MODULE 2

GENDER DATA LITERACY AND AVOIDING COMMON MISTAKES

TRAINING SYLLABUS

Curriculum on Gender Statistics Training

This product was developed under the guidance of the Subgroup on Gender Statistics Training, within the Asia-Pacific Network of Statistical Training Institutes.

Introduction

This syllabus has been designed to guide trainers on how to conduct related training. The syllabus can also be used by learners who wish to know more about this topic and people who are generally interested in gender statistics.

This syllabus is part of a wider module on this area of gender statistics. Other materials within this module might include exercises, PowerPoint presentations and example quizzes. Please refer to the additional set of materials for a comprehensive and effective learning experience.

Who is this module for?

- Statisticians that wish to understand the specificities around select areas of gender statistics, such as violence against women and time use (for expert statisticians, however, it is recommended to skip through the initial part of the module, as some of the content might already be known)
- Policymakers and decision-makers who are looking to enhance their understanding and use of gender data for evidence-based decision-making
- Academics who wish to focus or inform their research through the use of gender data
- Civil society organizations that wish to enhance their use of gender data for advocacy or communication purposes
- Media personnel interested in integrating gender data into their media products, and presenting a more accurate and comprehensive picture
- Anyone who wishes to find out how to use gender data

What do I need to know before going through this module?

This is an introductory module on gender statistics, targeted to non-experts in the area of statistics. No advanced knowledge of statistics is necessary. However, it would be good for the learner to have an idea of what the Sustainable Development Goals (SDGs)[1] are, including their targets and indicators[2]. It is also recommended for learners to have gone through Module 1 and understand the definitions of gender, sex and gender indicators.

[1] For additional information on the SDGs see: https://www.un.org/sustainabledevelopment/sustainable-development-goals/

[2] See: https://unstats.un.org/sdgs/indicators/indicators-list/

Learning objectives

The expected learning outcomes for this module include:

- After going through this module, the learner is expected to become familiar with basic concepts of data and statistics and the proper use of semantics of statistics.
- The module also provides an introduction to key concepts of statistics from a gender data perspective. Therefore, the learner is expected to gain knowledge on specific issues of gender data, such as time-use, violence and crime data, etc.
- Finally, trainees will be introduced to the issue of misinterpretation of data and how to avoid it.

Note to trainers: Depending on the pace of the trainer and trainees, it is expected that training for this module can be delivered in 30 minutes to 1 hour.

Table of content

1. Understanding definitions and key concepts in the area of gender statistics

1.1 Official vs. non-official statistics

Official statistics are statistics produced either by the National Statistics Office (NSO) or another government body in charge of data production (e.g. line ministries, Central Bank, National Meteorology Agency, etc.). They are usually produced in accordance with the National Statistics Law/Act and in line with the Fundamental Principles of Official Statistics[3]. In select cases, official statistics might also be produced by third-party organizations, such as private sector entities, civil society organizations or academic institutions, with the involvement of the NSO or other National Statistical Authority. It is important to note that, in these cases, the NSO's involvement and validation of such statistics is essential for the figures to be treated as 'official'.

Some examples of official statistics include:

- Figures derived from Census data (e.g. population count)
- Estimates derived from official surveys (e.g. employment and unemployment rates from labour force surveys)
- Aggregates calculated using administrative records compiled by government institutions (e.g. birth registration)
- Increasingly, select official statistics are also starting to be produced using nonconventional sources (e.g. big data, crowdsourcing, etc.). Although these data sources have great potential, it is important to examine their comprehensiveness in terms of coverage to avoid bias.

Non-official statistics are those produced without any involvement of the National Statistical Office or any other member of the National Statistical System. These statistics, when derived from surveys, are often narrower in coverage, as sample sizes tend to be larger in data collection exercises conducted by National Statistical Authorities (due to availability of financial and human resources for data collection). While most official statistics tend to be produced periodically (e.g. the Census often takes place once a decade, demographic surveys often have a five-year periodicity), non-official statistics are more likely to be ad-hoc studies and one-off data collection experiments.

When to select official vs. non-official statistics?

Official statistics are almost always preferable over non-official statistics, as they are usually more comprehensive, and periodicity tends to be more frequent. This is often the case because:

- National Statistical Systems have larger amounts of financial resources allocated to data collection.
- Official data producers are able to make use of Census data for sampling purposes,

[3] UNSD https://unstats.un.org/unsd/dnss/gp/fp-english.pdf

resulting in more accurate estimates.

- NSOs and other producers within NSSs, for whom data production is a core responsibility, are able to access and train large numbers of enumerators, who are well prepared to collect data. Larger teams of enumerators also mean that the chances of them speaking the local language of respondents are higher, and therefore the data compiled might be more reliable than that compiled through non-official statistics.
- Most official statistics tend to be produced periodically, as funds are allocated accordingly by national governments.

There are two instances, however, when non-official statistics might be preferable: when the user is looking for data on a particular topic that might not be available through official statistics, and when there might be a conflict of interest in the official statistics (e.g. statistics about government corruption, good governance, etc.).

1.2 Metadata

Metadata refers to the range of information, generally textual, that fosters understanding of the context in which statistical data have been collected, processed and analyzed with the objective of creating statistical information (legal and regulatory texts, methods and concepts used at all levels of information processing, definitions and nomenclatures, etc.)[4]. In other words, metadata is information about data. Metadata might provide information about an indicator, a data series or a data point. Generally, for SDG monitoring, two types of metadata are used: indicator/series metadata, and data point metadata.

Let's focus on indicator/series metadata first.

Indicator metadata: What does it usually include?

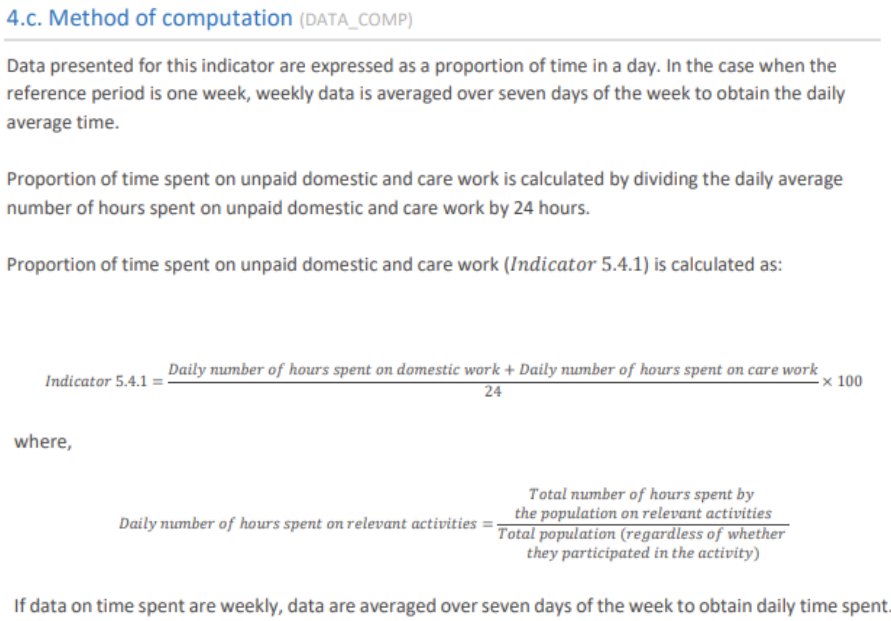
- Official indicator name
- Definitions
- Rationale
- Methods of computation / Formulas
- Information about exceptions, methodological concerns and limitations
- Information about usual data sources utilized to derive the indicator
- If the metadata refers to an SDG indicator, it often also includes information about custodian agencies and methodology for the production of regional aggregates.

Indicator metadata: Where can you find it?

- In on-line repositories (e.g. https://unstats.un.org/sdgs/metadata/)
- In indicator handbooks, normally developed to accompany indicator sets. For example, an excerpt of the metadata for indicator 5.4.1 is given below. It elaborates on the method of calculating the estimates.

[4] https://www.afdb.org/fileadmin/uploads/afdb/Images/Photos/eng-charte.pdf

Figure 1: Excerpt of metadata for SDG Indicator 5.4.1



4.c. Method of computation (DATA_COMP)

Data presented for this indicator are expressed as a proportion of time in a day. In the case when the reference period is one week, weekly data is averaged over seven days of the week to obtain the daily average time.

Proportion of time spent on unpaid domestic and care work is calculated by dividing the daily average number of hours spent on unpaid domestic and care work by 24 hours.

Proportion of time spent on unpaid domestic and care work ($Indicator$ 5.4.1) is calculated as:

$$Indicator\ 5.4.1 = \frac{Daily\ number\ of\ hours\ spent\ on\ domestic\ work + Daily\ number\ of\ hours\ spent\ on\ care\ work}{24} \times 100$$

where,

$$Daily\ number\ of\ hours\ spent\ on\ relevant\ activities = \frac{Total\ number\ of\ hours\ spent\ by\ the\ population\ on\ relevant\ activities}{Total\ population\ (regardless\ of\ whether\ they\ participated\ in\ the\ activity)}$$

If data on time spent are weekly, data are averaged over seven days of the week to obtain daily time spent.

Data point metadata: What does it include?

Information about specific datapoints. Such information often includes explanations about exceptions, coverage, methodological limitations and specific details about one particular data point.

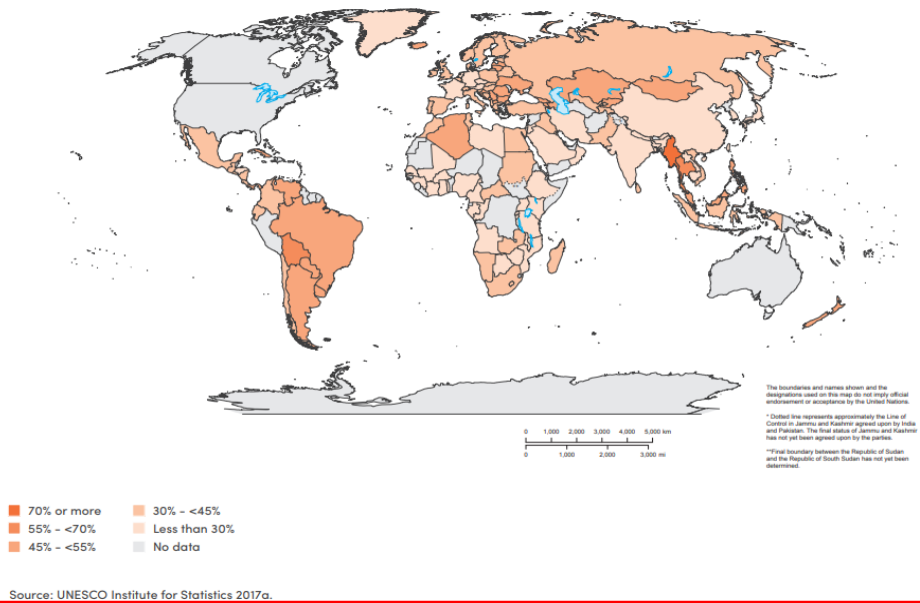
Data point metadata: Where can you find it?

Data point metadata is often found in the form of footnotes, alongside data tables or in data cells. Below is an excerpt of the data point metadata for the share of female researchers, by country, for the years 1999–2015[5].

Figure 2: Example of data point metadata

[5] UN Women http://www.onumulheres.org.br/wp-content/uploads/2018/02/SDG-report-Gender-equality-in-the- 2030-Agenda-for-Sustainable-Development-2018-en.pdf

SHARE OF FEMALE RESEARCHERS BY COUNTRY, 1999-2015



Source: UNESCO Institute for Statistics 2017a.

Notes: Data refer to latest available from 143 countries. Data are based on headcounts (HC), except for Congo, India and Israel, which are based on full-time equivalents (FTE). Data for China are based on total research and development (R&D) personnel instead of researchers. Data for Brazil are based on estimations.

Why is metadata important[6]?

- Metadata makes data meaningful: Without metadata, one would not be able to understand data. For instance, look at the data table below. Would you be able to understand what it refers to? Without the appropriate metadata, which in this case would be the name of the indicator, definition of the indicator and unit of measurement, data is meaningless.

Figure 3: Data without metadata is meaningless

Country	Age	Reporting Type	Sex	2016
India	15-19	G	FEMALE	17.6 [c]
India	15-49	G	FEMALE	22.0 [c]
India	20-24	G	FEMALE	23.2 [c]
India	25-29	G	FEMALE	22.5 [c]
India	30-34	G	FEMALE	23.0 [c]
India	35-39	G	FEMALE	22.3 [c]
India	40-44	G	FEMALE	21.6 [c]

- Metadata improves comparability of data: Differences in data and their interpretation can arise due to the use of different definitions, concepts, units and classifications. When comparing data between countries or across time, make sure to look at the metadata fully

[6] Yongyi Min. 2019. "Introduction to metadata for global SDG indicators," Regional workshop on strengthening monitoring for the SDGs and selected SDG indicators, 23–25 September 2019, Nadi, Fiji.

for any inconsistencies or changes that may have taken place over time. For instance, in the case of child marriage, two different indicator series might be produced – girls married or in union before age 15 and girls married or in union before age 18. Metadata is important to understand the exact information being considered.

Figure 4: Data showing estimates for proportion of women married before age 15

Indicator 5.3.1, Series: Proportion of women aged 20-24 years who were married or in a union before age 15 (%)

Country	Age	Reporting Type	Sex	2014
Afghanistan	20-24	G	FEMALE	-
Bangladesh	20-24	G	FEMALE	22.4 [C]
Cambodia	20-24	G	FEMALE	1.9 [C]
India	20-24	G	FEMALE	-

- Metadata can provide information about inconsistencies in computation methods: For instance, the SDG indicator on adolescent birth rates is defined as the annual number of births to women and girls aged 15 to 19, per 1,000 women and girls in the respective age group. The metadata for this SDG indicator, however, clarifies that depending on the type of data source used to calculate this indicator, the method of computation differs. This information is important to understand any possible discrepancies in estimates over time.

Figure 5: Data showing estimates for proportion of women married before age 18

Indicator 5.3.1, Series: Proportion of women aged 20-24 years who were married or in a union before age 18 (%)

Country	Age	Reporting Type	Sex	2014
Afghanistan	20-24	G	FEMALE	-
Bangladesh	20-24	G	FEMALE	58.6 [C]
Cambodia	20-24	G	FEMALE	18.5 [C]
India	20-24	G	FEMALE	-

Figure 6: Computation method may vary depending on type of data

Civil registration data	Survey data	Census data
• The numerator is the registered number of live births by women aged 15- 19 during a given year. The denominator is the estimated or enumerated population of women aged 15-19 years.	• The numerator is the number of live births obtained from retrospective birth histories of the interviewed women who were 15-19 years of age at the time of the births during a reference period before the interview. The denominator is person- years lived between the ages of 15-19 years by the interviewed women during the same reference period.	• The adolescent birth rate is computed on the basis of the date of last birth or the number of births in the 12 months preceding the enumeration. The Census provides both the numerator and the denominator for the rates.

1.3 International definitions

Internationally agreed definitions exist for almost all statistical concepts. In fact, when new indicators are developed, the obtention of international agreement on definitions is often the first step towards data production. The use of such definitions ensures the international comparability of the data. For SDG indicators, these definitions and classifications can be found in the SDG metadata repository. When interpreting data, these must be kept in mind.

For example, when measuring the proportion of urban population of women living in slums, one would expect that a slum dweller is a person living in poverty. However, understanding the official definition highlights that wealth is not necessarily a precondition for slum-dwelling. A person is statistically classified as a slum dweller if they live in an urban area in a household that lacks at least one of the following:

- Improved water source
- Improved sanitation facilities
- Sufficient living area
- Durable materials
- Security of tenure

Additional definitions are also in place for each of these concepts. Such definitions help identify whether or not a household should be statistically classified as a 'slum' household. As such, a household is classified as a slum household if it lacks:

- Access to improved water: A household is considered to have access to improved drinking water if it has a sufficient amount of water (20 litres/person/day) for family use, at an affordable price (less than 10% of the total household income) and is available to household members without being subject to extreme effort (less than one hour a day for the minimum sufficient quantity), especially to women and children.
- Access to improved sanitation: A household is considered to have access to improved sanitation if an excreta disposal system, either in the form of a private toilet or a public toilet shared with a reasonable number of people, is available to household members.
- Sufficient living area/overcrowding: A dwelling unit provides sufficient living area for the household members if not more than three people share the same habitable room
- Structural quality/durability of dwellings: A house is considered 'durable' if it is built on a non- hazardous location and has a permanent and adequate structure able to protect its inhabitants from the extremes of climatic conditions – such as rain, heat, cold and humidity.
- Security of tenure: Security of tenure is understood as a set of relationships with respect to housing and land, established through statutory or customary law or informal or hybrid arrangements, that enable one to live in one's home with security, peace and dignity.

Understanding the official definition of a 'slum' and of women living in slums is therefore essential for data users to be able to understand and interpret gender data. Another example where international definitions are extremely important is disability. Disability statistics are often derived from census data, due to the fact that disability is a phenomenon of relatively rare incidence and therefore large sample sizes are needed to derive estimates. However, traditionally, different countries have included different questions in their census questionnaires to identify whether a respondent was disabled. Asking someone directly whether or not they are disabled might be inappropriate and often yields underreported results. For this reason, the Washington Group[7] (a city-group within the UN Statistical Commission) developed a set of standardized questions and countries are encouraged to use this common set to assess the prevalence of disabilities.

According to the Washington Group's set of questions, respondents are asked to respond to the following:

1. Do you have difficulty seeing, even if wearing glasses?
2. Do you have difficulty hearing, even if using aid?
3. Do you have difficulty remembering or concentrating?
4. Do you have difficulty with self-care?
5. Using your usual language, do you have difficulty understanding or being understood?

A larger set of questions has also been produced by the Washington Group for countries who might wish to assess disability in more detail. Understanding how to classify respondents as disabled and non- disabled, as per the questions above, is important for data users to properly interpret any disability figures. This, once again, highlights the importance of metadata.

1.4 Data

Data are measurements or observations that are collected as a source of information. There are a variety of different types of data and different ways to represent data[8]. In statistical circles, two main types of data are used: macrodata and microdata.

Macrodata is aggregated data usually obtained from aggregating individual-level records into a figure that is representative of a population group. Some of the most widely used statistics are macrodata. For instance, figures about a country's Gross Domestic Product, unemployment rate, or inflation rates are all microdata, as one single value is representative of a country (e.g. individual values have been aggregated to come up with a single value that is nationally representative). When national estimates are disaggregated by sex, location, geographical unit or other variables, this is still macrodata. Macrodata aggregates are often prepared by a country's National Statistics Office or other members of the National Statistical System. You should choose macrodata when you are looking for readily available estimates, representative of a country or select groups within the country.

Microdata is data made up of individual-level records. For instance, in survey datasets there are individual records for each of the survey's respondents. This is microdata. You should turn to microdata when the aggregated data that you are looking for is not available from official statistics

[7] http://www.washingtongroup-disability.com/

[8] See Australian Bureau of Statistics. "What are data".

or if you want to conduct further testing (such as finding associations between variables by running correlation tests, attempting regression analysis or working on modelling).

Typically, both macrodata and microdata are stored in databases, online repositories and data servers. While macrodata is often openly available online, accessing microdata might sometimes entail the need to submit a formal request and sign a confidentiality agreement.

1.5 Variable

An element or factor that can vary or change and is not fixed[9]. In statistics, a variable is any factor that is capable of having multiple values. For example, age as a variable can have multiple values – such as, 25, 31, 42, 65 years old, etc. A variable may also be called a data item. When working with survey data, each of the survey questions is typically a variable, although some variables are composites of several questions.

Box 1: Additional examples of variables

Some more examples of variables:

-Age	-Age at first pregnancy
-Sex	-No. of children
-Marital status	-No. of people in house
-Age at death	

Figure 7: Age is a variable as it varies from person to person

[9] Ibid.

2. Avoiding common mistakes when interpreting data: Understanding the semantics

2.1. The difference between ratio, rate, proportion, percentage and percentage points

All of these measures are statistically different. It is important to understand the differences and avoid using these words interchangeably.

2.1.1. Ratio

A ratio compares the frequency of one value for a variable with another value for the same variable. For example, when a coin is tossed 20 times, let's suppose that heads turn up 12 times and tails turn up 8 times. In this case, the ratio of heads to tails is 12:8 (spoken as 12 to 8).

We know these figures are ratios because we are looking at one single variable (coin tosses), which can return two possible values (heads or tails).

In official statistics, ratios are often used with bases of 1, 10, 100, 1000, 10,000 or 100,000. Among development indicators, an example of ratio is maternal mortality ratio (MMR), which is defined as the number of maternal deaths during a given time period per 100,000 live births during the same time period. Here, again, we are looking at one single variable (women going into labour and delivering live children) and two possible outcomes: death or survival of the mother.

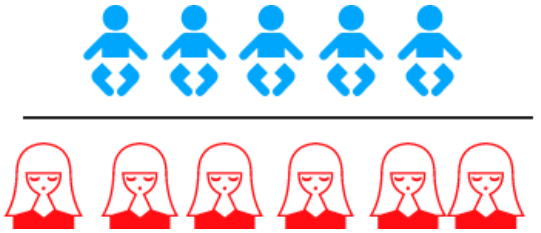
If a country's MMR is 200, it means that 200 mothers died for every 100,000 live births delivered.

Another example is sex ratio at birth which represents the number of male births per 100 female births. In this context, we are examining two possible outcomes (boy or girl) within the same variable, which is the live births. A sex ratio at birth of 105 means that per 100 female births are born 105 males, or per female are born five males (1:5).

2.1.2. Rate

Rate is a measurement of one value for a variable in relation to another measured quantity. When using rates, it is recommended that the reference period for the numerator and enumerator are equal. Among development indicators, one example of rate is the adolescent birth rate. Adolescent

birth rate is the number of births delivered by women aged 15–19 years per 1,000 women in that age group. Here, unlike the ratio, two different variables are being considered in the numerator and denominator respectively. These variables are: births (numerator) and women of a certain age group (denominator). Both statistics (births and number of women of a certain age) should ideally refer to the same year.



Observations being counted
(Number of births delivered by young women)

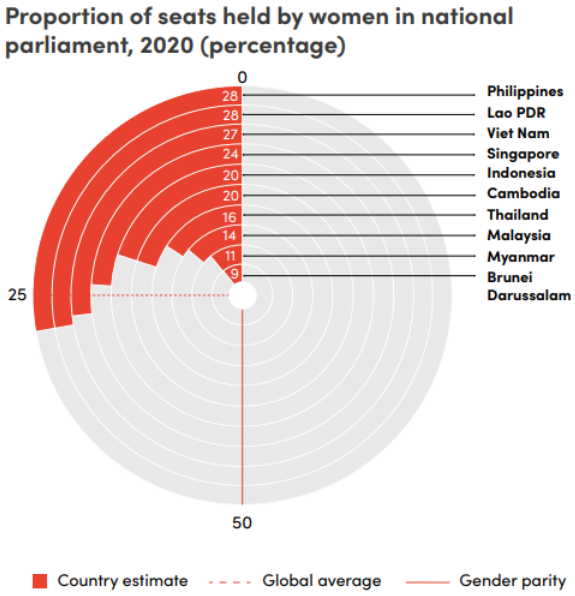
Measured quatity (total no. of young women)

2.1.3. Proportion

Number of times a particular value for a variable has been observed, divided by the total number of values in the population. Proportions are one of the most statistically used concepts in development indicators. They are easy to understand, as they represent the parts of a whole.

Proportions are always a number between 0 and 1 and can be expressed as a percentage when you multiply them by 100.

For example, the proportion of seats held by women in national parliaments is calculated by dividing the number of seats held by women in the national parliament by the total number of seats in the national parliament.

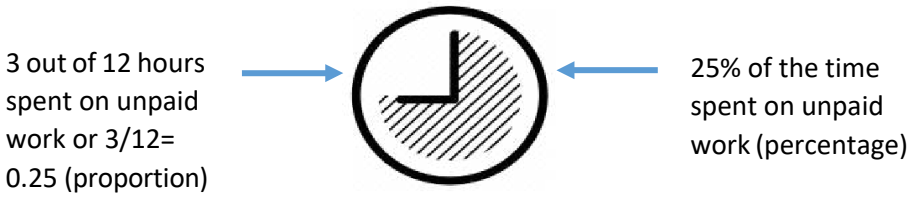


Note that although many SDG indicators use the term proportion, the values are in practice often expressed as percentages.

2.1.4. Percentage

A percentage is the expression of a value for a variable in relation to a whole population as a fraction of one hundred. Proportions are often expressed as percentages.

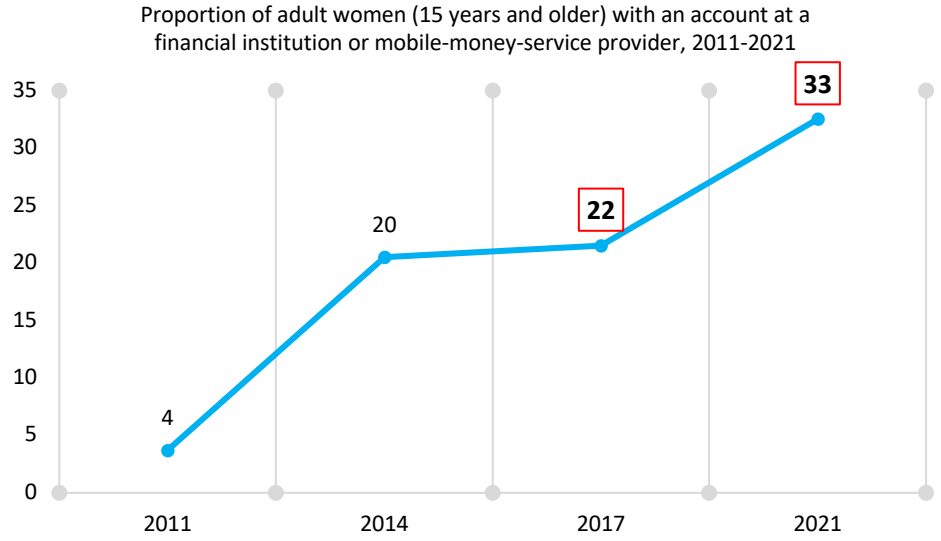
For example: Let's take the indicator 'Proportion of time spent on unpaid care and domestic work'. We could say that someone spends three out of 12 hours on unpaid care and domestic work, or we could express this value on the basis of 100 and say that someone spends 25 per cent of their time doing this work.



Do not use percentages when the total size of your sample or population is small, as the use of a percentage may be deceiving. This is particularly problematic when conducting trend analysis. For instance, if a peace negotiation was attended by 3 men and 2 women, it is technically true that 60 per cent of the negotiators were men and 40 per cent women. However, if the second phase of this negotiation process takes place one month later, but one of the men does not attend (e.g. only two men and two women attend), the percentages will stand at 50 per cent men and 50 per cent women. As you can see, the use of percentages might give a deceiving impression of progress in gender parity.

2.1.5. Percentage points

Percentage points are used to express increments, drops or differences. Percentage points often represent decimal points. It is very important to understand that percent and percentage points are completely different concepts and cannot be used interchangeably.



In the example above, the proportion of women with an account at a financial institution has increased between 2011 and 2021. To talk about the increase, we must use percentage points. In other words, the increase has been of 11 percentage points between 2011 and 2021 (33− 22 = +11). It would be incorrect to note that the increase has been of 11 per cent instead, because that refers to a "rate of change" or "percentage change" instead. We use the later if the figure is multiplied by 100 and thus appears in percentage form.

To calculate a rate of change (or the rate of increase in this example), one needs to divide the difference in percentage points by the initial value. This is to see how much change has taken place with respect to the starting point. In this case, the rate of increase will be calculated as: (33-22)/22 = 50%.

Thus, in this example one could note that between 2017 and 2021, the proportion of women with access to financing has increased by 11 percentage points, or it has increased by 50 per cent.

2.2. The difference between mean, median, average and total

2.2.1. Mean

Mean is the sum of all the values in a set, divided by the total number of values. It is the most commonly used measure of central tendency. The mean is affected by changes in any of its values. The mean may or may not be an observed value in the dataset.

Take the following dataset, for example:

2,3,5,6,20

The mean is calculated by adding all the values and dividing by the total number of values, as shown below:

$$\text{Mean} = \frac{\text{Sum of all observations}}{\text{Total number of observations}}$$

$$\text{Mean} = \frac{2+3+5+6+20}{5} = 7.2$$

The 'mean' is a good measure for symmetric continuous distributions, but it is not a robust measure, meaning it is influenced by outliers[10]. Outliers are data points that are far from other data points. In other words, they are unusual values in a data set. An outlier will pull the value of the mean in its direction and away from the location of majority of the observations.

For instance, in a distribution such as [1,1,1,1,1,1,1,1,1,1,1,1000] The mean is 77.8 - although the majority of the values are actually 1. Thus, in the presence of outliers, the mean might not be a suitable measure of central tendency because it may not be a good representation of the observations in the distribution.

Average

Statisticians don't really use the word average. The more precise terms are mean or median. If using the word average, please specify whether you are referring to the mean or the median. Most of the times, non-experts use the word "average" to refer to the mean.

2.2.2. Median

The 'median' is the numeric value separating the higher half of a sample, a population, or a distribution, from the lower half. In practice, it is computed by arranging the numbers in ascending order and locating the middle number in the centre of that distribution. It is also a measure of central tendency, as it indicates the relative position of an observation in the distribution.

The 'median' is much more robust than the mean, as it is not influenced by outliers. For instance, take a look at the distribution in Figure 8. This is the same distribution of numbers that we had in the previous example. Here, the median is 5, which is different from the mean value.

How to choose median?

If your distribution has an even number of observations, the mean would be the sum of the two middle numbers, divided by 2.

If your distribution has an odd number of observations, choose the number that falls in the middle.

[10] Outliers are observations that are markedly different from the rest of the data items.

Figure 8: Distribution to calculate median

2 3 **5** 6 20

Similarly, for the second example considered above, a distribution such as [1,1,1,1,1,1,1,1,1,1,1,1000], the median would be 1, rather than 77.8. Thus, mean and median aren't words that can be used interchangeably.

When to use the mean and the median?

It is best to use the mean as a measure of central tendency when you have a symmetric distribution. However, in a skewed distribution, it is best to use the median.

2.2.3. Total

The total value is a whole number or amount. For instance, when measuring poverty, we can say that roughly 730 million people lived below the poverty line in 2015, or about 10 per cent of the world's population. The former value (730 million) is the total value.

Exercise caution when utilizing total values for comparisons. For instance, if we only say "200,000 more people are now living in poverty" it appears as a negative development. However, due to overall population increases, it is possible that the actual poverty rates might have dropped over time, which is a positive development. When interpreting data and making comparisons, it is important to look at the relative values, not just the totals.

2.2.4. Minimum, Maximum, and Range

Minimum is the lowest value and maximum is the highest value in the data set. If we get the difference between the maximum and the minimum, we get the range. In data analysis, it is good to always include the minimum and the maximum values. For example, the lowest monthly household income is $5,000 and the highest monthly household monthly income is $15,000.

If we get the range, it will be $15,000 - $5,000 = $10,000. However, range can be misleading when we have outliers in the dataset. Range is a measure of spread in the data. If you only have one set of data, there are no other values to compare it to, so interpreting the range does not provide any useful information because there are no other data points to establish a range with. Rather, we simply say, the monthly income range of households is from $5,000 to $15,000.

2.3. Other misinterpretation issues specific to gender data

As we saw in Module 1, gender statistics is a field of statistics which cuts across many other fields of statistics to reflect the realities of the lives of women and men and policy issues relating to gender equality[11]. The following issues include select statistical areas relevant for gender equality that are often prone to misinterpretation. This list of areas is not necessarily comprehensive but can provide a rough idea of some of the key issues that gender data users often encounter, and how to deal with them to interpret gender data correctly.

2.3.1. Interviewing only the household head to obtain data

When sex-disaggregated data does not exist because individual-level surveys are not conducted, statisticians and policymakers often turn to disaggregating the data by sex of the household head as a gendered measure. Household head, however, is not an adequate measure as it does not capture some gender differences appropriately. For instance:

- It often provides biased responses, as male household heads might not have accurate information about women's reproductive health choices, use of time, etc.
- It is unsuitable to provide unbiased information about violence against women, control issues, etc.
- It fails to capture intra-household inequalities, including inequitable use of resources, issues of agency and decision-making power.
- Because most women-headed households are single-parent households or unmarried women's households, but most male-headed households are two-parent family-type households, the information between women-headed and men-headed households may carry bias for some indicators.

It is therefore important to obtain gendered information through individual-level records where both adults in the household are interviewed. That is, questions about women must be asked to women.

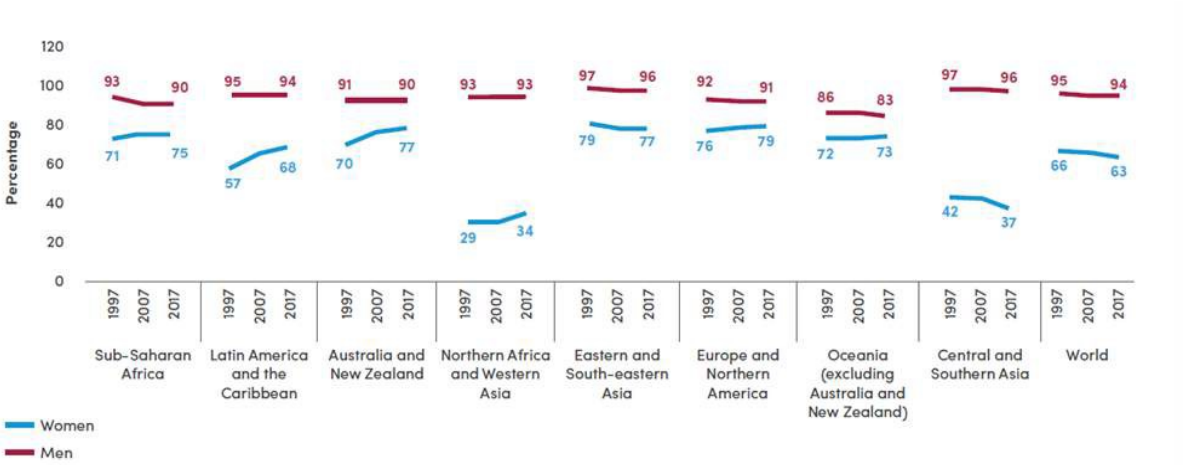
[11] See UNECE. 2010. Developing Gender Statistics.
https://www.unece.org/fileadmin/DAM/stats/publications/Developing_Gender_Statistics.pdf

2.3.2. Measuring gender gaps

Measuring gender gaps might provide an interesting picture of the degree of equality in a country. However, always keep in mind that the trend of the gaps might differ from the overall trend of an indicator.

For instance, take a look the graphic below. It is possible that, utilizing the below information, someone would write an article headlined: "Sex gap in labour force participation finally shrinks in Oceania". From this headline, readers might assume this is good news, as women might be accessing more jobs. However, by looking at additional information, such as the sex-disaggregated labour force participation rates for women and men provided in the graph below (rather than considering the gap only), it is obvious that women's participation has only increased by 1%. The reduction in the gender pay gap may also be attributed to a decline in men's participation in the labor force. The headline would still hold true, as the overall gap in Oceania has decreased overall, but women's labour force participant remains lower than men's throughout the time period.

Figure 9: Labour force participation rate among population aged 25–54 by sex and region, 1997–2017



Source: UN Women 2018, Turning Promises into Action

2.3.3. Violence and crime data

Violence and crime data must always be interpreted carefully. As these estimates refer to very sensitive issues, they are consistently underreported. For instance, if an enumerator is sent out to the field to compile information on intimate partner violence, it is almost certain that not every single woman victim of violence will admit being a victim. As violence and crime are sensitive topics, disclosure rates are low, and statistical estimates never capture the full extent of the problem.
It is also for this reason that survey data is always preferable to administrative data when it comes to violence and crime statistics. For instance, police records of instances of violence could never be comparable to prevalence data derived from specialized surveys. This is the case because most victims do not report instances to the police. Some of the reasons behind this include:

- Victims in many countries fear for their own safety if they report cases
- Victims often believe reporting to the police or to the justice systems won't lead to results, as many perpetrators aren't brought to justice and cases are often dropped
- In some cases, the stigma associated with violence also prevents victims from reporting

Respondents are more likely to talk about cases of violence/crime when asked (as opposed to voluntarily reporting or registering cases). There are a number of reasons why specialized violence surveys are more likely to yield more reliable estimates:

- Enumerators are specifically trained to build rapport with victims.
- Trained enumerators for these surveys are more sensitive to confidentiality issues and also aware of the psychological harm a woman can go through while recalling and reporting violent instances.
- Women are interviewed separately, at a time and/or place when the possible perpetrator (e.g. husband or other) is not around.
- Specialized survey questionnaires are designed thoughtfully, with the question order and wording carefully crafted to introduce the topic slowly and produce more reliable estimates.
- Very specific questions are asked to potential victims. For instance, rather than asking directly if someone has been a victim of violence, an enumerator might ask a general question such as: "Do you think it is justified for a man to beat his wife if she burns the food?". After the victim appears comfortable responding to these kinds of questions, enumerators might move to more targeted questions such as "Does your husband ever push you, shake you or throw something at you?". Targeted questions such as this allow enumerators to classify violence cases as physical, social or psychological violence.

Similar issues are associated with other forms of crime statistics. As people are usually unlikely to disclose being perpetrators or even witnesses of crime, victimization surveys are also more reliable in capturing these instances than police records.

Please note: Due to the many difficulties associated with collecting accurate violence and crime data, when violence/crime estimates increase, it doesn't necessarily mean that violence/crime increases! It might be a result of increased disclosure rates, better trained enumerators, or increased rapport between enumerators and victims.

Please also note: Due to all the complexities associated with violence statistics, specialized training on this topic is necessary before conducting a violence survey. For those interested in this line of work, it is highly recommended to undergo specialized training. UNFPA and the University of Melbourne have made specialized training on this topic available through: https://asiapacific.unfpa.org/en/publications/project- overview-knowvawdata

2.3.4. Time use data

Time-use statistics are quantitative summaries of how individuals "spend" or allocate their time over a specified period – typically over the 24 hours of a day or over the 7 days of a week[12]. Time-use surveys (TUS) capture activities and time spent doing those activities. In many instances, time-use surveys also capture the location for certain activities, and the number of people that might have been present at the time. Thus, among other purposes, time-use surveys are useful to capture the proportion of time spent by women and men on unpaid care and domestic work.

Some key issues to keep in mind while interpreting time-use data:

- Because people usually spend their time differently during weekdays and weekends, time-use surveys yield better estimates if the information is collected over different days in a week. In addition, in countries where seasonality influences jobs and daily activities, time-use surveys must be repeated in different seasons to capture such differences.
- The International Classification of Activities for Time Use Statistics (ICATUS)[13] is a classification of all the activities a person may spend time on during the 24 hours in a day. Its purpose is to serve as a standard framework for time-use statistics based on activities grouped in a meaningful way.
- When measuring unpaid care and domestic work, we only refer to work that is for own-use and that takes shape in the form of services. If someone is producing goods for own use or family use, these won't be classified as unpaid domestic work.
- Time-use information can be compiled through stylized questions (where respondents are asked to estimate how much time per day they spend doing one specific activity) or through time diaries (where respondents list all activities performed in a certain time interval). Statistics obtained through diaries are more accurate, as the diary method is the only method that captures simultaneity. For instance, if a woman is asked how much time she spends taking care of her child she might say 5 hours. If she is asked how much time she spends cooking, she might say 2 hours. With stylized questions, a statistician might be calculating a total of 7 hours of unpaid care and domestic work (5+2). However, if these two activities happened simultaneously, the total amount of time spent on unpaid care and domestic work is overestimated through double counting (in reality only 5 hours were spent in total). Diaries are helpful in identifying activities that happen simultaneously, and therefore yield more reliable estimates.

[12] See UNSD. 2005. https://unstats.un.org/unsd/publication/SeriesF/SeriesF_93E.pdf

[13] See UNSD for more on ICATUS https://unstats.un.org/unsd/statcom/48th-session/documents/BG-3h- ICATUS-2016-13-February-2017-E.pdf

Table 1: Example of a time diary

Activity categories		04:00-05:00				05:00-06:00				06:00-07:00				07:00-08:00			
Sleeping and resting	1																
Eating	2																
Personal care	3																
School (also homework)	4																
Work as employed	5																
Own business work	6																
Framing	7																
Animal rearing	8																
Fishing	9																
Shopping/getting services	10																
Weaving, sewing	11																
Cooking	12																
Domestic work	13																
Care for children	14																
Commuting	15																
Traveling	16																
Watching TV	17																
Reading	18																
Sitting with family	19																

For additional details about time use, you are encouraged to consult Module 5, which discusses the methodology for SDG Indicator 5.4.1 on 'Time spent on unpaid care and domestic work'. In addition, please refer to some of the resources listed in the 'list of resources' for this module, to access details about time-use surveys and ICATUS.

2.3.5. Sex-disaggregated poverty rates

Poverty rates are typically calculated at the household level. That is, income or expenditure data is often compiled for a household. Assessments of how many men and women live in poverty have therefore traditionally been calculated by utilizing household measures of income and/or consumption and matching this information with the number of men and women that live inside each household (based on survey or census information on household composition). However, such measures fail to capture intra- household inequalities. That is, resources – monetary or otherwise – are often unequally distributed among household members. In order to capture accurate measures of individual poverty, separate assessments of income and/or expenditure at the individual level are necessary. In practice, few surveys ask for this level of information. When interpreting sex-disaggregated measures of poverty, it is important to check the indicator metadata to assess whether the data source pertains to household-level or individual-level surveys, to see if the estimates capture intra-household inequalities.

2.3.6. Gender pay gap

The gender pay gap is a measure of what women are paid relative to men. It is commonly calculated by dividing women's pay by men's pay, and this ratio is often expressed as a percent, or in dollar terms. This tells us how much a woman is paid for each dollar paid to a man.

Estimates on the gender pay gap refer to mean hourly earnings from paid employment of

employees by sex[14]. It is important to note that the data for this indicator should always be interpreted by occupation and level and taking into consideration total time worked. Therefore, the indicator can be used to assess if equal pay is in place for equal work. A common misconception about this indicator is that the pay gap just reflects the average pay of women vs. the average pay of men in a certain country. However, the indicator in fact compares earnings for a certain occupation and level. That is, it refers to the gross remuneration in cash or in kind paid to employees for time worked or work done, together with remuneration for time not worked, such as annual vacation, other type of paid leave or holidays. It excludes employers' contributions on behalf of their employees paid to social security and pension schemes and also the benefits received by employees under these schemes. Earnings, as considered for this indicator, also exclude severance and termination pay.

The gender pay gap is measured in two ways: (i) the unadjusted gender pay gap, and (ii) the adjusted gender pay gap.

The unadjusted gender pay gap is measured as the difference between the average hourly earnings of women and men expressed as a percentage of the average hourly earnings of men. This is a simple measure, which gives a first indication of the gender gap in pay, and may be easily understood and used by policy makers[15].

$$Unadjusted\ Gender\ Pay\ gap = \frac{(Mean\ hourly\ earnings\ of\ men - Mean\ hourly\ earnings\ of\ women)}{Mean\ Hourly\ earnings\ of\ men}$$

This gap is not the pay gap between a man and a woman with the same observable characteristics, doing the same work; it is the difference between the average wage levels of all working women and men. That being said, this measure does not differentiate between the portion of the gender pay gap attributed to discrimination and the portion resulting from various factors such as education, work experience, etc.

In order to decompose the gender pay gap and distinguish between possible discrimination and differences in endowments of characteristics of the labor force, the method of Oaxaca Blinder is applied for the adjusted gender pay gap. The adjusted gender pay gap takes into account various factors that can contribute to differences in pay between men and women. The pay differential between males and females, is decomposed into three parts accounting for: 1) differences in endowments, 2) differences in coefficients, and 3) the interaction between endowments and coefficients.

The adjusted gender pay gap is calculated using statistical methods, such as regression analysis, to estimate what the pay gap would be if men and women had the same characteristics and were equally distributed across occupations and other relevant factors. The formula for the adjusted gender pay gap is more complex due to the statistical methods involved, but the key idea is to quantify the pay gap that remains after accounting for factors other than gender.

[14] See: https://unstats.un.org/sdgs/metadata/

[15] SDG Indicator Metadata 8.5.1. https://unstats.un.org/sdgs/metadata/files/Metadata-08-05-01.pdf

The difference between the non-adjusted gender pay gap and the adjusted gender pay gap is that the non-adjusted gap is a straightforward comparison of average earnings between men and women, while the adjusted gap takes into account various factors to provide a more accurate understanding of the pay disparity that can be attributed to gender discrimination. The adjusted gap provides a more nuanced view and helps to identify areas where gender-based pay disparities exist even after accounting for other relevant variables.

3. KEY TAKEAWAYS

- *Always refer to metadata and international definitions when interpreting data*
- *Percentage is different from percentage points*
- *Rate is different from ratio*
- *Mean is different from median, although both are measures of central tendency*
- *Median is a better measure of central tendency when a distribution is skewed because it does not get affected by extreme values*
- *Data disaggregated by household head is not a good substitute for sex-disaggregated data*
- *Violence statistics are always underreported*
- *Time-use statistics are more accurate when compiled using time diaries, because they capture simultaneity*
- *Poverty rates are difficult to calculate at the individual level. If applying household composition to perform sex disaggregation, the estimates will fail to capture intra-household inequalities*
- *Gender pay gaps attempt to capture whether men and women receive equal pay for equal work*