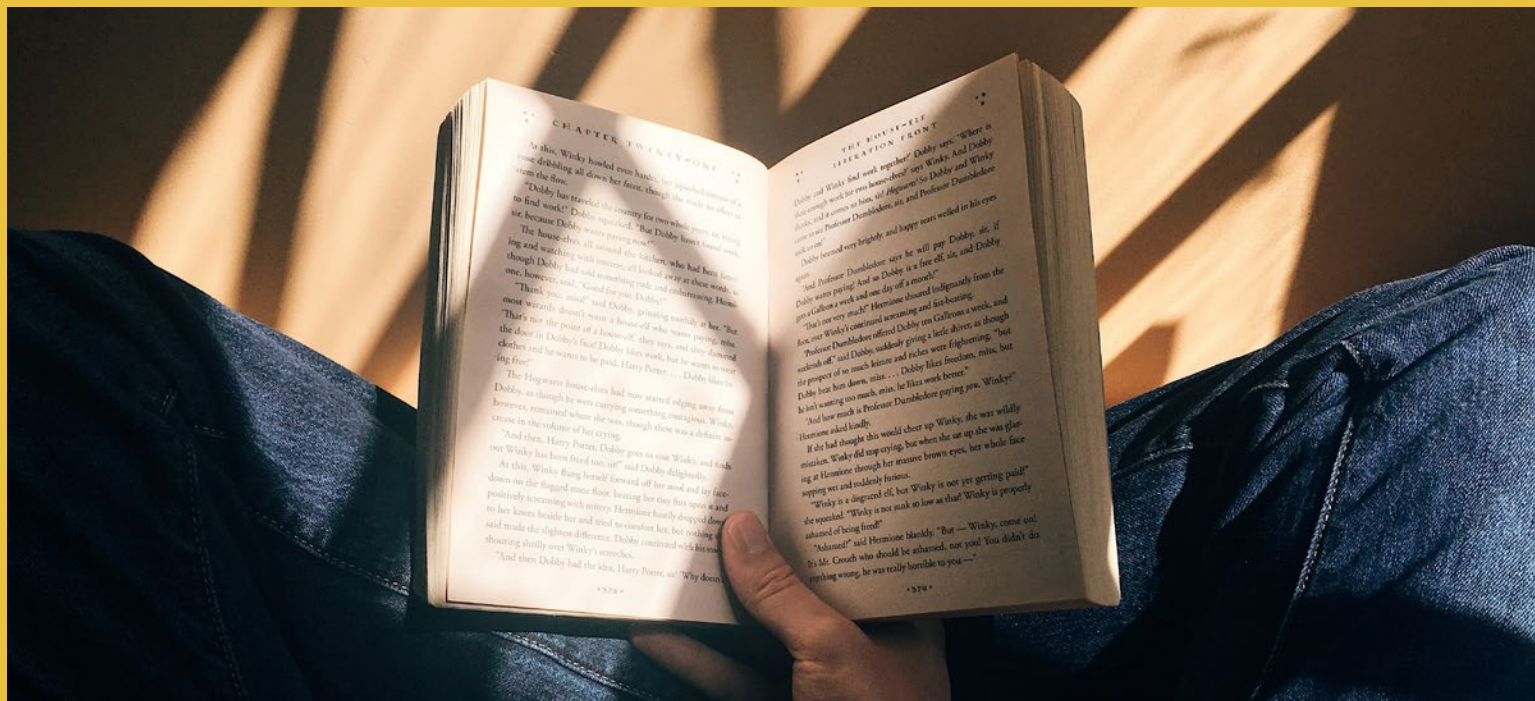# ILLUMINATE 01
DRAWING ON SECONDARY DATA

The second step to *Decoding Injustice* is to *illuminate* the underlying issues behind the inequalities ingrained in our economic systems. In this note, activists and changemakers will find an introduction to the use of data (the preferred language of policy-makers), uncovering patterns and trends that might otherwise remain hidden, and will learn about the benefits of drawing insights from already existing, or secondary, data.

Key Questions

What is secondary data?

How does secondary data fit into the OPERA Framework?

Where can secondary data be found?

What are factors to consider when using secondary data?

ILLUMINATE 01
DRAWING ON SECONDARY DATA

Introduction

Data is a type of knowledge that can be highly effective in making the case for change. In essence, it answers the questions posed by indicators. This makes it a really useful way to help illuminate the problems in the economic system that you've interrogated, as a first step in decoding injustice. This note focuses on secondary data— in other words, data that already exists because it was collected by someone else, or for some other purpose, or a combination of the two — which has a number of particular advantages for our research.

First, in our information age, data has come to play a central role in how government officials, public servants and other policy-makers see the world and understand social problems. Data is a language they are typically fluent in and use regularly in their work, and it's seen as a way to make decisions that are evidence-based. It's a language that often prioritizes technical expertise because it's seen as neutral or apolitical. The amount of secondary data is growing exponentially. In reality, data is far from neutral or apolitical. The way it is created and interpreted is filled with assumptions and biases. For this reason, it's important for us as activists to be able to critically read data in order to make use of it astutely. Doing so can help to illuminate trends and patterns that help debunk myths, uncover ideological assumptions and reveal new insights.

Second, there are also practical reasons for activists to draw on secondary data in their research. As the breadth of available data grows and the cost of accessing it shrinks, it's increasingly possible to curate, rather than *create*, the data we need to answer the questions posed by indicators. This means selecting, organizing and presenting secondary data. Drawing on secondary data can help to avoid unnecessary duplication and save a significant amount of time and resources when conducting research.

For these reasons, secondary data can be a strategic tool for decoding injustice. That said, it is not without its challenges! In particular, a lot can get hidden behind numbers, and so it's important to read data with a critical eye. This note introduces different types of secondary data and explains how they can be used with indicators (and benchmarks) under each dimension of the OPERA Framework. It also discusses where to find secondary data and, importantly, how to evaluate whether or not it is strategically useful. Finally, it flags some of the shortcomings of secondary data and highlights how primary data (which we will cover in Illuminate 2 - Collecting Primary Data) might help to address them.

As we will discuss below, this can reveal the deep disparities between the bargaining power of these different groups, which can distort or "capture" policy-making. The influence of international economic institutions is particularly significant.

What Types Of Secondary Data Can We Use?

The major categories of data are quantitative and qualitative:

- **Quantitative data** refers to a number; for example, the size of a population, the price of an item, or a score on a test.
- **Qualitative data** describes the non-numerical characteristics of something; for example, a description of an experience, an agreement with a statement, or an opinion about an issue.

Quantitative data is well suited for spotting patterns and mapping trends. A larger and wider number of facts can be captured, categorized and compared across groups and over time. This helps us make more specific, more precise findings that support advocacy for change. Consider findings such as "*a large number* of people do not have access to water" or "the

---

This document is organized according to an innovative method for collecting, analyzing and presenting evidence around three steps:

INTERROGATE
Map the problem in depth using OPERA to identify indicators and benchmarks.

ILLUMINATE
Spotlight the underlying issues by collecting, analyzing and visualizing data.

INSPIRE
Take action to build power and hold decision-makers accountable.

proportion of children underweight for their age *rose dramatically*". The words highlighted in italics are open to challenge, and to different interpretations by different people, and the statements are therefore imprecise. Replacing these subjective words with numbers that can be objectively verified can give greater credibility to findings.

Relevant categories of secondary data include:

EVENTS-BASED DATA

Events-based data is qualitative or quantitative data that documents the *occurrence of an incident*, such as the number of people whose homes were destroyed in a wildfire. This kind of data can be compiled and consolidated using common definitions and classifications. It is often captured in the databases of organizations that handle complaints from or deliver social services to particular communities. Other sources might include information provided by the media or contained in reports by government, civil society organizations and international human rights oversight bodies.

SOCIO-ECONOMIC AND ADMINISTRATIVE STATISTICS

Socio-economic and administrative statistics are aggregate datasets compiled and disseminated by state agencies. This is usually done in collaboration with national statistical agencies, and it may be carried out with guidance from international organizations, or be outsourced to private companies. This type of data is used for policy-making and planning purposes, and can be categorized into three sub-types:

Administrative records are generated when ministries, government agencies and other public institutions interact with the public and record information related to the number and profile of their beneficiaries (e.g., data held by schools about their students and hospitals about their patients, or a finance ministry about taxpayers). Administrative records can be extremely useful. However, because data is limited to those segments of the population that use a particular public service, their coverage is not always reliable or complete. For example, crime statistics cover only crimes that are reported, and we know that some crimes, particularly sexual assault and domestic violence, are severely underreported.

Statistical surveys, also called sample surveys, collect data from a subset of the population, with the goal of drawing inferences for the entire population. Surveys can be a cost-effective way to collect information. Many indicators are tracked using statistical survey data. Examples include demographic health surveys, labor force surveys, and household income and expenditure surveys.

Censuses collect data about all members of a population. In other words, everyone is counted. Countries typically conduct censuses of population, housing, agriculture and industrial establishments. A population census is usually conducted at 10-year intervals because of the complexity and cost of the operation. Censuses provide baseline data on the structure and key characteristics of the population that do not change rapidly.

BUDGETARY DATA

As budgets are original government documents, they are considered primary data sources. Budget documents are important sources of public finance data, but they are typically long and largely technocratic in nature. Trying to read and make sense of them can be a daunting exercise. The good news is that not all budget analysis techniques are highly technical and sophisticated. With some basic familiarity with budget terminology and using straightforward arithmetic — adding and subtracting, multiplying and dividing — we can use some simple techniques to analyze budgetary data. First and foremost the principles of accountability, transparency and participation demand that all budgetary documents at all tier levels of governance are open and accessible to the public - although this is not always the case in practice. Secondly, qualitative techniques can help to gather feedback from particular rights holders. Quantitative perception surveys and indicators (e.g. the Open Budget Index) can provide a general overview of the country's situation. Rights-based methods can be used to reveal whether the allocations made in previous years and the current fiscal year in key social sectors are done in an equitable and efficient manner. Some relevant methods/ calculations include:

1. Calculate the percentage of the State's budget allocated to social spending relevant to the total budget. Place the relevant benchmarks and the outcome-output indicators set by the respective department/ Ministry in comparison with the specific human right.

2. Identify which community groups are benefitting either directly or indirectly from the spending; contrast spending disparities with progress made towards each of the outcomes of the identified state interventions from a rights-based lens.

3. Compare allocations to previous budgets to see how spending has evolved over time, taking into account economic growth over the period.

4. Track utilization levels of key social sector schemes for the current fiscal year with previous years (e.g. using public expenditure tracking surveys or social audits).

5. Collect feedback on public participation in the design, implementation and evaluation of fiscal and monetary policies (e.g. through interviews or other qualitative methods and quantitative data, if available). Perception surveys too can be useful in analyzing whether policy processes are participatory, transparent and accountable.

6. Analyze indicators related to transparency of economic policy process.

A human rights-based approach ties the main objectives of development-related policies, legislation, regulations to budgets that have the potential to actually fulfill rights equally for all people. In future Decoding Injustice resources, we will dive into budget analysis in more detail.

Budgets: a glossary

The terminology used in budgets differs significantly from country to country, so it is difficult to compile a common glossary for all countries. Nevertheless, there are some basic concepts that can help us better understand where, and how much, public money is being allocated.

The first distinction is between two main kinds of expenditures:

Recurrent expenditures are expenses incurred year after year for running public administration, such as salaries, procuring goods and services, and providing subsidies.

Capital expenditures are one-off expenses for building assets to improve the productive capacity of the economy, such as purchasing land and developing infrastructure.

The second distinction relates to the way allocations are presented in the budget:

Economic classifications provide limited information, simply identifying how much money is being allocated to different "inputs", such as salaries and wages, utilities, travel, or printing costs. They may also be referred to as line items or object classifications.

Administrative classifications identify the entity responsible for managing the allocated funds, such as the ministry of education and health or, at a lower level, schools and hospitals.

Functional classifications categorize expenditure according to the purposes and objectives for which they are intended.

Source: International Monetary Fund, *Manual on Fiscal Transparency: Glossary*, 2007

ECONOMIC DATA

Finance ministries, international financial institutions, financial firms and civil society organizations also publish economic statistics. These are used to measure the health of a country's economy over time (e.g., monthly, quarterly, annually). Examples include: Gross Domestic Product (GDP), inflation rates, consumer price index, credit ratings, and inflows of development assistance and foreign direct investment. Government budgets are of course another hugely important piece of economic data; those are dealt with in the next module on primary data.

Other relevant data comes from public procurement monitoring, oversight from parliamentary bodies and government audits. Public sector audits generally take one of the following three forms:

Financial audits: when the auditor verifies the accuracy and fairness of the State's financial statements. The auditor scrutinizes a sample of transactions to establish their authenticity and determine whether the accounts fairly present the government's financial affairs. In many countries, the audit report also contains a formal opinion by the auditor on whether the financial statements present a true and fair picture of the government's financial position and whether the receipts and payments have been applied as per the budget law.

Compliance audits: when the auditor answers the following questions:

- Has the expenditure been authorized by a competent authority?
- Has the expenditure been authorized by the budget appropriation law and made in accordance with the terms of the law?

WHAT IS AGGREGATED AND DISAGGREGATED DATA?

Aggregated data is collected without making any distinctions: for example, the number of people without access to water in a community is 6,000. Disaggregated data is divided according to set criteria, such as sex, age, or the area where people live. The above data, disaggregated, might tell us that 70% of people without access to water in this community are women. Disaggregated data help to uncover disadvantage that would otherwise remain hidden, and show where needs are greatest.

- Does the expenditure conform to the procedures (relevant rules, regulations and orders) promulgated under the country's various public finance and other laws?

Performance audits: where auditors broaden their mandates to include measuring "value for money", and typically report on three factors:

1. **Economy:** Can the government program being audited (e.g., building a road, operating a hospital, rolling out unemployment benefits) run at less expense?
2. **Efficiency:** Is the program achieving maximum impact with minimum cost?
3. **Effectiveness:** Is the program delivering its intended results, assessed by measuring program performance indicators against actual results?

Due to growing community action in several countries in the past few decades that has led to successful implementation of access to information and right to information laws, **social audits** have become increasingly popular. Social audits are a powerful tool that furthers public disclosure by involving local communities and groups in order to hold social sector programmes, schemes and interventions to account. Going beyond traditional audits, social audits involve all stakeholders including service providers, local civil society organizations, public officials, multilateral donor agencies and immediate and potential beneficiaries of the intended public services by examining the intended and unintended impact of the given government intervention. As a part of the process, public consultations are held to produce qualitative data that help in generating progress indicators and plugging information gaps generally not found in public documents. Social audits also allow face-to-face community sit-ins with public officials as well as are helpful in raising awareness. In certain countries, some government departments themselves have incorporated the practice of conducting social audits in an effort to strengthen citizen engagement and participation.

PERCEPTION AND OPINION SURVEYS

Perception and opinion surveys ask a representative sample of individuals for their personal views on a given issue. The information collected is, for the most part, subjective. Questions are often asked in a way that the answers are predetermined, or in "closed" formats (such as multiple-choice questions) to make it easier to interpret the data. Scales are also often used (e.g., rating satisfaction with a policy from 0 to 10). Well-known household perception and opinion surveys include those carried out by Gallup International Association, Afrobarometer, Latinobarómetro, Asian Barometer and Eurobarometer. They act as international measures of public opinion or perception on issues such as democracy, governance, livelihoods, participation, conflict and crime.

DATA BASED ON EXPERT JUDGEMENTS

Data based on expert judgements is data generated from a limited number (or sample) of "informed experts". Like household perception and opinion surveys, the information generated is essentially subjective, but it can be translated into a quantitative form to make it easier to interpret. These surveys, which capture only the opinion of experts who presumably should be well informed about the topic, have drawn criticism for not being representative of the realities on the ground, and for the lack of transparency in the selection of experts. Examples include reports by Freedom House, which uses expert judgements to assess and rank countries according to political and civil freedom, and the Minorities at Risk Project, which follows the status and conflicts of politically active groups using multiple sources of information, such as the media, government reports, non-governmental reports and expert opinion.

What Types of Secondary Data Can We Use With Opera?

There is likely to be secondary data that is relevant for many — if not most — of the indicators identified across all four dimensions of OPERA. These are summarized in the table and unpacked further below.

Outcomes	Policy Efforts	Resources	Assessment
• Socio-economic and administrative statistics • Perception and opinion surveys • Events-based data	• Categorical data about legislation and policies • Expert judgments • Socio-economic and administrative statistics • Perception and opinion surveys • Events-based data	• Audit data • Economic statistics • Perception and opinion surveys	• Perception and opinion surveys • Expert judgments

DATA ON OUTCOME INDICATORS

Data on outcome indicators seeks to measure:

- **Aggregate levels of rights enjoyment:** Socio-economic statistics can be assessed against relevant benchmarks and/or compared to those of similar countries. Deviations, or departures from the norm, can point to whether or not a country's performance is reasonable.

- **Disparities in rights enjoyment:** Disaggregated socio-economic statistics are particularly useful for uncovering inequalities by looking at disparities; for example, between women and men, or among ethnic groups or income levels.

- **Progress over time:** Socio-economic statistics can be compared over time to indicate progress or deterioration and whether disparities are growing or reducing.

This can be complemented with data from perception and opinion surveys that capture people's own views of their ability to enjoy their rights.

DATA ON POLICY EFFORT INDICATORS

Data on policy efforts seeks to:

- **Identify legal and policy commitments:** Categorical data (which puts items described into a category) can be helpful in comparing commitments made, relevant laws enacted and policies developed to other countries or against international standards and guidelines. For example, inventories of laws, policies and regulations in the area of land tenure can paint a picture of government efforts to address land rights. Expert judgment can be useful in assessing whether laws and policies are in line with good practice.

- **Examine implementation:** This data allows us to see how laws and policies translate into goods and services on the ground. Socio-economic and administrative statistics (such as teacher-to-student ratios, distance to nearest health clinic, number of beneficiaries of social welfare programs, quality of roads) can be compared, disaggregated and viewed over time to determine who is accessing what. Events-based data and perception and opinion surveys can offer insights on what it's like for people to access goods and services in practice.

- **Analyze whether legal and policy processes are participatory, transparent and accountable:** Perception surveys and governance data can be useful in this respect.

DATA ON RESOURCE INDICATORS

Resource indicator data is necessary to analyze:

- **Resource generation:** Evaluation of revenue sources, alongside relevant fiscal, monetary and macroeconomic policies, can be helpful in this regard.

- **Resource allocation:** data that shows allocation ratios (such as the share of the budget dedicated to particular

sectors or groups) can help to assess the reasonableness of amounts allocated when judged against reference points and over time.

- **Resource expenditure:** Various governance tools, such as social audits, can be helpful when reviewing the disbursement of allocated funds.

- **Policy processes:** This can be assessed by perception data that asks people how satisfied they are with the avenues for public participation in development of the government budget (e.g., whether people have access to information on budgetary processes). The Open Budget Index is a particularly useful source in this respect.

Where Can Secondary Data Be Found?

At the **national level**, a large amount of secondary data is likely to be publicly available on the website of the national statistical office or bureau of statistics. This is particularly true of censuses and surveys. Another source of data on government policy and programs is the evaluations carried out by the government itself. States often invest resources, time and money in the evaluation of their policies and programs to determine whether they are operating efficiently and effectively. Many of these evaluations will be available on the website of the relevant government ministry, or possibly through the country's ombudsman or auditor-general.

There has been a big push in recent years for governments to provide **open data** in order to increase transparency, accountability and citizen engagement. Numerous governments around the world have joined the Open Government Partnership, committing to reforms that should increase the amount of publicly available information. There are many examples of open data initiatives, including website portals and apps that allow citizens to access relevant government data.

When relevant data is not already publicly available, it may be necessary to **file an access to information request** in an attempt to obtain relevant documents.

At the **international level**, sources of secondary data include the United Nations Development Programme (UNDP), the World Bank, the World Health Organization (WHO) and many others. Their databases compile and aggregate national data,

which makes it easier to make comparisons between countries and among regions. Note that there may be differences in the data provided by international databases and that of national databases.

Civil society organizations also produce important data—both nationally and internationally. Where possible, comparing government statistics and reports with those of the international community or civil society organizations can help to ensure reliability, as well as identify discrepancies in the form of conflicting information or interpretations.

How Can We Judge Secondary Data?

Given the proliferation of data, it is important to consider whether specific secondary data is appropriate for what you want to do with it. It should always be interpreted cautiously, and there are a number of issues to weigh in this regard.

For example, what happens if there is a difference between data provided by global databases and national databases? This is not uncommon, due to differing methodologies for collecting, analyzing and interpreting data. As outlined in the table that follows, there are various statistical concepts that can help when comparing data sources.

Concept	Explanation	Potential Issues
Validity	Data must reflect what it is trying to measure (i.e., the fulfillment of a right) as closely and as accurately as possible.	Most secondary data is collected for uses other than human rights monitoring and so must be re-purposed. This reinterpretation can be challenging.
Reliability	Refers to the consistency or dependability of the data. In other words, data collected multiple times in the same way should deliver similar results.	Ambiguities or biases in the way data is collected (e.g., how survey questions are framed or how a population is sampled) may make data unreliable.
Impartiality	Data must be collected in a way that respects scientific independence and in an objective, professional and transparent manner.	A national statistical office that is not independent may present numbers in such a way that a situation appears better than it is.
Ability to be disaggregated	To be most useful, data should be disaggregated (or able to be disaggregated) by, for example, gender, ethnicity, age, economic status, disability and geography.	Access to the original dataset is often necessary in order to disaggregate the data, and it is not always easy to obtain.

How particular data compares against these concepts are influenced by the following questions:

- **How** are the questions for the data collection framed?

- **How broadly** is it collected (i.e., what is the sample size)?

- **Who** is the data collected by?

- **How frequently** is the data collected?

Reviewing metadata, which is essentially data *about* the data, can help to answer these questions. It provides information such as the computation methods used, the measures of variability and the description of errors in the data, such as bias and sampling errors.

CENTER for **ECONOMIC** and **SOCIAL RIGHTS**

The following table considers these questions for common types of secondary data:

Characteristics	Administrative Data	Surveys	Censuses
Breadth	All events registered	All sampled units	All units (100%)
Frequency	Ongoing	3-5 years	10 years
Bias	Can be incomplete	Usually biased, but good design can minimize this	Theoretically no bias
Potential for rights monitoring	Very good	Good	Good, but not in the short term

Source: OHCHR, HRBA Training Materials *(undated)*

If you find yourself with conflicting datasets, note the variation of the data in your findings and, if possible, explain why you think there is a difference between the datasets (e.g., due to variation in the data collection method). Choose one data point to use and explain why, in your opinion, it is the most valid, reliable or impartial. Reviewing the metadata will help you to make this kind of assessment.

Finally, careful attention must be paid to what data exists, what data does not exist and, importantly, *why* there are gaps in the data. The availability of data is often a **reflection of existing, unequal power relations**. If this is not acknowledged and the data treated with care, using it can actually reinforce existing discrimination and hurt more than it helps.

This is particularly important when considering the gender dimensions of an issue; data about women is notoriously lacking. The absence of gender-disaggregated data, in and of itself, says something about the marginalization of women or non-binary people, for example. The undervaluing of women's activities and priorities in every sphere has been replicated in the statistical record. In many countries, we do not even know how many people who identify as women or non-binary there are, let alone how poor they are, how healthy they are, how educated they are, what work they do or how they spend their time. This is not just a problem of a lack of gender-disaggregated data. Some of the issues most important to their lives are simply not counted. For example, labor market statistics rarely take into account the time spent on domestic work.

EXAMPLE: **DATA ON THE GENDER WAGE GAP IN EGYPT**

One of the labor rights indicators included in the Egypt Social Progress Indicators (ESPI) is the Gender Wage Gap. However, the secondary data available to measure it is not reliable. Specifically, the data sample size and scope of the survey used to determine women's employment in Egypt is very limited; in fact, it includes only 5.4% only of women's employment. For example, the sample used in the annual bulletin on employment, wages, and labor statistics does not include farmers, while the agricultural sector accounts for about (21.4%) of total female employment. To overcome this challenge, and avoid reinforcing patterns of economic inequality and invisibilization of women's struggles in Egypt, the ESPI team redefined the indicator and calculated the gender wage gap on the basis of the weighted average in the top three occupations held by women according to the available official data, which constitute about (67.9%) of the formal employment of women.

CONCLUDING THOUGHTS

There are both strategic and practical reasons to draw on secondary data to illuminate issues interrogated under each of the four dimensions of OPERA — including socio-economic statistics, administrative data, budgetary documents, perception surveys and expert judgments. From a strategic perspective, data is a language that government officials, public servants and other policy-makers are fluent in. When activists use this language, they are demonstrating that these actors do not have a monopoly on technical expertise, and that alternative ways of understanding a problem can also be evidence-based. From a practical perspective, a significant amount of secondary data is available online. If there are gaps in what you need, more is often publicly available from relevant government agencies, municipalities and service providers.

Because secondary data isn't neutral or apolitical, activists must be able to read data critically in order to make use of it astutely. Doing so can help illuminate trends and patterns that help debunk myths, uncover ideological assumptions and reveal new insights. For a lot of socioeconomic and administrative statistics, reviewing metadata against agreed criteria can help you decide if it helps you gain new insights into the issue you're researching and, if not, why not. Budgetary data can be explored with some basic familiarity with budget terminology and using straightforward arithmetic.

Reading data critically is important in its own right. But it may also point to the need to gather primary data. We discuss primary data in the second note in this module.