputs. This proved to be time-consuming and challenging, as there was a general scarcity of gender-inclusive data that matched the needs of the project.

KEY TAKEAWAY

Thinking about biases in datasets is critical to any data journey. However, identifying and unpacking these must be an intentional process and can be one that is time-consuming and challenging. Make sure to budget for this accordingly in project planning and leave room for iteration as needed.

More details on this process will be described later in the report. Details on the profile breakdowns can be found in **Appendix A** and resources for those looking to ensure gender inclusivity in their data science work can be found in **Appendix B.**

CREATE AND DISSEMINATE A DATA VISUALIZATION PLATFORM

Fraym's deliverables for the project, as outlined in the original scope of work, were to use Fraym's spatial data to 1) create customer profiles; 2) create an entrepreneur recruitment profile; and 3) map those profiles at the local level to identify "hotspots'' – where the profiles produced by Fraym indicates a large concentration of potential customers or entrepreneurs reside – for sales and Solar Sister business expansion. The resulting geographic map of hotspots was originally designed to be presented as a series of static maps. However, as the partnership progressed it became clear that these maps would not provide the community-level data that was ultimately needed to achieve Solar Sister's goals. Solar Sister expressed the need to dynamically select and zoom into communities to be able to provide more localized data tailored to individual Solar Sister Entrepreneurs and field staff. In response, Fraym adjusted the contract to provide Solar Sister with access to its interactive DATAfraym platform. DATAfraym includes comprehensive visualizations and allows users to highlight specific locations and export data for additional analysis in a user-friendly way.

DATAfraym users can use dashboards to zoom in and out of each layer, analyze community variation in target customers at the one-by-one square kilometer, filter using a combination of layers, explore correlations between layers, and

KEY TAKEAWAY

When beginning your data journey, remember that you don't know what you don't know. While you may think that one type of data output or visualization is what you need at the outset of a project, be prepared to evaluate this and pivot as you go along. Data science and implementation partners should discuss early and often, using visuals, what deliverables make the most sense so that both partners understand what is needed and what they are agreeing to.

export socio-demographic statistics for the chosen localities. By combining different data layers, it is possible, for instance, to analyze average levels of energy spending, women's decision-making power, and informal savings group participation for communities that have high or low concentrations of target customers. This analysis generated insights about why some communities might have low or high concentrations of target customers in addition to highlighting customer concentration areas.

More details about the technicalities of the DATAfraym dashboard created for Solar Sister can be found in **Appendix C**.

PILOT PROFILES IN THE FIELD AND USE LEARNINGS TO REFINE THE PROFILES

Prior to the pilot, Fraym conducted two training sessions for Solar Sister's Leadership Team on DATAfraym's features and functionalities for hotspot identification, accompanied by a training guide with practical exercises. However, it became apparent that there was a significant gap between Solar Sister field staff's data literacy and DATAFraym's training and platform. To address this gap, field staff were trained on how to interpret and analyze the data produced by DATAfraym to build their capacity to make decisions about where to prioritize their sales efforts across a number of potentially strong customer areas.

Following the training, field staff received data on five geographic areas near them that were identified as potential recruitment hotspots, and applied their training to use the data they received to analyze those areas and make data-driven decisions on where to prioritize their recruitment focus. Solar Sister believes that this initial training will serve as a building block for future capacity development for hotspot reading, and plans to introduce field staff to the interactive DATAfraym platform in 2023 so that they can continue to grow their data literacy and be able to use the tool themselves.

Additionally, Solar Sister faced challenges in determining how best to disseminate data to Solar Sister Entrepreneurs, many of whom did not have access to the technology needed or the data literacy skills to utilize the DATAfraym platform themselves. To combat this, Solar Sister's team extracted tailored data from DATAfraym and matched ideal customer hotspots to each individual entrepreneur. To disseminate this data, Solar Sister created an SMS campaign to message entrepreneurs who were matched with a hotspot with the geographic information for the hotspot.

KEY TAKEAWAY

Meet data users where they are and tailor data dissemination and visualization tools accordingly. Don't underestimate the ability of new users to learn how to read, understand, and use data for decision-making, but understand that this is a process and design trainings and ongoing coaching and support accordingly to build data literacy and skills.

Solar Sister gathered feedback from Solar Sister Entrepreneurs and field staff following the pilot. This feedback and the learnings from the pilot were leveraged to develop and refine the profile

and hotspot parameters and to improve the quality of the data Solar Sister Entrepreneurs and field staff received. Feedback from Solar Sister Entrepreneurs included requests that future hotspots take travel and security challenges into account. Additionally, feedback revealed that mapping prominent landmarks or institutions is helpful so that Solar Sister Entrepreneurs and field staff know how best to reach the recommended areas. Based on these comments and suggestions, Fraym and Solar Sister worked together to update DATAfraym to include a mapping option on the dashboard which now shows the closest health clinic to a hotspot location - this information was also included in the SMS campaign, which gives Solar Sister Entrepreneurs and field staff a landmark to know where to go when we send them a hotspot.

DEVELOP AND REFINE

Effective matching of Solar Sister Entrepreneurs with potential sales locations is essential to ensuring Solar Sister Entrepreneurs can be successful in reaching customers efficiently, and at a minimum cost, time, and effort. However, the task of matching Solar Sister Entrepreneurs with customers has proven both technically challenging as well as time-consuming. With hundreds of both Solar Sister Entrepreneurs and hotspots in each country, the number of potential combinations of assignments of Solar Sister Entrepreneurs to customer locations is extensive.

Initially, Solar Sister team members performed this task manually using export tools in DATAfraym. To expedite this process, following the pilot Fraym developed a more advanced hotspot identification export utility for future use. Even after this improvement was made, it became evident that hotspot identification was still too time-consuming, and thus would not allow for the eventual scale-up of hotspot matching for Solar Sister's cadre of more than 4,000 active Solar Sister Entrepreneurs. Thus, Solar Sister contracted a consultant to develop an automated hotspot matching tool. The product of that effort was an interactive application that accepts input lists of both Solar Sister Entrepreneurs and hotspots exported from DATAfraym and performs algorithmic smart matching to identify the ideal hotspot for each Solar Sister Entrepreneur. This smart matching will allow Solar Sister to prioritize both the proximity of each individual Solar Sister Entrepreneur to each of the hotspots and the lowest cumulative distance for all Solar Sister Entrepreneurs to travel to their respective sales areas, automatically optimizing hotspot suggestions for all Solar Sister Entrepreneurs. The hotspot matching tool optimizes potentially millions of combinations of Solar Sister Entrepreneurs and hotspots in just a few seconds, saving many hours of previously manual effort.

The result of this work was not just a significant reduction in the level of effort to spread the insights derived from the customer profiles to Solar Sister Entrepreneurs, but also further insights into the distribution of customers and Solar Sister Entrepreneurs in Solar Sister's network. When the distance to travel to customers was limited to ten kilometers or less, it quickly became apparent that many Solar Sister Entrepreneurs do not live close enough for practical travel to any hotspots identified by the customer profiles, while many hotspots have no nearby Solar Sister Entrepreneurs. One sample dataset used for developing the matching tool

produced a suggested hotspot for only about one-third of the active Solar Sister Entrepreneurs. This has implications for future Solar Sister strategies in recruiting new Solar Sister Entrepreneurs and supporting existing ones. It also raises questions of prioritizing hotspots based on criteria other than the volume of customers and geographic distance that Solar Sister continues to explore.

The matching tool will continue to be refined and used to streamline Solar Sister processes and to inform future strategic decision-making among Solar Sister's Leadership Team.

KEY TAKEAWAY

It is important to think about the quality and validity of the data you are producing and how it will be visualized and disseminated. However, don't forget to also consider how you will generate this data and what level of effort will be needed to produce it. If data is too time-consuming to package in a usable fashion, you may need to find new ways to streamline the process to ensure that it is possible to scale.

CHALLENGES AND LESSONS LEARNED

The overarching takeaway from Solar Sister's data journey is to expect the unexpected. As with any new project or innovation, not everything went as planned, and the involvement of decision-makers from different fields created both new opportunities and challenges. From this project, Solar Sister learned to continuously adapt and to always be ready to learn as you go. They also learned the importance of creating a plan and process for and adding time into the work plan for both partnership development and ongoing reassessment.

Below are the key considerations that emerged from this project, drawn from challenges encountered and lessons learned.

EXPECT THAT PARTNERS WITH DIFFERENT AREAS OF EXPERTISE WILL HAVE DIFFERENT PERSPECTIVES AND THAT PARTNERSHIP DEVELOPMENT WILL TAKE TIME AND INTENTIONALITY

This project required flexibility from both understandings. As an implementing partner, Solar Sister brought expertise to the project in the programmatic areas of social enterprise, clean energy, gender, and the geographical contexts in which they were working. Fraym brought valuable expertise in data science and analytics, which addressed a gap identified during the data needs assessment at the project's onset. Although Solar Sister is well-versed in how best to use data and analytics within their own programs for monitoring and evaluation and decision-making, the assessment and collaboration that followed

highlighted gaps in Solar Sister field staff's understanding of data science and visualization techniques as well as in end-user data literacy.

It took time and intentionality on the part of both partners to build and address these and other gaps in understanding and communication. For example, Solar Sister had to build upon their knowledge of what type of data insights were possible, what visualization tools made the most sense given the implementation context, and how much effort various data-related tasks would require. Similarly, Fraym had to learn how to apply a gender lens to data analysis and how to generate data that would be relevant for and useful in last-mile communities.

This collaborative learning process was highlighted with the customer mapping deliverable; as the partnership progressed, both Fraym and Solar Sister realized that the original deliverable of static maps was ultimately not a functional product for Solar Sister due to end-user usability and overall functionality. Met with this challenge, the partners went back to the drawing board to realign goals and deliverables and adapt the mapping product to the interactive data system. This required both partners to listen and to learn from one another to ensure that project deliverables were meeting programmatic needs and that both partners were aligned on expectations as the project evolved.

MAKE SURE TO UNDERSTAND END-USER DATA LITERACY GAPS

Once the mapping for customer hotspots was created, Solar Sister was faced with the challenge of how to integrate these new and valuable insights into their Solar Sister Entrepreneur model. Solar Sister Entrepreneurs often lack regular access to technologies typically used in data analysis. None own a computer, and very few have access to a smartphone. The obvious ways to distribute data – through an interactive platform, emailing charts and tables, or through an app – were not possible within the end-user context of Solar Sister. Distribution of physical materials, such as the initial static map print-outs, was logistically challenging given the rural and remote environment where many Solar Sister Entrepreneurs live. Additionally, data concepts were unfamiliar to many Solar Sister Entrepreneurs, which means there was little established trust in data or understanding of how to use it.

To deliver the data-based recommendations, the partners learned that it was important to find a method of delivery that beneficiaries were already comfortable with instead of introducing new technology. The solution identified for Solar Sister Entrepreneurs' comfort was for the Solar Sister Leadership Team to deliver data recommendations from DATAfraym via SMS in short, basic language. SMS is widespread in countries where Solar Sister works, including in more rural, last-mile areas, so it was easy to roll out for users.

Using regular, consistent messages for hotspot recommendations in plain language provides a channel of repeated and non-confrontational exposure to data. However, it's important to remember that one size does not fit all, and each organization on its own data journey should figure out what works for them and for its end-users. Also, of note, is that end-users access to technology and capabilities to understand and use data change over time. Methods of data dissemination, data visualization, and data literacy and use training should adapt and be tailored accordingly.

GENDER INCLUSIVITY IN DATA SCIENCE TAKES DELIBERATE CREATIVITY

Significant gaps exist in data on women and girls globally. Commonly-used indicators often unintentionally leave women out (employment data, for example, may neglect to include informal workers like entrepreneurs, market vendors, and those providing childcare, which fails to account for many women, who are more likely to work within these informal structures). Data-driven decision-making in the context of programming for women is a significant challenge across the sector, and the gendered gaps that exist in data are vital to consider in projects such as this one to ensure women are not being left out of the conversation.

With this in mind, the project team had extensive discussions to think through ways existing data could be used or altered to ensure that all profiles created were gender-inclusive. Discerning how to do this with the data that currently exists in the field was extremely challenging, but ultimately the team was able to find ways to mitigate the effects. Several modifications were made to the customer and Solar Sister Entrepreneur recruitment profiles to ensure that they were capturing all relevant women and being inclusive of traditionally marginalized and excluded populations.

For example, the project team explored ways to define "creditworthiness" that would not unnecessarily exclude potential female borrowers based on employment types predominantly filled by women. The team ruled out a requirement that at least one employed adult in the household must be paid in either cash or in-kind, given that many women in Solar Sister's operating context are paid through more informal channels and not necessarily captured in this way. For example, many women work directly for family businesses where wages are not expected or negotiated ahead of time (i.e. herding cattle, working in family shops, etc.) or they are dependent on the long-term financial gain from their labor as opposed to regularly-scheduled/anticipated payments. Being mindful of these gaps in data was critical given Solar Sister's beneficiaries are women, and gender data gaps are an important consideration for any organization working in this space.

Please review **Appendix B** for more thorough information on gender inclusivity indicators.

CONCLUSION AND RECOMMENDATIONS

Partnerships between implementing organizations and data science firms can lead to valuable capacity-building and data-driven products. Solar Sister's partnership with Fraym, for example, to provide female clean energy entrepreneurs with access to tailored customer data, was an innovation that will drive data-informed decision-making, which will provide Solar Sister Entrepreneurs with a valuable tool for improving their businesses and quality of life and allow them to market, sell, and deliver clean energy technologies to underserved communities in a more efficient way.

As with any new partnership or innovation, this project was not without challenges. However, both partners were committed to learning from this experience and are keen to share lessons learned with the broader community. To this end, the following recommendations for mission-driven organizations hoping to partner with experts in the data science field to build organizational capacity for data-driven decision-making have been compiled:

KEY RECOMMENDATIONS

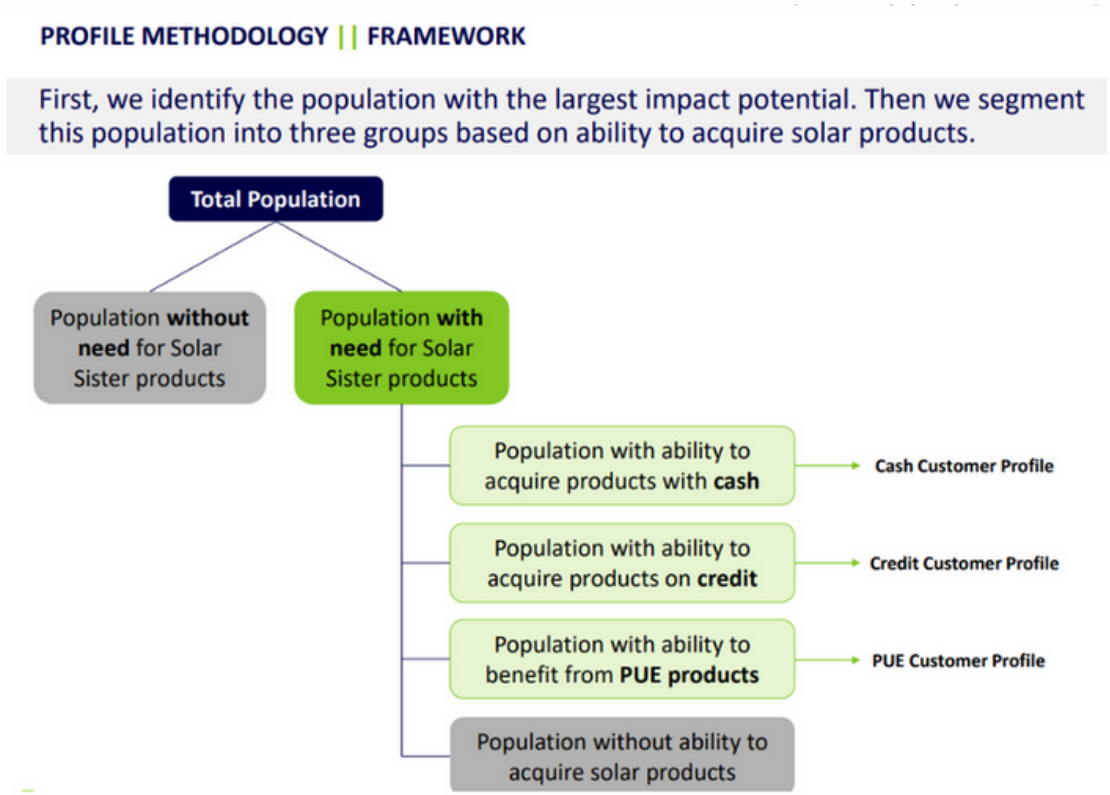
- Before getting too far along in your data journey, make sure to take stock of the data capabilities, needs, and desires of the population(s) you are looking to serve.
- Anticipate and actively seek to identify any possible data limitations from the outset and be ready to pivot and try different strategies to test out assumptions and to ensure data validity and usefulness.
- Think about biases in datasets, as this is critical to any data journey. Understand that identifying and unpacking these must be an intentional process and can be one that is time-consuming and challenging. Make sure to budget for this accordingly in project planning and leave room for iteration as needed.
- Remember that you don't know what you don't know. While you may think that one type of data output or visualization is what you need at the outset of a project, be prepared to evaluate this and pivot as you go along. Data science and implementation partners should discuss early and often.

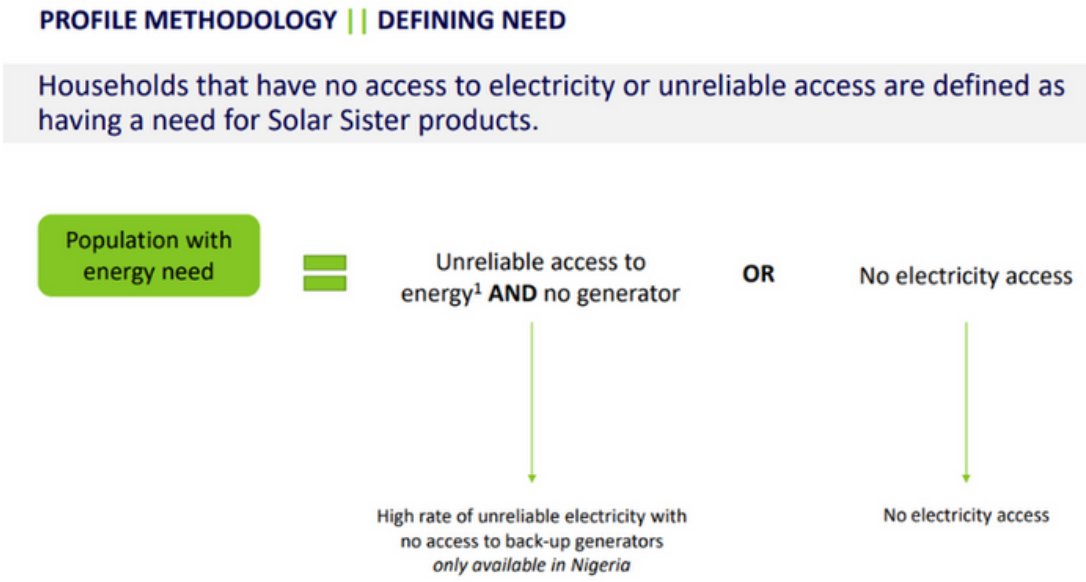
- Meet end-users where they are and tailor data dissemination and visualization tools accordingly. Don't underestimate the ability of new users to learn how to read, understand, and use data for decision-making, but understand that this is a process and design trainings and ongoing coaching and support accordingly to build data literacy and skills.
- It is important to think about the quality and validity of the data you are producing and how it will be visualized and disseminated. However, don't forget to also consider how you will generate this data and what level of effort will be needed to produce it. If data is too time-consuming to package in a usable fashion, you may need to find new ways to streamline the process to ensure that it is possible to scale.

APPENDICES

APPENDIX A: FINAL CUSTOMER AND SOLAR SISTER ENTREPRENEUR (SSE) RECRUITMENT PROFILES

The final customer and recruitment profiles are outlined in detail below:





PROFILE METHODOLOGY || CASH CUSTOMERS

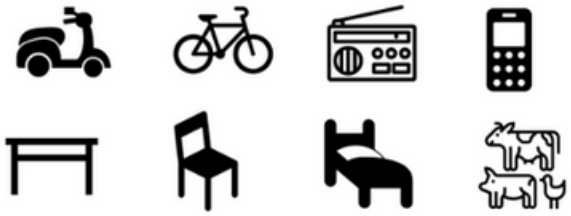
Cash customers are households that have a modest ability to pay for Solar Sister products demonstrated through ownership of one or more assets.

Population with ability to acquire products with cash

=

Some purchasing power
Household owns at least one asset (scooter, bicycle, radio, mobile phone, bed, table, chair or livestock[1])

Most common assets owned by households affected by poverty[2]:



Defining the profile:

- *We simplified the first brainstormed list of consumer characteristics by lowering the requirements for purchasing power (we removed durable housing and employment conditions)*
- *Added bed, table and chair to list of assets to be inclusive of female-headed households*

Note 1: Livestock includes goats, sheep, chicken, cattle and horses.
Note 2: Bed, table, and chair are missing in the data available for Tanzania

PROFILE METHODOLOGY || CREDIT CUSTOMERS

Credit customers are households with some level of income indicating they are credit-worthy, in addition to having modest purchasing power.

Population with ability to acquire products on credit

=

Some purchasing power
Household owns at least one asset (scooter, bicycle, radio, mobile phone or livestock[1])

+

Credit-worthy
Household has a stable income that can be used for credit payments

Employment Characteristics of Cash Customer Households in Tanzania[1]			
Indicator	**All Households**	**Female-Headed Households**	**Male-Headed Households**
Credit Profile Option 1: At least one adult in the household is currently working	82%	75%	90%
Credit Profile Option 2: At least one adult in the household is currently working and not as a seasonal subsistence farmer	49%	56%	48%
Credit Profile Option 3: At least one adult in the household is currently working, with pay either in cash or kind, and not as a seasonal subsistence farmer	41%	49%	40%
The only employed adults in the household are subsistence farmer	49%	39%	52%
The only employed adults in the household are employed seasonally	45%	40%	47%
The only employed adults in the household are seasonal subsistence farmers	33%	30%	38%
The only employed adults in the household are unpaid	28%	25%	29%
The only employed adults in the household are unpaid and not as seasonal subsistence farmers	8%	7%	8%

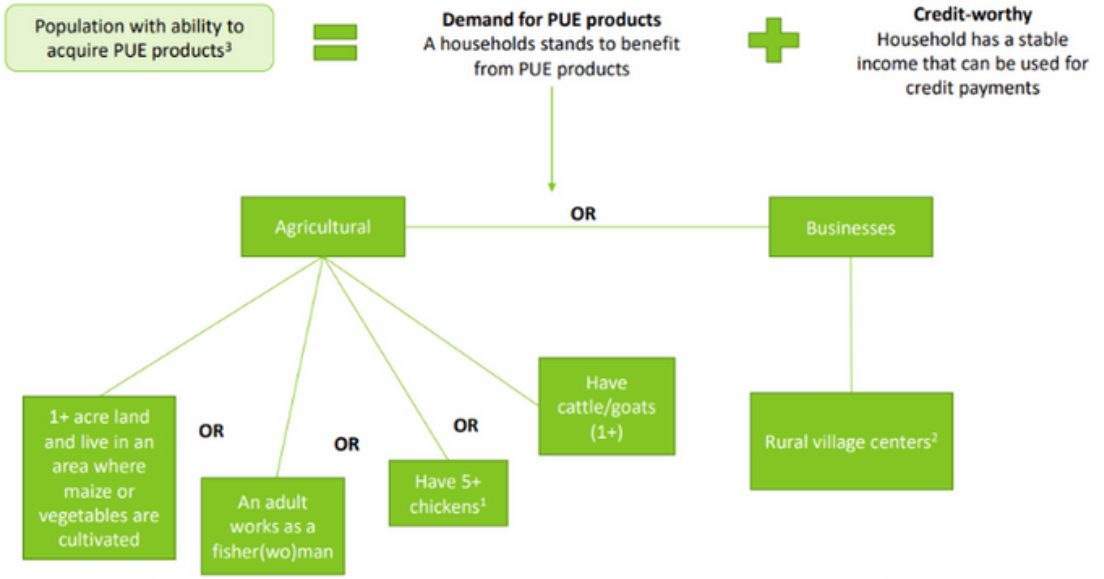
Defining the profile:

- *We chose credit profile option 2, which excludes seasonal subsistence farmers from employment after analysis showing that these households have lower savings and a significantly lower ability to purchase on credit*
- *Adding the requirement that at least one employed adult in the household must be paid in either cash or kind only removes an additional 8 percent of households. Since this employment situation is most common among women, we did not choose option 3.*

Note 1: Adults are defined as individuals aged 15 to 49. All differences between female-headed and male-headed households in the table are statistically significant at the 95 percent confidence interval.
Source: Fraym

PROFILE METHODOLOGY || PUE CUSTOMERS

A PUE customer is an individual who lives in a household with a demonstrated demand for PUE products, and an ability to pay.



Note 1: The threshold of at least five chickens was calculated assuming one chicken can produce six eggs per week, to match the need for 100 eggs in 21 days to fully benefit from the Solar Sister egg incubator product.
Note 2: Rural village centers can be identified with a combination of the Global Human Settlement Layer (GHSL) and a population threshold using WorldPop.
Note 3: This profile may not identify all the potential users of water pump and solar incubator products, as these products could be sold in communities have

PROFILE METHODOLOGY || SSE RECRUITMENT

Solar Sister is aiming to recruit **women** living in communities with **a need for solar products** who have a **modest level of purchasing power**, **entrepreneurship experience**, and **at least two children or young adults** in the household.

Ideal SSEs for recruitment are women over 30 years old who meet the following qualifications

Qualification	Definition
Lives in communities with a need for solar products	At least half of the local community (1km squared) doesn't have access to electricity.
Has modest purchasing power	The woman lives in a household that is in the top 60th percentile of the DHS wealth index, indicating she may have access to the capital required to buy into the business.[1]
Has entrepreneurship experience	The woman has been self-employed in the past year, indicating she has some entrepreneurship experience.
Has at least two children or young adults in the household	The woman lives in a household with at least two people 30 years old or younger.

APPENDIX B: GENDER FRAMEWORK TABLE

The following table presents the indicators that were considered as part of the target customer and Solar Sister Entrepreneur recruitment profiles. It details if each indicator was included or excluded, and why, highlighting gender-inclusivity considerations where applicable.

Variables Considered for Inclusion in Profile Filtering Criteria				
Indicator	**Indicator Inclusion in Customer Profiles**			**Gender Considerations for Inclusion/Exclusion of Indicators In Each Profile**
	Cash	**Credit**	**PUE**	
Durable housing (material of floor, wall, or roof)		x	x	Previous success by Solar Sister in selling to households (including and especially to female-headed households) without durable housing, as traditionally defined, showed that this criteria needed to be excluded from the cash profile, but it was determined to be a necessary condition for other customer types who would be purchasing larger solar products, including PUE, on credit.
Employment		x	x	Contextual knowledge from Solar Sister indicates that women living in lower-income households with enough purchasing power for low-cost solar lights may not always be employed, especially as it is defined in the Demographic and Health Survey, which tends to focus on more formal, male-dominated forms of employment. Thus, it was left out of the cash customer profile. However, employment indicates a household has a stable income that can be used for credit payments, so it was included in the credit and PUE customer profiles.
Asset ownership of a bed, table, and chair	x	x	x	Contextual knowledge from Solar Sister indicated that women are more likely to drive the purchase of assets for the household and that the existence of these assets can be treated as a necessary condition to be a solar product customer at all levels. Fraym and Solar Sister worked together to evaluate which assets women are most likely to own, as the original list of assets was more male-dominated and would have unintentionally excluded women customers. The final list included a bed, table, and chair to be more inclusive of female-headed households.
Seasonal subsistence farming		x	x	Analysis showed that households where the only employed adults are employed in seasonal subsistence agriculture have low rates of savings and monthly spending levels that are too low to accommodate credit payments. Thus, this was included as a requirement in the credit and PUE customer profiles. However, Solar Sister and Fraym decided not to include it in the cash customer profile as they did not want to exclude women, who often participate in seasonal subsistence farming but who Solar Sister knows to be good cash customers for Solar Sister products.
Employed adult is paid in either cash or kind				Adding the requirement that at least one employed adult in the household must be paid in either cash or kind only removed an additional eight percent of households. Since this employment situation is most common among women, Solar Sister and Fraym chose not to add this requirement into any of the profiles.
Employed year-round				A large majority of women aged 31 – 49 who have been self-employed in the last year are currently employed year-round. If Solar Sister and Fraym exclude women who are employed year-round, they would be excluding a large portion of entrepreneurial women. Thus, this was not included in any of the profiles.

APPENDIX C: DATAFRAYM DETAILS

Fraym's DATAFraym platform was designed to combine machine learning-enhanced datasets with population indicators and Solar Sister project data. DATAFraym is easy to use and requires no GIS expertise. It includes comprehensive visualizations to facilitate understanding by a non-technical audience, is compatible with modern browsers, monitors, laptops, and tablets, and allows users to highlight specific locations and export data for additional analysis.

In addition to providing visualization of data layers, DATAFraym allows Solar Sister to export contextual statistics at different geographic levels – national, regional, and community (local user-defined areas). DATAFraym users can use dashboards to freely zoom in and out of each layer, analyze community variation in target customer concentration at the one-by-one square kilometer level, filter using a combination of layers, explore correlations between layers, and export socio-demographic statistics for the chosen localities. By combining different data layers, it is possible to analyze average levels of energy spending, women's decision-making power, and informal savings group participation for communities that have high or low concentrations of target customers. This analysis generates insights about why some communities might have low or high concentrations of target customers.

Fraym created three dashboards for Solar Sister in their areas of operation as of 2021-2022: Nigeria, Tanzania, and Kenya. Each dashboard includes the proportion and number of people that belong to cash, credit, PUE, and Solar Sister Entrepreneur recruitment profiles.

Each dashboard also includes four contextual layers to aid Solar Sister in their analysis. The four layers are:

1. Energy-based wealth inequality;
2. No electricity or unreliable electricity access;
3. Participation in informal savings groups; and
4. Women's participation in the informal economy.

Energy-based wealth inequality is a data layer created to highlight gaps in wealth between electrified and non-electrified households. When combined with a customer profile, this layer can help target high-wealth gap communities that could benefit from Solar Sister products. Similarly, no electricity or unreliable electricity access can help identify communities that could be enriched with solar energy products. Participation in informal savings was added as a contextual layer to serve as an extra filter during hotspot analysis. By combining participation in informal savings with the credit profile, it is possible to target only those communities that have high savings and a high density of credit customers, for example. Finally, women's participation in the informal economy can provide context on entrepreneurial behavior among women in communities being considered for Solar Sister Entrepreneur recruitment.

In addition, Fraym added points of interest data to each dashboard. This included the location of Solar Sister Entrepreneurs and the location of health facilities to serve as reference points for Solar Sister Entrepreneurs traveling to the area.

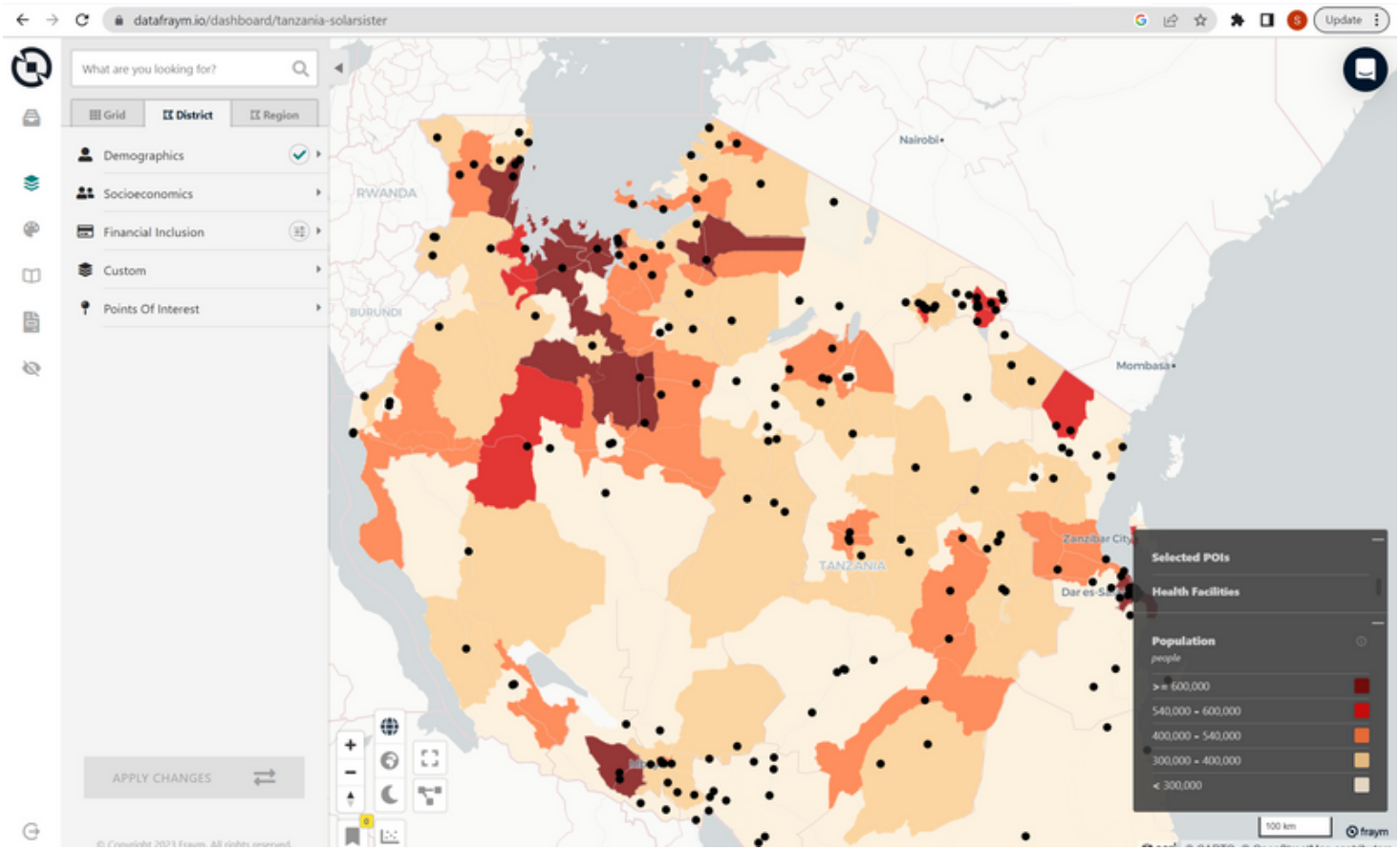


Figure 2. Screenshot from DATAFraym depicting cash customer hotspots in Tanzania