GENDER 101 COURSE

Module 1: Fundamentals of Gender Data

Overview

This module's goal is to help you build a foundational understanding of both of gender and data as we begin to navigate gender data.

Gender Basics

In this section, we will establish a shared vocabulary for the course and answer questions such as: What is gender? What is sex?

Data Basics

In this section, we will cover the basics of a data lifecycle and introduce our data lifecycle model for this course. We'll also answer questions like: What is qualitative versus quantitative data? With the purpose of establishing a common data vocabulary.

Building Alliances for Gender Equality

In this section, you'll find a summary of Ladysmith Collective's 2020 report on gender data and the Sustainable Development Goals.

Introduction to Case Study

In this section, we will introduce you to a gender data case study in order to frame this course. You will also meet two personas who will join us on our Gender Data 101 journey. We will periodically return to our case study and personas during our discussions and activities.

Activities

Asks & Offers Board
We've set up an "Asks & Offers" board to help you connect. Ask or offer resources, data, a consultation expertise, and more.

Data Lifecycle Discussion
Compare and contrast the two data lifecycles and consider the data lifecycle of your organization. If you were to make your own data lifecycle, what might you emphasize, add, or remove?

Open gender data scavenger hunt
Explore an example of a public dataset with sex-disaggregated data.

Gender and Language discussion
Discover how language affects how we perceive the world.

Sourced Participant Resources

This section provides a place for you to share resources that you have used or created to help you when it comes to gender, data, and gender data. Click here to add resources you would like to share with other participants! There is no limit to how many you can add. You can also review those submitted by others.

Gender Basics

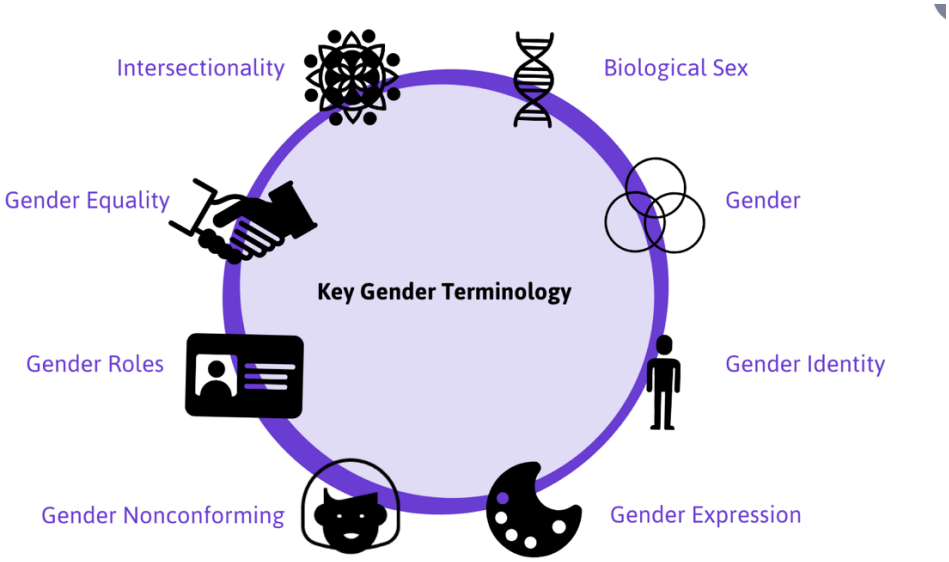
Establishing a Common Vocabulary

The term 'gender', is not synonymous with the word 'women'. Gender goes beyond exclusively women and women's issues. Gender is a set of socially constructed roles, behaviors, activities, and attributes that a given society considers appropriate. These constructions specifically surround the ideas of masculinity and femininity. Therefore, when looking at gender, we have to remember it is not simply equivalent to a person's biological sex. Biological sex refers to anatomical, physiological, genetic, or physical attributes that determine if a person is male, female, or intersex. Gender goes beyond

biological sex and examines the individual's societal role, responsibilities, and expectations.

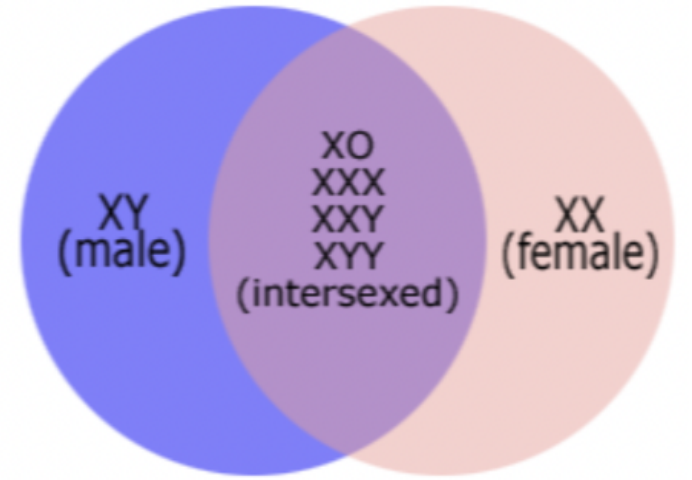
Furthermore, we live in a complex world! No one factor is isolated from other factors. Other identifying factors (religion, race, sexual orientation) interact with gender. The interaction between gender, sexuality, race, socio-economic status, ability, and other categories of social difference is known as intersectionality.

Throughout the course, we will be using the terms 'gender', 'sex', and 'intersectionality' frequently in addition to building upon the following key terms below.

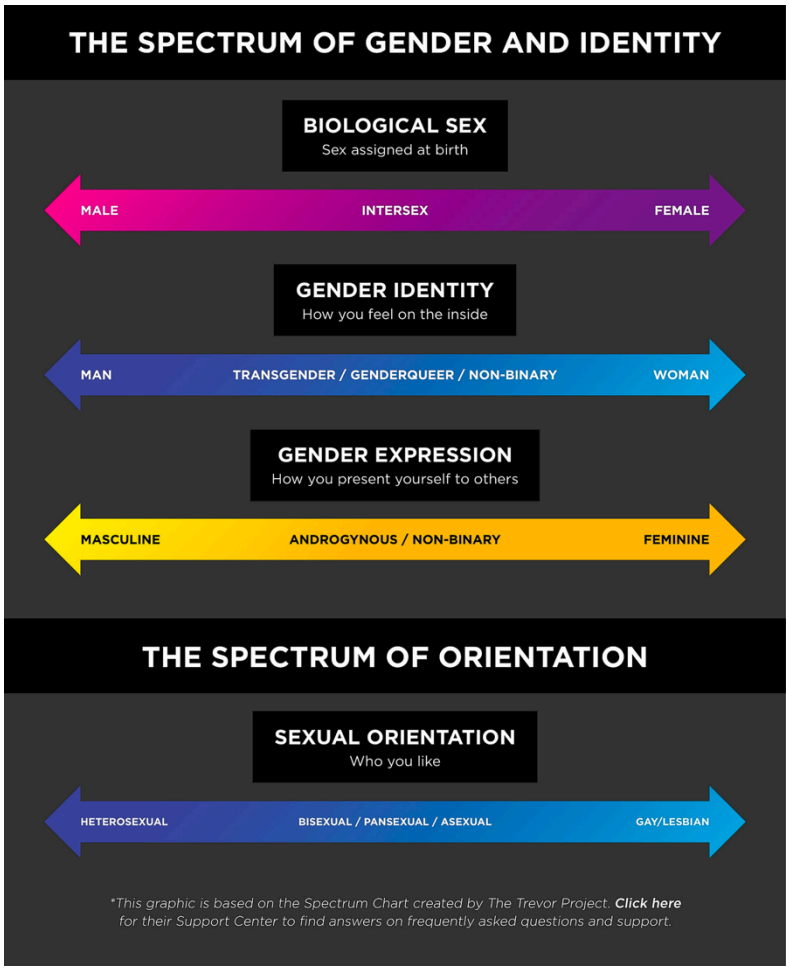


Understanding Spectrums

It's also important to understand that many of the terms introduced exist on a spectrum - not a binary. For example, we're socialized to think of biological sex as either male (XY chromosomes, shown below) or female (XX chromosomes). When in fact, biological sex exists on a spectrum. It's possible for a human to have anywhere from one to three X chromosomes! Intersex is the umbrella term to describe these variations (see the diagram below).



Gender identity, gender expression, and sexual orientation exist on spectrums as well. The interaction of all four of these spectrums make up parts of our identity. Therefore, someone who is intersex (biological sex) could identify as a pan-sexual (sexual orientation), androgynous (gender expression), man (gender identity).



Visualizing Sex as Spectrum. In 2017, the Scientific American put together a diagram of how sex determination can be visualized as a spectrum, rather than a simple XX female and XY male. [Please check out this complex diagram here.]

Gender and Language

Why is establishing a common vocabulary important?

Why did we spend a full sub-module on terms? You may be wondering this, especially if you may already have a basic idea of what gender and sex are. Language is a critically important way of how we understand concepts. Gender Data 101 Edition 2.0 is a global course. We each bring our own unique experiences and perspectives into this community, including our understanding of language and the languages we speak. Language shapes how we view femininity and masculinity. It can also be a key way to break down our own understanding of gender biases, including the gender binary. The gender binary is a system of gender classification in which all people are categorized as being either male or female.

Excerpt from Lera Boroditsky, Cognitive Science Professor at the University of California, San Diego:

"...One study that I love is a study that asked monolingual speakers of Italian and German and also bilingual speakers of Italian and German to give reasons for why things are the grammatical genders that they are. And so for example, if the word chair is masculine in your language, why is that? And if you're a monolingual speaker of one of these languages, you're very likely to say that the word chair is masculine because chairs are, in fact, masculine.

Right, so these speakers have internalized this idea from their language, and they believe that it's right. They believe that their language reflects the true structure of the world. But if you ask bilinguals, who have learned two languages, and now they know that some genders disagree across the two languages, they're much less likely to say that it's because chairs are intrinsically masculine. They're more likely to say, well, it's a formal property of the language. They're more likely to see through this little game that language has played on them..."

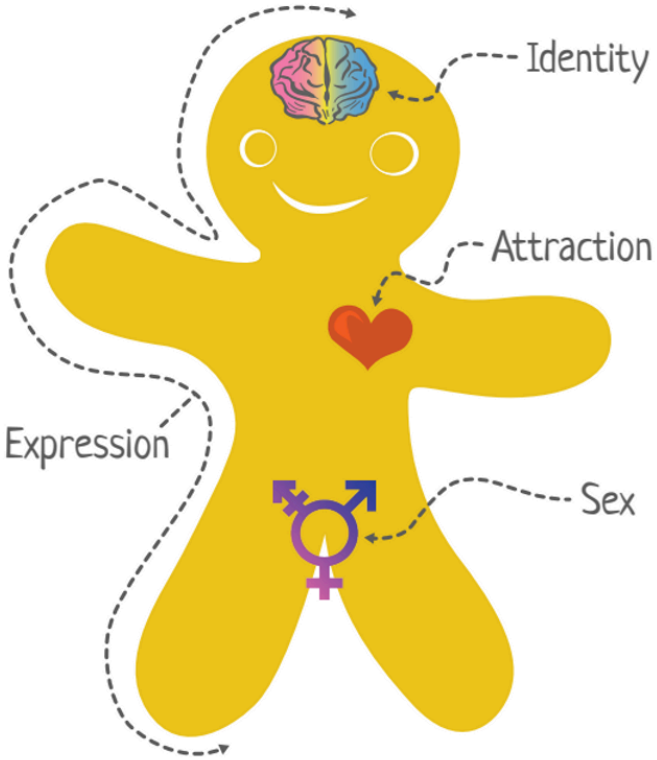
Optional: Explore YOUR Gender

Gender is also a personal topic. In this course, we would like to provide you the opportunity to explore your own personal understanding of gender. Below are some exercises you may consider completing.

What does gender mean to you?

The Genderbread Person

The Genderbread Person is a teaching tool for breaking the big concept of gender down into bite-sized, digestible pieces. This resource has evolved from version 1 to version 4. You are welcome to download the Genderbread Person and fill-out the worksheet for yourself. More information can be found on the Genderbread website.



The Genderbread Person Version 1 Limitations

Spectrum vs. Continua. This version relies on a spectrum model of understanding gender, placing two aspects of identity on opposite ends of one scale. This pushes one to (1) see themselves as somewhere within those two aspects; (2) see them as opposing ideas; and (3) if they embody both, to see themselves as in between them, instead of as a lot of both, or a little of both. This issue is cleared up with version 2, which introduced the “-ness” concept.

Labels. The language describing the different points of the spectrums erases folks who don’t identify with any of the concepts presented (e.g., someone who isn’t bisexual, but asexual). Also, it reinforces an unhelpful oversimplification that the middle term exists directly between the two at the end (e.g., that “genderqueer” is a term for someone “between woman and man”).

Download Genderbread person version 4 (.pdf)

Data Basics

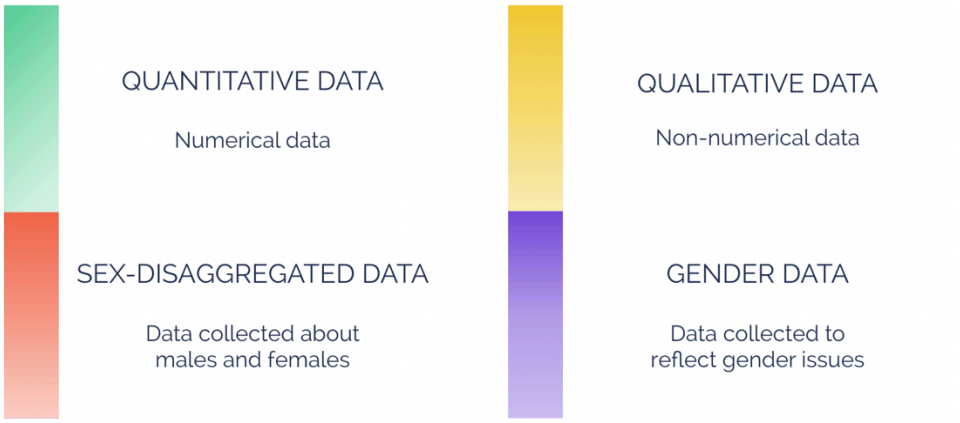
Establishing a Common Vocabulary

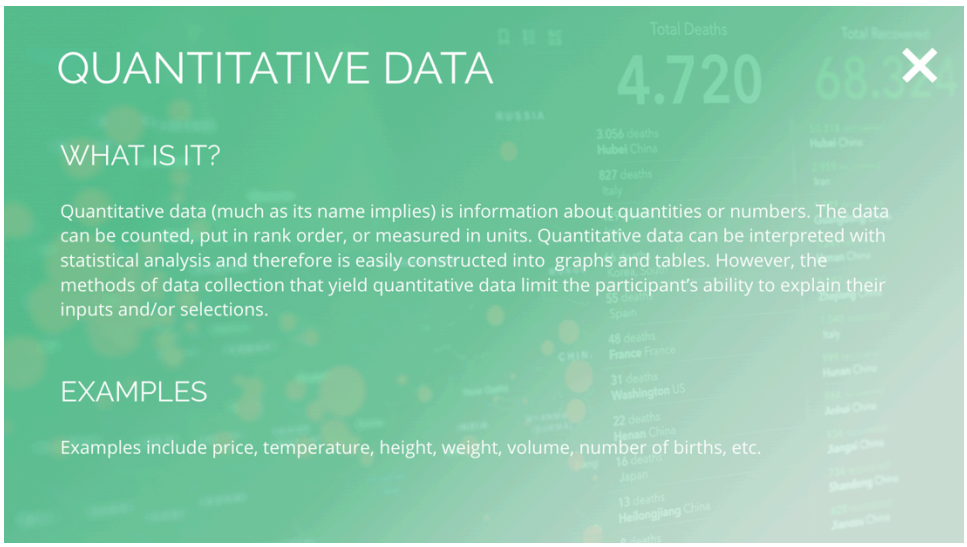
Data is simply any fact or figure that can be collected for reference and analysis. The term 'data', often conjures visions of spreadsheets with endless rows and columns of inputs and numbers. Sometimes (depending on where you are in the data life-cycle), these spreadsheets contain raw data. Raw data (also known as dirty data or unclean data) is data that's directly from the source, often unprocessed, and therefore not immediately valid or ready to be analyzed or applied.

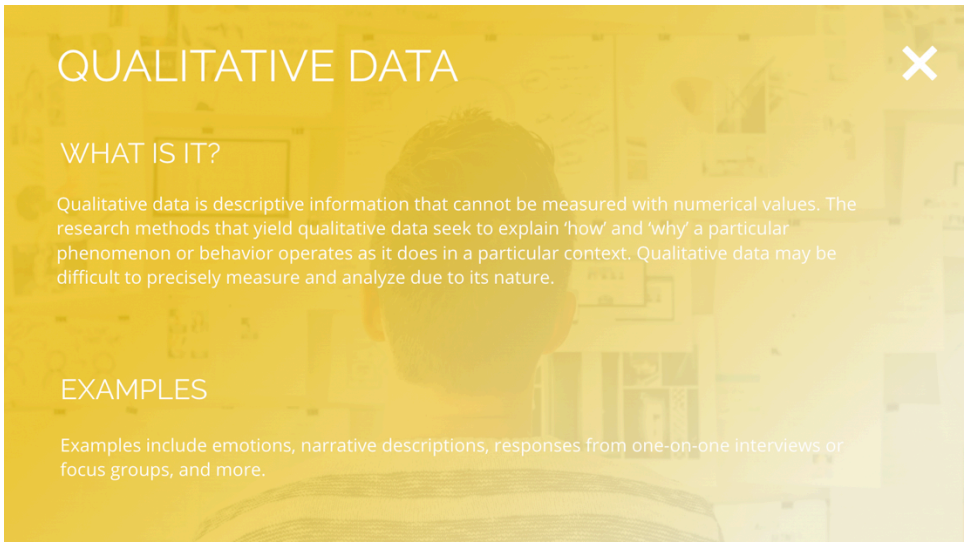
Throughout this course, you will learn about data and the stages it must go through in it's 'life-cycle' order to be understood and applied. You will also learn why capturing gender data accurately and thoughtfully is so important.

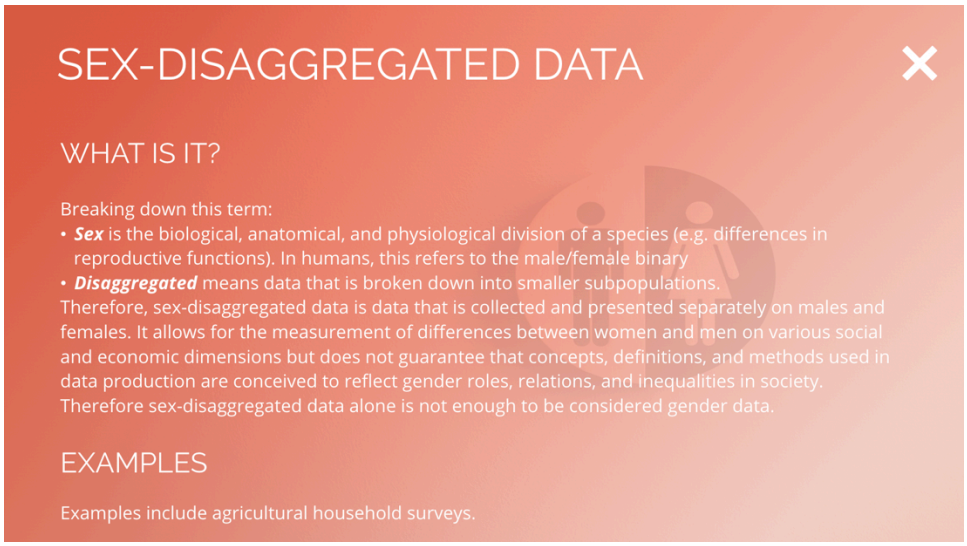
As we continue in the course, we will be building upon the foundation of the following key terms.

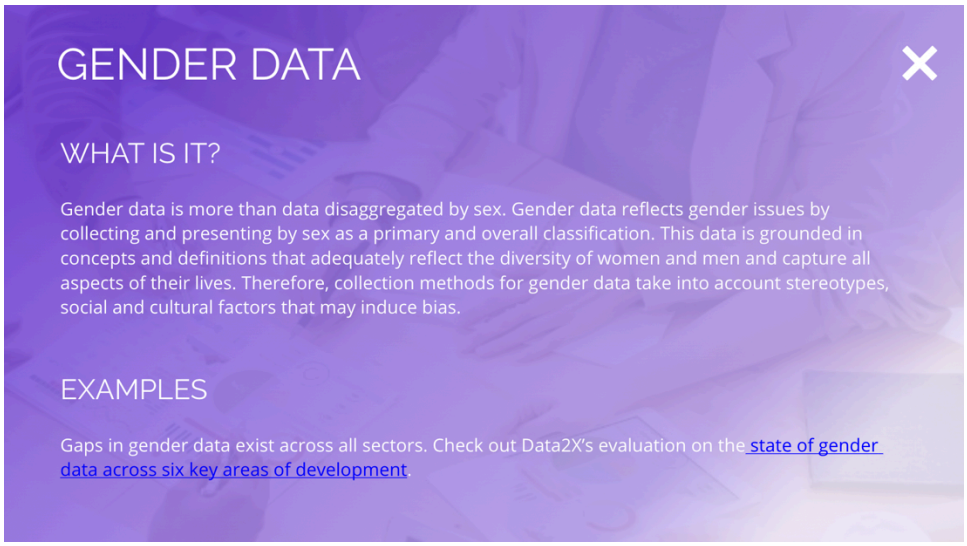
Key Data Terminology






QUALITATIVE DATA
WHAT IS IT?
Qualitative data is descriptive information that cannot be measured with numerical values. The research methods that yield qualitative data seek to explain 'how' and 'why' a particular phenomenon or behavior operates as it does in a particular context. Qualitative data may be difficult to precisely measure and analyze due to its nature.
EXAMPLES
Examples include emotions, narrative descriptions, responses from one-on-one interviews or focus groups, and more.


SEX-DISAGGREGATED DATA
WHAT IS IT?
Breaking down this term:
• Sex is the biological, anatomical, and physiological division of a species (e.g. differences in reproductive functions). In humans, this refers to the male/female binary
• Disaggregated means data that is broken down into smaller subpopulations.
Therefore, sex-disaggregated data is data that is collected and presented separately on males and females. It allows for the measurement of differences between women and men on various social and economic dimensions but does not guarantee that concepts, definitions, and methods used in data production are conceived to reflect gender roles, relations, and inequalities in society. Therefore sex-disaggregated data alone is not enough to be considered gender data.
EXAMPLES
Examples include agricultural household surveys.



https://data2x.org/where-are-the-gaps/

Take this poll!

In your day-to-day, what data do you work with the most?

- ☐ Qualitative data
- ☐ Quantitative data
- ☐ Both

Data Life Cycle

A data lifecycle is simply the process that data goes through from its inception (when it is first collected) to its application (when its shared, analyzed, visualized, etc). While various organizations may have slightly differing data lifecycles, each share key overlapping stages. Data lifecycles typically include a collection phase, processing phase, and using phase. "Using" can include any sort of sharing, analysis, and/or visualization. "Using" may even include uptake and impact!

Below are two data lifecycles, from the GovLab and Data2X, respectively. Compare and contrast the two data lifecycles below. If you were to make your own data lifecycle, what might you emphasize, add, or remove? What might be some ways to thread gender considerations into or throughout the data lifecycle? What is the data lifecycle for your organization and how does it compare to either of these? If you don't know, this is a good time to find out! We will answer these questions together in our discussion activity.

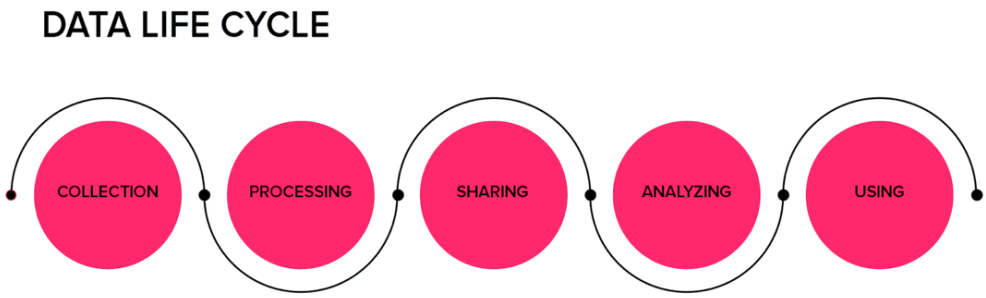
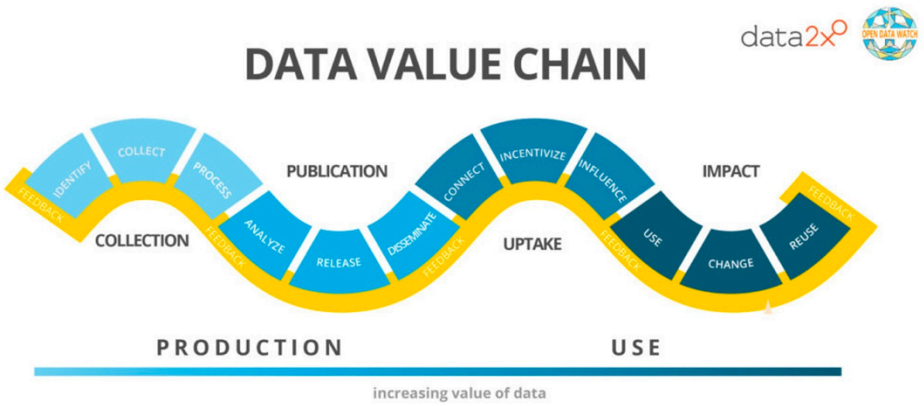


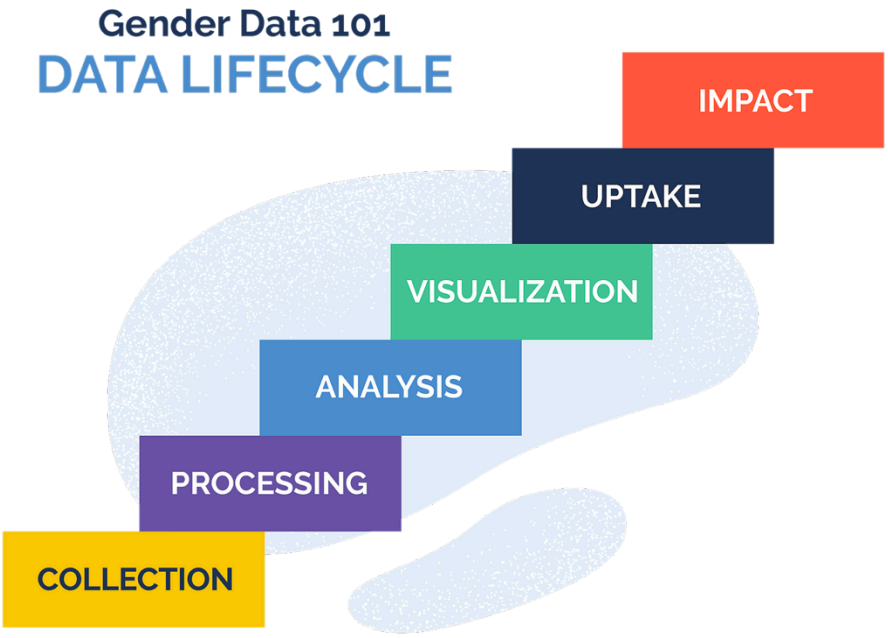
Figure 2: Data Life (The GovLab)



Gender Data 101's Data Lifecycle

As you can see from above, there is no one-size-fits-all data lifecycle. For this course we have created our own data lifecycle. This lifecycle was influenced by the two shown above from the GovLab and Data2X, respectively.

In summary, as we move forward, we will be referring to the data lifecycle with these steps.



Building Alliances for Gender Equality

Ladysmith's Report: Building Alliances for Gender Equality

In March 2020, Ladysmith Collective, a feminist research consultancy, released the report Building Alliances for Gender Equality: How the Tech Community Can Strengthen the Gender Data Ecosystem.

In this report, Ladysmith built a bridge between two communities: gender development and tech. In order to establish your gender data foundation, we strongly encourage reading this report. Though if you're pressed for time, we've summarize some key takeaways from the report below.

What is Gender Data?

Gender data does not simply mean data about women. Recall, that gender is the socially constructed roles of women and men, boys and girls. Sex refers to biological differences. Gender data is not the same as sex-disaggregated data. Gender data can be both quantitative and qualitative. High quality gender data requires the understanding of how to address embedded gender biases in the entire research process.

For example, on the quantitative side, gender data can be sex-disaggregated statistics on the use of paternity leave benefits. On the qualitative side, it can be the data from interviews and participant observation with mothers and fathers about why paternity benefits are or aren't used.

Additional Resource

For more information on why gender data matters, read Data2X's embedded infographic. Additionally, Data2X made a commitment to Gender Data. Read more on Data2X's site.

Sustainable Development Goals

In 2015, the United Nations adopted The 2030 Agenda for Sustainable Development. In this agenda, there are 17 Sustainable Development Goals, commonly referred to as SDGs. Interact below to familiarize yourself with the SDGs. Click on the flag in the bottom left corner to link to additional information and resources.

Gender Data Gaps in the Development Community

Ladysmith's report focused on the gender data gaps that the tech sector could potentially have the best impact on. Below are the gaps highlighted from that report.

- **Family Life -** Household data is difficult to capture. However, there is an opportunity for social network data to shed light on how family's share resources and expend unpaid labor. Family life is captured in SDGs 1, 3, 5, and 8.
- **Last-mile Healthcare -** In order to aid the world's most isolated communities, there needs to be more data in how these communities engage with healthcare. Last-mile healthcare is captured in SDGs 1, 3, and 5.
- **'Hidden' Economies -** There are current data gaps on women and men's informal employment. 'Hidden' economies is captured in SDGs 1, 8, 9, and 16.
- **Climate Change and Migration -** Women and men have different migration patterns. Gender data is critical to inform climate action and rights-based policies. Climate change and migration is captured in SDGs 8, 9, 13, 14, and 15.
- **Gender-based Violence and Discrimination Online -** How can we best ensure women and girls' rights online? Gender-based violence and discrimination online is captured in SDGs 3, 5, and 16.
- **STEM 'School to Power' Pipeline -** Data is needed on when and why girls in STEM fields drop out or are pushed out. STEM 'school to power' pipeline is captured in SDGs 4, 8, 9, and 17.

Take this poll!

Which of these gender data gaps most closely relates to the work you do?

- ☐ Climate Change and Migration
- ☐ Family Life
- ☐ Last-mile Healthcare

- ☐ 'Hidden' Economies
- ☐ Gender-based Violence and Discrimination Online
- ☐ STEM 'School to Power' Pipeline

Avenues for the Tech Community to Strengthen the Gender Data Ecosystem

Ladysmith's report summarizes that the tech community should consider the following:

- ☐ Share privacy-preserved data
- ☐ Process and visualize existing external data for new insights
- ☐ Innovate new data products and tools

Optional: Please rank the elements by level of priority and/or need for you and your organization. (Top - Most Important, Bottom - Least Important)

- ☐ Process and visualize existing external data for new insights
- ☐ Share privacy-preserved data
- ☐ Strengthen the capacity of feminist and women's organizations
- ☐ Innovate new data products and tools

Gender Data Case Study: Women and Global Vaccine Distribution

Throughout Gender Data 101, we will engage with real-world data in our course material and activities. For the purposes of this course, we will focus on women and

health. Specifically, we will examine women and the global COVID-19 vaccine distribution

The COVID-19 pandemic did not affect everyone equally. The vaccine response to the COVID-19 pandemic has been far from equal. "Disease outbreaks affect women and men differently...[they make] existing inequalities for women and girls and discrimination of other marginalized groups...worse" (Source: UNFPA). It is important to be aware of many factors as we analyze COVID-19 data. For example, it is not enough to simply take sex-disaggregated data at face value. The underlying gender factors that may affect new datasets need to be considered and analyzed. This is imperative as the next steps of widespread access to treatment and vaccinations are created.

Underlying Gender Factors

There are many underlying gender factors to a global pandemic that must be analyzed in order to draw appropriate conclusions from the data. These factors go beyond sex and physiology. Some examples include:

- Women are more likely to be at the forefront of community care efforts. A woman's gender role places them "in a prime position to identify trends at the local level that might signal the start of an outbreak and thus improve global health security" (Source: The Lancet).
- Women, particularly in lower and middle income countries, including in Zambia, do not have access to tests and many of them lack knowledge and awareness of testing. This is particularly the case in remote and rural areas. Empowered women community health workers can play a huge role in scaling testing to ensure it is available to everyone (Source: WHO).
- Nine out of 10 people living in the poorest countries are set to miss out on a vaccine this year. Production delays put even this figure in doubt. Unjustifiably high prices block access and threaten to push more countries into an ever-deeper debt crisis. If we continue to pursue the vaccine model we have, we will fail to get this pandemic under control for years to come. (Source: UNAids.org)

- Stay-at-home orders may have profound effects on gender-based violence. In China's Hubei province, "the number of domestic violence cases...has nearly tripled in February after many people were quarantined in January due to the virus" (Source: Time and Axios).
- Single parents have been hit dramatically by COVID-19. "...There are 20 million single parents in the US, three-quarters of whom are women" (Source: BBC).
- "COVID-19's total burden of disease extends beyond those who get sick, and this has potentially deadly consequences for women and girls" (Source: Development Pathways)

Two Personas: Health and Advocacy

Meet Dr. Carolina Lopez (Goals drawn from The Lancet)

Dr. Carolina Lopez is on the front lines of the COVID-19 vaccine response. She is particularly concerned at the disease and its vaccine distribution's gendered issues. Her goals include:

- Ensure availability of sex-disaggregated data, including on differing rates of infection and vaccination, differentials in vaccine rates, differential care burden, and incidence of domestic violence and sexual abuse
- Ensure equal voice for women in decision making in the response and long-term impact planning
- Ensure that public health messages properly target women including those most marginalized

Meet Rumy

Rumy is a program manager at a non-profit focused on women's health. She has a background in advocacy and gender studies. However, she is new at data analysis. She will be consulting Dr. Lopez as she leverages her organization's gender data. Her goals include:

- Understand the importance of data and how it relates to her organization's efforts
- Limiting biases in data collection
- Utilizing data tools
- Implementing a sustainable vaccine distribution awareness program to aid women during the COVID-19 crisis informed by her data analysis

Activities

Asks and Offers Board

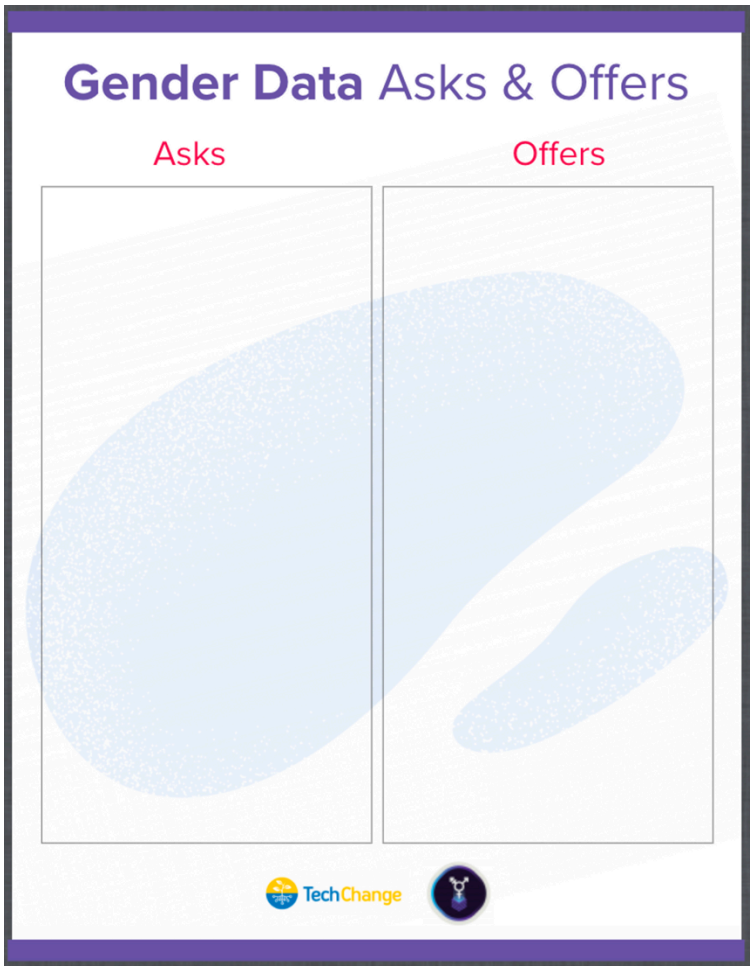
There is a wide variety of participants enrolled in this Gender Data 101 course. We recognize that with this unique mixture of gender experts, data experts, program specialists, field practitioners, and more there's a lot of opportunities for collaboration. This is your opportunity to connect with others also enrolled in the course and grow as professionals in your respective fields (gender or data) as we concurrently grow in gender data!

We've set up an "Asks & Offers" board to help you connect. Ask or offer materials/resources, data, consultation expertise, and more!

Here are some best practices for interacting with the Asks & Offers board:

- Be specific and relevant. Providing context when necessary makes it easier for others to follow up with you.
- Don't be afraid to make an ask! If it's on your mind, it's probably on several other people's radar as well.
- It's a community of give and take. While it's tempting to just post an ask, this is not a help desk board. It's an asks and offers board. Try to add an offer for every ask or two.

- Check back often!



Activity: Data lifecycle discussion

In the Data Basics submodule, we discussed how a data lifecycle is simply the process that data goes through from its inception (when it is first collected) to its application

(when its shared, analyzed, visualized, etc). While various organizations may have slightly differing data lifecycles, each share key overlapping stages.

Below are two data lifecycles, from the GovLab and Data2X, respectively. Compare and contrast the two data lifecycles and respond to this discussion thread with your responses to the following questions.

- If you were to make your own data lifecycle, what might you emphasize, add, or remove?
- What might be some ways to thread gender considerations into or throughout the data lifecycle?
- What is the data lifecycle for your organization and how does it compare to either of these? If you don't know, this is a good time to find out!
- Bonus: If you are able, then please upload a photo or sketch of your organization's data lifecycle to the discussion thread.

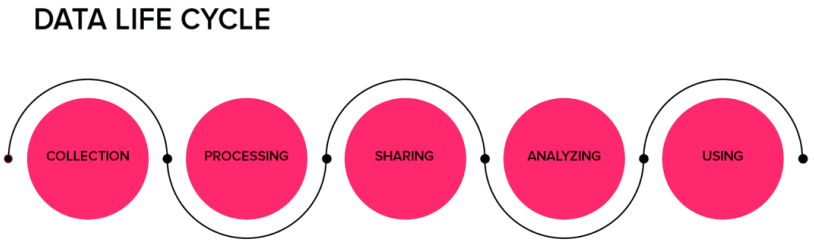
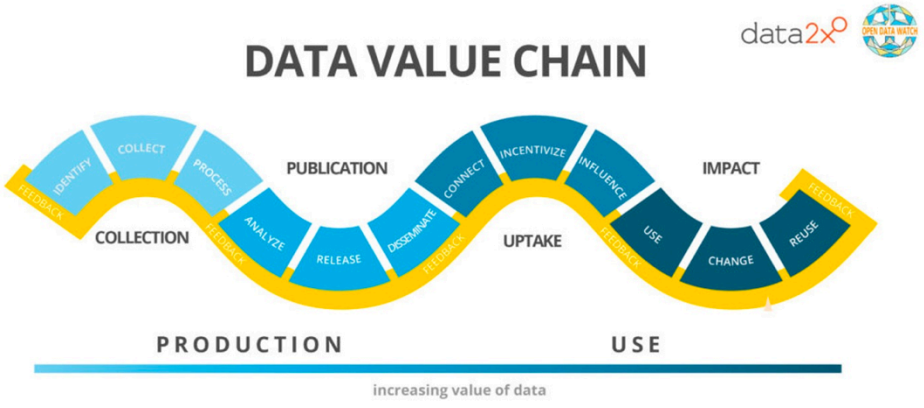


Figure 2: Data Life (The GovLab)



Activity: Open gender data scavenger hunt

We've compiled a short list of open sex-disaggregated data from various sources. The goal of this activity is to select a dataset, briefly explore it, and answer a few questions. Afterwards, you will engage in a discussion with the rest of the class.

Step 1: Choose a dataset

If you have a sex-disaggregated dataset already in mind, then please use that one and move on to Step 2.

Choose a dataset to explore:

- Survey on Gender Equality at Home - This survey covers topics about gender dynamics and norms, unpaid caregiving, and life during the COVID-19 pandemic.
- REACH-UNICEF WASH Household Baseline Survey - Supported by UNCEF, REACH collected household-level data on key WASH indicators across all 35 recognized ISCG camps between 1-22 April (2019).
- COVID-19 Vaccine Tracker - The biggest vaccination campaign in history is underway. More than 6.23 billion doses have been administered across 184 countries, according to data collected by Bloomberg. The latest rate was roughly 27.8 million doses a day.
- COVID-19 Sex-Disaggregated Data Tracker - In collaboration with CNN, Global Health 50/50 began compiling publicly available sex-disaggregated data reported by national governments to date and is exploring how gender may be driving the higher proportion of reported deaths in men among confirmed cases so far.
- UNICEF Youth Literacy Rate for 14-24 Years - Number of people age 14 to 24 years who can both read and write with understanding a short simple statement on their everyday life, divided by the population in that age group.

If you would like to find another dataset that is not listed, then please refer to the 'additional resources' section.

Step 2: Explore your dataset

Set Up: In order to explore your dataset, you should download the Excel (.xlsx) version of your file. If you do not have access to Excel, then upload your data to Google Drive and open as a Google Sheet.

Exploration: Some things to look for when you explore your dataset.

- Size of your dataset file (KBs, MBs, etc)
- Dataset description
- Date last updated
- Column names
- Key or legend
- Source

Step 3: Answer some questions

- Does the dataset have quantitative or qualitative data? Or both?
- Does the dataset take an intersectional approach? If so, to what extent?
- Would you consider your dataset to be "gender data"? Why or why not?
- Considering your answers to questions 2 and 3, what does the dataset do well and what is it missing?
- What are additional questions or comments you have for the dataset? For example: What methodologies were used to collect this data?

Step 4: Join the discussion

In this week's discussion, fill out a brief summary of the dataset you explored and the questions you answered. Engage with your classmates on your challenges and insights.

- What are the gender options?: (Example Answer 1: Gender options are binary, only M/F, Example Answer 2: Gender options are Male, Female, Other, etc.)

- Does the dataset have quantitative or qualitative data? Or both?
- Does the dataset take an intersectional approach? If so, to what extent?
- Would you consider your dataset to be "gender data"? Why or why not?
- Considering your answers to questions 2 and 3, what does the dataset do well and what is it missing?
- What additional questions or comments you have for the dataset? For example: What methodologies were used to collect this data?

Additional Resources

If you would like to do this scavenger hunt with your own dataset or find a different one, then feel free to use the resources below.

- Humanitarian Data Exchange
- The World Bank's Gender Data Portal
- Kaggle
- UNData
- Pew Research Center

Additionally, feel free to check out this list of 100 free data sources (2019)

Activity: Gender and Language discussion

Recall, the Gender and Language portion of Gender Basics. For this activity, follow these steps:

Listen to minute 16:41 - 29:29 of the Hidden Brain podcast. Answer the following questions:

- What language(s) do you speak?
- Give a specific example of how your language does or does not utilize gendered language OR Give an observation of how your language expresses gender.

- What question or comments do you have about the podcast material?