**Data for Democracy Project**

Data for Decision Making Training

*Annotated Instructor Guide*

**Module 1:** Introduction to Data for Decision Making

**Total expected time:** 3 hours

**Objectives:**

* Learn how data are used in everyday life, and understand the importance of data collection, analysis, and sharing.
* Understand the concepts and terms around data for decision-making.
* Learn basic questions to ask and how to define problems or needs within the participants’ community or organization.
* Understand how to be a consumer of data, and what questions to ask in order to use the data for decision making.

**Materials:**

* **Projector**
* **Blackboard/whiteboard (ideally)**
* **Paper**
* **Pencils**
* **Printout of Images**
* **3-4 Sample data sets (still need to find and attach to module)**
* **Activity packet 1.1**
* **Activity packet 1.2**

|  |  |  |
| --- | --- | --- |
| Time | Action | Instructor Notes |
| *5 minutes* | Introductions | Begin by greeting the class and introducing yourself. Provide time to discuss your background, any experience you have with data, and any other relevant information you wish to share.  Outline the structure of the day (two modules with a break in between for lunch), and provide time to address any other administrative issues that need to be brought up before class begins.  Pause to ask if anyone has questions so far. |
| *15 minutes* | Critical Thinking | Next, introduce the name of the workshop, taking time to write “Data for Decision Making” on the whiteboard (if possible). Underscore the word “data”, or verbally highlight it.   * Ask the class to work with a partner and take a moment and think about what they think of when they hear the word “data”. Have the class write down 5 words that they associate with data. If they are having trouble thinking of words, feel free to gently guide them by providing a few words: numbers, words, etc. If you provide this information, ask the participants not to use these words as their answers. (*5 minutes)* * Ask for volunteers to report back on what they wrote. If possible, write their answers on the board. What did they write? Did some people write similar answers as others? Do people agree with other people’s answers? Discuss their answers with the class. *(5 minutes)*   Before transitioning, make sure to underscore to the participants that we are going to jump right in and define “data”, “decision making”, and “data for decision making” and then outline course goals and objectives. |
| *20 mins* | Introducing key concepts | Note to instructor\* This section comes before describing the workshop to provide context for the workshop.  After that the class has shared with each other what they think data are and what they associate data with, offer a formal definition. On the board, write   * Data:   At its most basic level, data are collections of facts and information. This information can come be packaged as numbers, words, measurements, observations, or descriptions.  An example:  If I didn’t eat breakfast and am hungry, I could ask my colleagues where the best place to get Mohinga is nearby. When I ask that question and receive responses, I collect information. Try to solicit the type of information collected from the class. Is it price? How the breakfast tastes? How close the breakfast place is to the office? How long it will take to eat the breakfast?  The information that I collected can be considered data.  Questions to ask:  Now that data have be re-framed to simply be information, ask the class what kinds of data (information) they collect in their everyday life. If the class is quiet and has not provided a response, feel free to nudge them along with a few more examples, such as do they ask their families about their days when they get home? If they want to buy a new cell phone, do they ask for recommendations?   * Decision Making:   After collecting data (information), you are free to make a decision. When you make a decision, you can use the data you collected as evidence to make an informed decision.  For instance, in the previous breakfast example, if you hadn’t asked your colleagues for their suggestions for breakfast options, you run the risk of eating a bad breakfast. When you collect data from your colleagues, you can use that information (their response) as a solution to your problem (the fact that you are hungry) in order to make an informed decision that leads to a great breakfast.  The more people you ask, the more data (information) you have in which to make your decision.  Questions to ask:  Take some time to relate this information back to the participants’ own lives. Ask the class how they’ve used data to make their informed decisions in their own lives. As some guiding questions, ask would they rather have a good breakfast or a poor one? If you want to buy a new cell phone, would you rather take your chances on getting a phone that seems like a great choice, only to realize later it doesn’t suit your needs? Gathering information and using it to make an informed decision is a fact of life that may happen on a daily basis for many people.  Once the discussion has finished, Underscore that data are information that is all around us, and that we are all willing or unwilling consumers of data.   * Putting it all together: Data-driven decision-making is the practice of basing decision on the analysis of data rather than using your intuition, guess, or estimate.   Take a moment to ask the class why they think it is important to base decisions on evidence.  Once finished, relate data-driven decision making back to organizations or local governments, saying that using data to make these decisions allows us to improve programs, respond more effectively to organization and community needs,create solutions to pre-existing problems in organizations or communities, and allows the public to use data to make decisions as well  Take a moment to answer any questions from the participants. Then, transition by mentioning that with this greater context and understand of data for decision making, we can move on to , the goals and outcomes for this workshop. |
| *10 mins* | Defining workshop goals and outcomes | * Goals:   The goal of this 3 day workshop is to equip you with ability to understand issues and resources around data and data collection, and learn the importance of using the data to make an informed, evidence-based decision. This workshop is meant to help you and your organization better understand, use, and manage data for decision making.   * Outcomes:   By then end of this workshop you will be able to:   * Use a basic set of processes and questions to assess the data environment at your own organization. * Critically examine data visualization approaches for different audiences. * Understand data lifecycle theory, basic skills when working with data, and the resources needed to support working with data * Be able to create a data project plan that is specific to your own organization.   Take a moment to walk the participants through the structure of the course. Each day will be structured as two, three hour workshops with a one hour break for lunch in between that focus on addressing key outcomes and goals.   * Day 1: Introduction to Data for Decision Making (morning), Assessing your own organization's data culture (afternoon) * Day 2: Building data skills and resources * Day 3: An applied project using your new skills in which you will use the data available to you to solve a problem or assess a need at your organization.   Pause before moving on and allow time for the participants to ask any clarifying questions. |
| *30 mins* | Case Studies:  Armenia, Egypt | Provide context for the participants about how decisions made with data can lead to innovative, evidence based practices for organizations and governments. Take a moment to introduce the following case study to the participants, and provide context.  Guiding questions before beginning: How can data be used to solve a problem?  Are there any examples of governments using data to make decisions that the participants can think of?  What about organizations that use data to make decisions? ?  Many countries around the world are using data to make decision on a daily basis, and are seeking ways to solve existing problems using new forms of data.  As a prior to beginning the case study, provide context by briefly outlining the case study below in Armenia (adapted from UNDP, A guide to Data Innovation for Development, 2016).   * In Armenia, Tourism has great potential for growth. Commonly used data sources such as hotel booking logs and border control records are expensive and time-consuming to collect, and only provide part of the picture. * The UNDP office in Armenia saw an opportunity to help the tourism industry by using a new form of data collection: understanding tourist preferences by analyzing the number of roaming foreign SIM cards in use. * The UNDP team in Armenia used this new form of data collection to track origin and travel patterns for tourists in Armenia, and could then share the data with government decision-makers and local businesses to understand and adapt to shifting tourism trends.   Discussion Questions:  What was the problem the UNDP in Armenia was trying to solve? How did they use data collection to solve i How did they use new data to solve the problem? How can they now use that data to make a decision??  **Myanmar Specific Case (MIMU):**  Provide context by briefly outlining the case study from Myanmar Information Management Unit’s (MIMU) use of cell phone data and GPS coordinates to create and update baseline data on schools across Myanmar:   * In Myanmar, smartphones have become increasingly accessible and popular since foreign investors entered the country in 2014. Most individuals now have a smartphone and a data-enabled SIM card. * MIMU wanted an easier and more accurate way to compile and update baseline data on school locations across Myanmar. Data collection where individuals visit each school and mark the location is time consuming and costly. Accurate and up-to-date baseline data are important to developing more sector-specific data. Information regarding the number and location of schools is also important to help communities identify where available schools are, which schools close, and where more schools are needed. In 2015, MIMU decided to utilize the growth of cell phones and mobile data to facilitate baseline data collection on schools. This was cost effectively, fast, and would allow for data to be collected and updated more frequently. * They created the “School Location Collector” tool, which collects the GPS location of schools. School officials use the service to get the GPS location and save it to the system’s database. They can then send the GPS coordinates over the internet or by SMS.   Now, allow each participant the time to read the following case study. On their own, have them write down their answers to the following questions on a sheet of paper:  What was the problem the UNDP in Colombia was trying to solve? How did they use new forms of data collection to solve it? How can they now use that data for decision making? |

**Improving agricultural yields through data innovation**

*UNDP, A Guide to Data Innovation for Development (2016)*

Climate change is increasingly threatening agriculture around the world, including in Egypt, where the UNDP office set out to address the problem through data, in collaboration with the Faculty of Computers and Information at Cairo University.

The Challenge to farmers, and the lack of data, comes from unexpected shifts in rainfall and planting seasons. Most agricultural producers reply on historic data, which are increasingly unreliable in a shifting climate. But data innovation is offering a new tool.

The team knew that in Colombia, for example, a team of researchers and a local industry organization were able to create a new predictive weather model for rice farmers.

The team used multiple data sources:

* Annual rice surveys
* Harvest monitoring data
* Experiments on rice sowing dates
* Weather data

Using these data, the researchers indemnified geographically specific relationships between climate and agricultural yields. The tool they created successfully forecast a drought, and helped 170 million farmers in Colombia avoid an estimated 3.6 million loss during the drought.

Following the previous experience from the Colombia research, the team in Egypt identified new sources of data that could be used within their own country. They established two sources: data from the sensors network for the Central Laboratory for Agricultural Climate, and the local and international weather station data, a combination that (Based on previous successes) proved to reveal insights for decision makers in the area of water management.

**This example is a hypothetical situation of how data could be used for decision making in Myanmar:**

Natural disasters in Myanmar threaten both crop production and populations in affected areas. Both the cyclone in 2008 and the excessive flooding in 2015 resulted in many deaths, displaced persons, and crop and paddy destruction. The new parliament is working to create new policies to address concerns following future natural disasters. They want to understand if crops and export potential or the displacement of local populations are affected more by these natural disasters. In order to do so, they need to draw from several data sources to evaluate the effect of natural disasters on crop production and on local populations:

- Annual rice surveys

- Monthly exports

- Weather data

- National population data and UN data on internally displaced persons

Using these data, politicians can use past weather data and data from UN agencies to determine which townships were most affected by the cyclone in 2008 and the rainfall and flooding in 2015. They can then analyze the population data published yearly by the Central Statistical Organization to identify if there was a significant population decrease in these townships and increases in neighboring townships. They can also use the population data to identify if populations returned to the affected areas by observing any growth in affected townships. The politicians can then identify if the UN data on internally displaced persons aligns with the affected townships. To evaluate the effect on crop production, politicians can use annual rice surveys to identify any decreases in rice production in the states and regions with affected townships. The politicians can also identify if rice production remained low in these areas in following years, which would indicate more permanent damage to crop fields. They can also use data on the amount of rice exported and changes in the price of rice to identify the effect of natural disasters on the nation’s economy. To identify if natural disasters in Myanmar have a larger effect on populations or crop production, politicians can then compare their findings. In this situation, data collection and sharing are instrumental in helping individuals make informed decisions.

**BREAK: 10 MINUTES**

**Time Action Instructor Notes**

|  |  |  |
| --- | --- | --- |
|  |  |  |
| *35 mins* | Activity 1.1 | Problem Definition Activity  See Activity 1.1 document for more details. |
| *10 mins* | Exploring different forms of data | Explain to the class that data exist in many forms and can be looked at in multiple ways.   * Qualitative vs. Quantitative Data (the following information has been adapted from *MathisFun.com* (2014)   Ask the class if they know the difference between the two types of data. Then, provide the following definitions for each:  Qualitative: information that cannot be expresses as a number or quantified. Qualitative data include descriptive data such as your friend’s favorite holiday, the most common given names in your town, or how to describe the smell of a freshly cooked meal.    Quantitative: information that can be counted or measured.  Application: Show the class a picture of this dog (either on screen or a hard copy)    Ask the class how can you use qualitative data to provide information about the dog? (For example, he has long hair, he is brown and black, he is running, etc.)  Now ask the class how can you use quantitative information to provide information about the dog? (For example, he has 4 legs, he has two ears, he is running 32 kph)   * Primary vs. Secondary Data   Ask the class if they know the difference between primary and secondary data. Then provide the following definitions for each:  Primary: Data that have been collected from an original source for a specific purpose, for example, if a school wanted to know what their students thought of the provided lunch they would question the pupils directly.  Secondary: Data that are not originally collected by a group for specific purpose, for example, finding out the average age of a population by using national survey data. |
| *15 mins* | Introducing key concepts | The following are ways that data can be applied. Similar to the format above, pause and ask the class for definitions of each word before providing the definitions in this instructor’s guide.  Dataset: Data selected and arranged in rows and columns. Has the class seen datasets before? What does their datasets look like? How are these datasets stored?  Take time to pass around several sample datasets. In groups of 2-3, have the participants look at each dataset. What information is provided? Is it qualitative or quantitative? What does this information show? Who can use it/what is it for?  Documentation: documents that provide proof or evidence of something, or are a record of something. What are some examples of documentation in the participants’ everyday lives? If the class is quiet, feel free to provide some sample answers (i.e. a passport, birth certificate, proof of residency, etc.)  Data visualization: a way to help people understand the significance of data by placing them in a visual context. Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns.  Pass around example visualized charts or project these images on the screen:    *Source: East Africa Business Daily, 2017*    *Source: T. Patil*  For each visualization, ask the class What information is provided? Is it qualitative or quantitative? What does this information show? Who can use it/what is it for?  Stakeholders: a person with an interest or concern in something. For example, who are the stakeholders in this class? Who are the stakeholders at a restaurant? Who are the stakeholders at a football match? |
| *35 mins* | Activity 1.2 | Understanding and Working with Data Visualizations  See Activity 1.2 document for more details |
| *5 mins* | Debrief | Briefly review key concepts identified, pause to answer any questions, and dismiss the class for lunch. |