

# PRASC



**Project for the Regional  
Advancement of Statistics  
in the Caribbean**

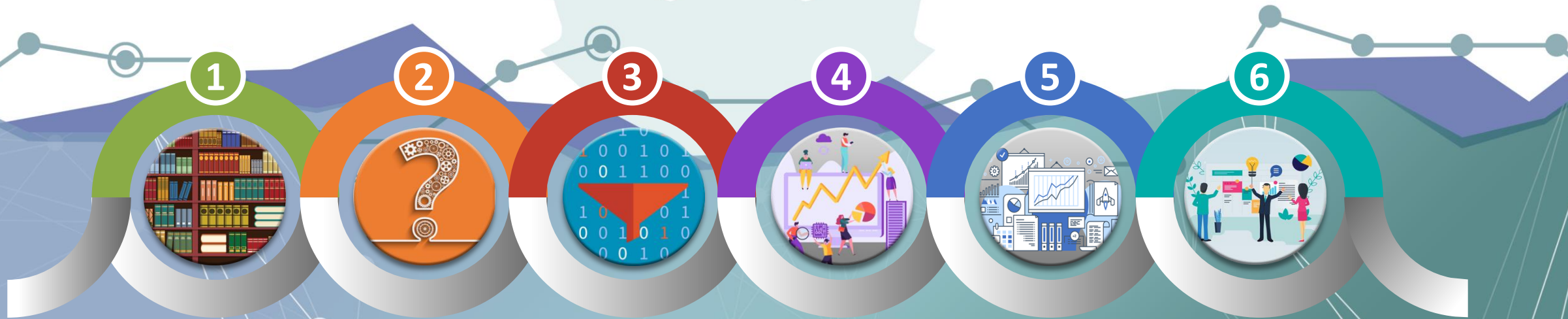
**Projet régional pour  
l'avancement de la statistique  
dans les Caraïbes**



In partnership with

**Canada**

# Steps of the analytical process



**Regional Census Analysis and Dissemination Workshop**  
**Statistics Canada**  
**December 1-7, 2022**

# Engagement

This course is meant to be very interactive.

Much of the classroom version of this course is dedicated to sharing stories as we work through the content.

When you see the slido icon, please enter the code provided to engage in the discussion!

Take  
notes  
[active  
listening]

slido

# Steps of the analytical process





# Objectives

Upon completion of this course, participants will understand the analytical process.

1. **Planning your analysis:** how to define your analytical objectives, how to create a formal analysis plan, and best practices for seeking early feedback;
2. **Performing your analysis:** how to best organize, prepare, and examine your data;
3. **Sharing your findings:** how to tell your data story and best practices for reviewing your work and presenting your findings to an audience.

This process applies to many different subject matters – social, business, health, economic...

# Data Literacy Training Initiative

Data literacy training

statcan.gc.ca/eng/wtc/data-literacy

Franglais

Statistics Canada Statistique Canada

Search website

Subjects Data Analysis Reference Geography Census Surveys and statistical programs About StatCan Canada.ca

Home > Workshops, training and conferences

## Data literacy training

As Canada's national statistical organization, Statistics Canada is committed to sharing our knowledge and expertise to help all Canadians develop their data literacy skills. The goal is to provide learners with information on the basic concepts and skills with regard to a range of data literacy topics.

The training is aimed at those who are new to data or those who have some experience with data but may need a refresher or want to expand their knowledge. We invite you to check out our [Learning catalogue](#) to learn more about our offerings including a great collection of short videos. Be sure to check back regularly as we will be continuing to release new training.

### What is data literacy?

Data literacy is the ability to derive meaningful information from data. It focuses on the competencies involved in working with data including the knowledge and skills to read, analyze, interpret, visualize and communicate data as well as understand the use of data in decision-making.

Data literacy also means having the knowledge and skills to be a good data steward including the ability to assess the quality of data, protect and secure data, and their responsible and ethical use.

**Data literacy competencies**  
Data literacy competencies are the knowledge and skills you need to effectively work with data.

**Data journey**  
The data journey represents the key stages of the data process starting with finding and exploring data through to telling the data story.

**Learning catalogue**  
Check out the data literacy training available from Statistics Canada

Report a problem or mistake on this page

Share this page

14:34 2021-05-04

Self-study suggestions to complement the course

Data Literacy Learning Catalogue:  
<https://www.statcan.gc.ca/eng/wtc/data-literacy/catalogue>  
Short videos on a variety of topics

# Agenda

**Friday, Dec 2 – 9:00 to 12:00 (EST)**

- Objectives
- Guiding Principles for Analysis
- Planning your Analysis
- Exercise 1

Health break at around 10h30

**Monday, Dec 5 – 9:00 to 12:00 (EST)**

- Implementing Your Plan
- Descriptive Analysis

Health break at around 10h30

**Tuesday, Dec 6 – 9:00 to 12:00 (EST)**

- Sharing Your Findings
- Exercise 2
- Guest Speaker: Dr. Godfrey St. Bernard

Health break at around 10h30

1



2



3



4



5



6



# Module 1: Planning your Analysis

- ✓ Generic Statistical Business Process Model
- ✓ Guiding principles for analysis
- ✓ Identifying analytical objectives that add value-added
- ✓ Planning your analysis and getting feedback





# The role of analysis at a statistical office

**Analysis is about converting “data into information”**

Analysis can serve different roles at a national statistical office:

1) Support **core information products**:

- Provides quality assurance on major statistical outputs (analysis as data verification)
- Enhances the interpretability of statistical outputs, improving the relevance of the data to users

2) Support the development of **in-depth research products** on topics of importance to stakeholders:

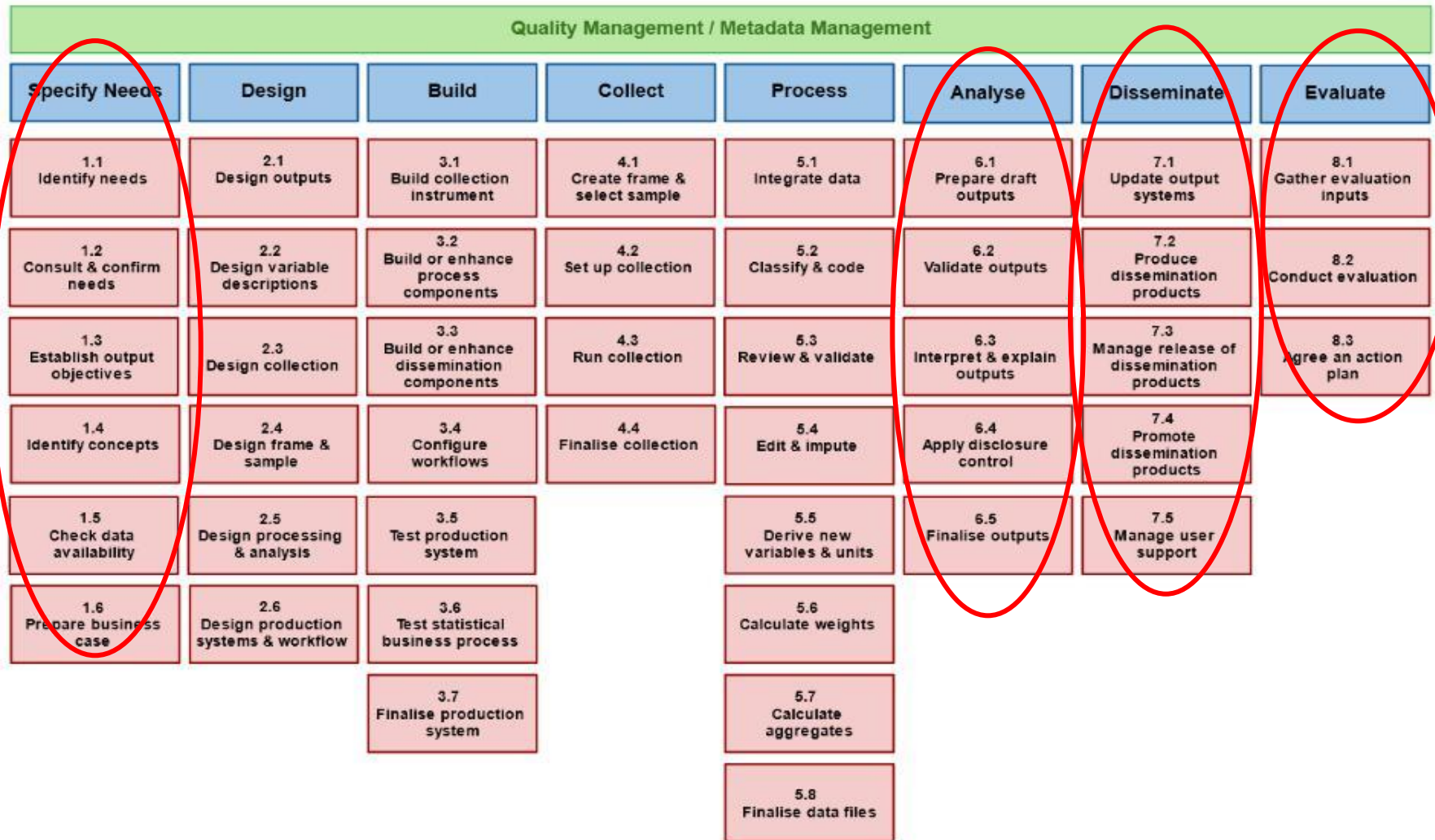
- Fosters a leadership role in key subject matter areas
- Leads to new data sources and new statistical outputs



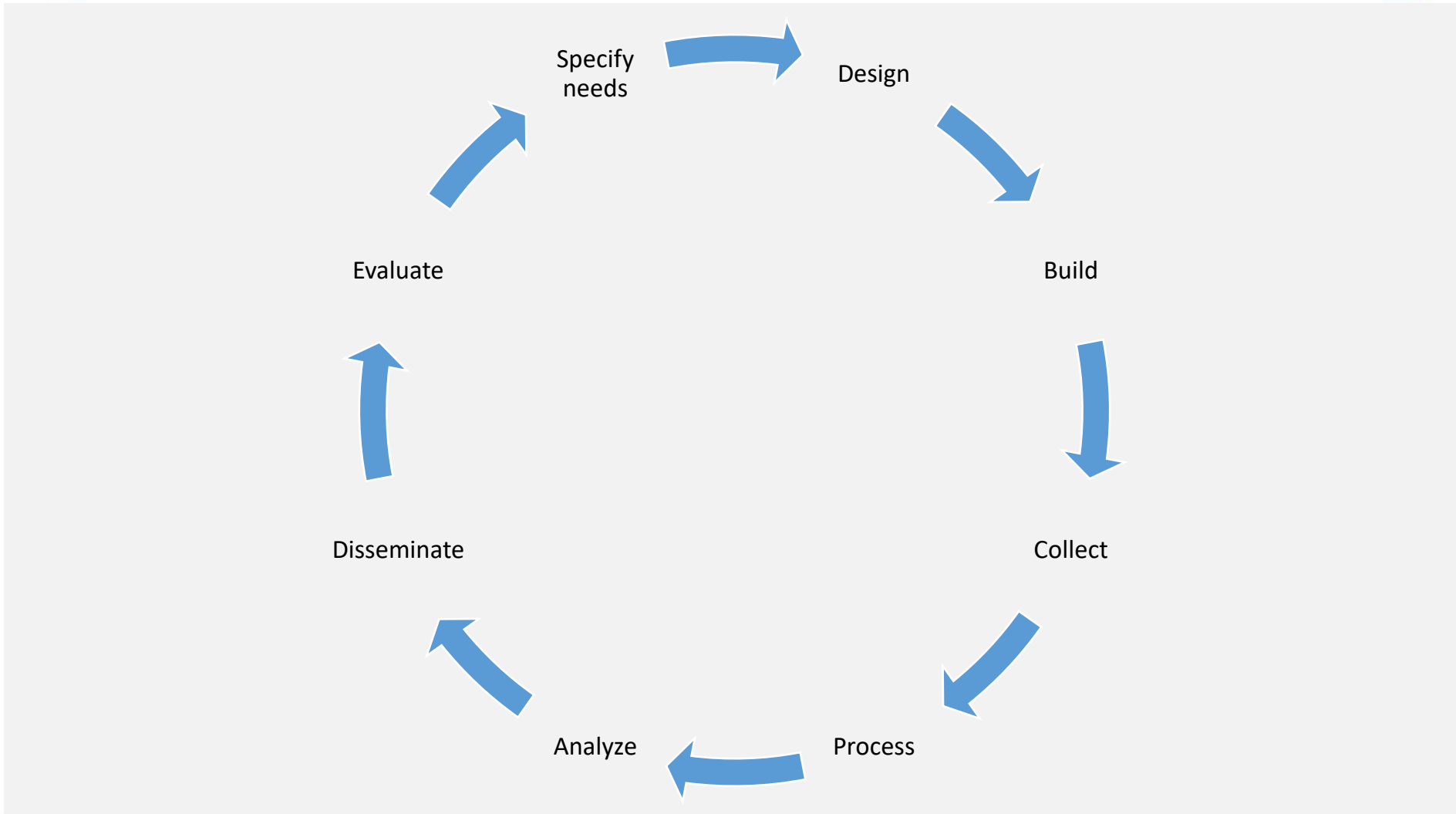
# Analysis as part of the statistical process

- The **Generic Statistical Business Process Model (GSBPM)** is a model that describes the various steps of the statistical process
- Widely used at national statistical offices, the GSBPM provides a framework for understanding how analysis and research support different stages of the statistical process
- **Analysis and research make important contributions at many stages of the statistical process**

# The Generic Statistical Business Process Model (GSBPM)



# The Generic Statistical Business Process Model (GSBPM)



# Guiding principles for analysis

slido

Which examples below are good general guiding principles for analysis?  
Select as many as you wish.

Neutrality

Relevance

Rigour

Interpretability

Accessibility

Coherence

Complexity



# Analysis is central to National Statistical Offices' mandate

- At Statistics Canada we have three guiding principles for analysis:
  - ✓ Neutrality
  - ✓ Relevance
  - ✓ Rigour

Our New MISSION Statement

**Serving Canada with high-quality  
statistical information that matters.**

# Neutrality

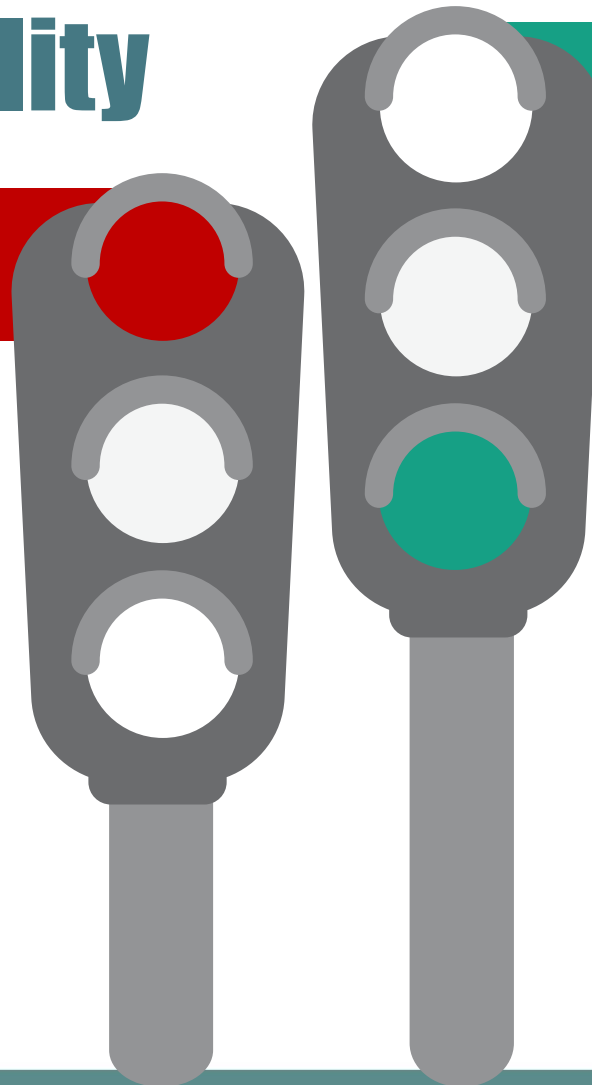
- Neutrality means we **inform** policy through analysis
  - We don't comment.
  - We don't "take sides".
  - We don't critique or evaluate policy.
- We aim for an **impartial** presentation of results.
- **We let the data speak for themselves.**



# Maintain neutrality

## Subjective

- Massive
- Alarming
- "Only"

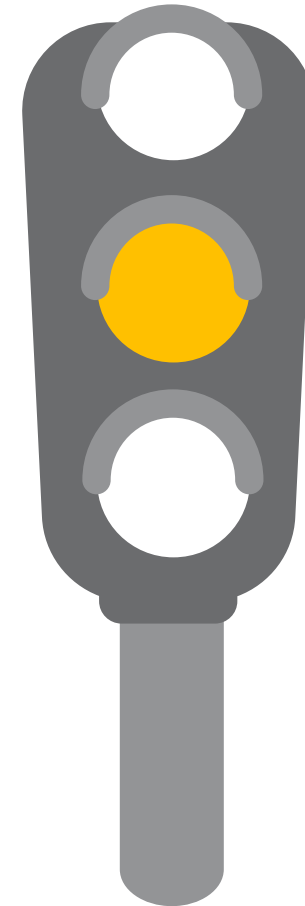


## Neutral

- Increased or decreased by X%
- Higher or lower by X times

# Maintain neutrality

- Bias or subjectivity can also happen earlier in the analytical process...
- Can you think of other ways neutrality may be compromised?



slido

# Relevance

slido

Which examples below are good ways to make sure that our work is relevant? Select as many as you wish.



My organization has information needs.



A stakeholder has information needs.



A subject matter expert identifies an emerging information need.

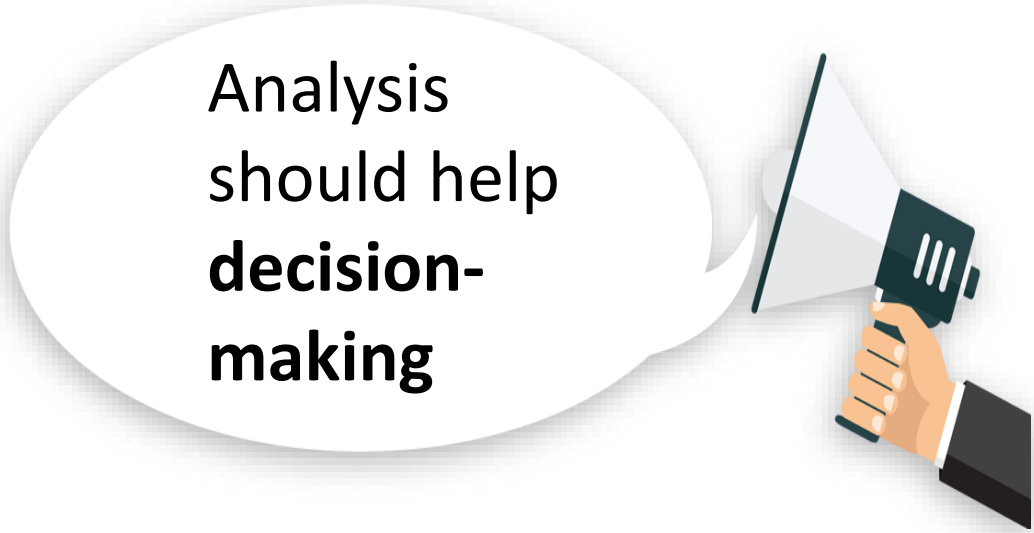


My community has information needs.

I am curious and I want to find answers to my existential questions.

# Relevance

- Analysis should address **issues of importance** to policy-makers, stakeholders, data users, and the general population.
- Analysis should contribute to a **better understanding** of current and emerging economic or social issues.



Analysis should help **decision-making**

# Rigour

- Analysis should be **statistically defensible** and supported by the data.
- Methodology is **transparent**.
- Conclusions are fact-based or **evidence-based**.
- Data go through an extensive quality assurance process (**quality control**).





# Quality control...

**Appropriate checks and balances exist to ensure that analytical products published by statistical offices meet expectations regarding analytical quality.**

- All products (information, analytical, methodological) for which Statistics Canada is solely or jointly responsible, are subject to **review** prior to release outside the Agency
  - Peer and institutional review
- The review should ensure that their content is compatible with the Agency's **mandate**, and that they adhere to generally accepted **norms** of good professional practices.

# Data Quality in Six Dimensions

## Relevance

- Degree to which the information meets the needs of users and stakeholders

## Timeliness

- Delay between the time when the data are meaningful and when they are available

## Accuracy

- Degree to which what is being measured is in line with what is actually true

## Accessibility

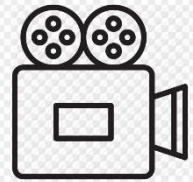
- Degree to which people are aware of and have access to the data

## Coherence

- The same concepts, definitions and methods are used overtime and across statistical programs

## Interpretability

- Metadata is provided to allow users to interpret the data properly



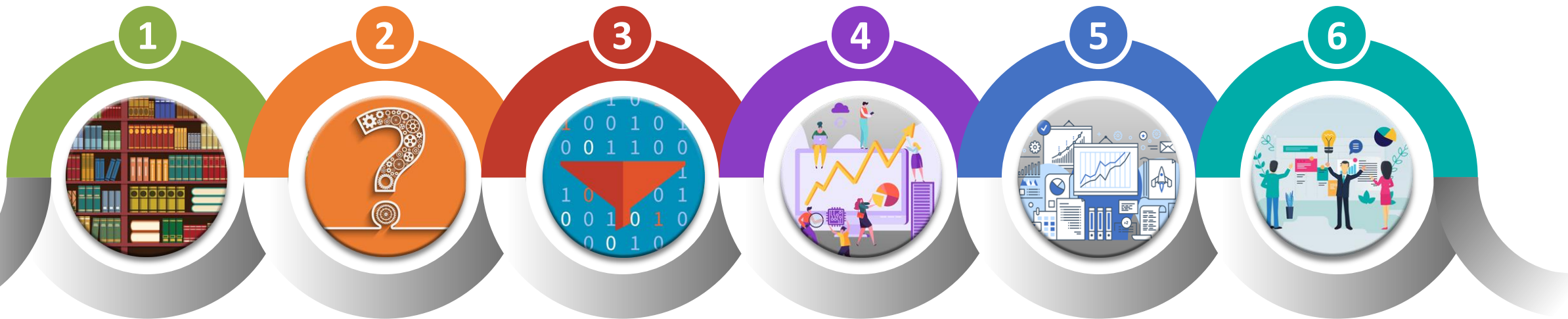
Self-study:  
Data Quality in  
Six Dimensions

# Steps of the analytical process



# Planning your analysis

How to define an analytical question and how to create a formal analysis plan



PLANNING YOUR  
ANALYSIS

# Step 1: What do we already know?



# Step 1: What do we already know?

- Understand the context and **previous work** on the topic.
- What are the knowledge or information **gaps**?
- What are the issues of interest or **concern** to stakeholders?



# Look for previously published information

- Government publications
  - All past Census releases
- Reports from reputable organizations (e.g., WHO)
- Peer-reviewed publications
  - Systematic reviews
  - Meta analyses
- \*Media reports
- \*Websites
- ...

## \*Ask yourself...

- ✓ Is the organization behind the report reputable?
- ✓ Is the journal reputable?
- ✓ Does the news story cite the original source? Can you find the original source?
- ✓ Websites: who is the organization putting out the information?



# Step 1: What do we already know?

What do we already know about ... in your country?  
Education,  
Housing,  
Employment,  
Migration, etc.

What do we already know about the topic?

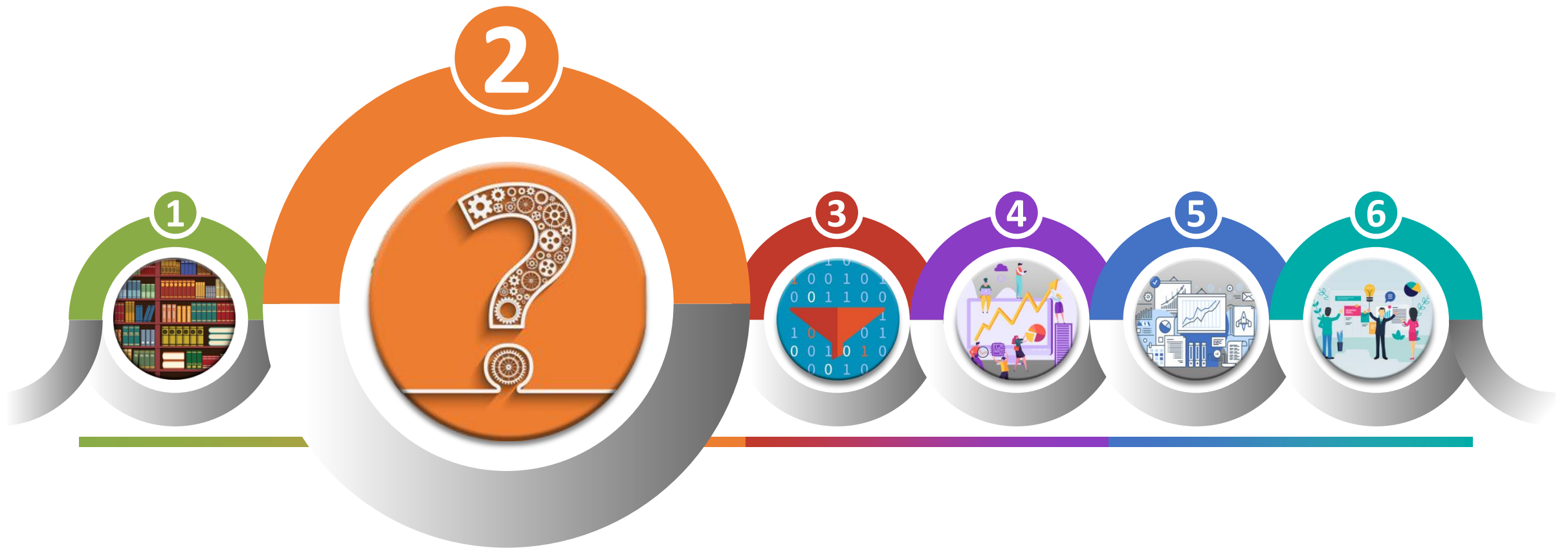
What are the information gaps or needs?

Question / Objective

Relevant contribution



# Step 2: What is your analytical objective?



# What is your analytical objective?

- What are you trying to find out?
- Why is this objective relevant?
  - Why should we care?
- What is the value added?
  - To your organization, your client, and/or our understanding of the topic.





# Best practice: Break down your results for sub-populations

- *Aggregates* can hide important differences in the experiences of particular population groups.
  - For example: an estimate for the total population of a country hides differences based on the lived experiences of men and women; the younger and the older; etc.
- By *disaggregating* data—that is, by breaking down statistics into **more refined categories**—we are able to uncover important differences across sub-populations.



# Best practice: Break down your results for sub-populations

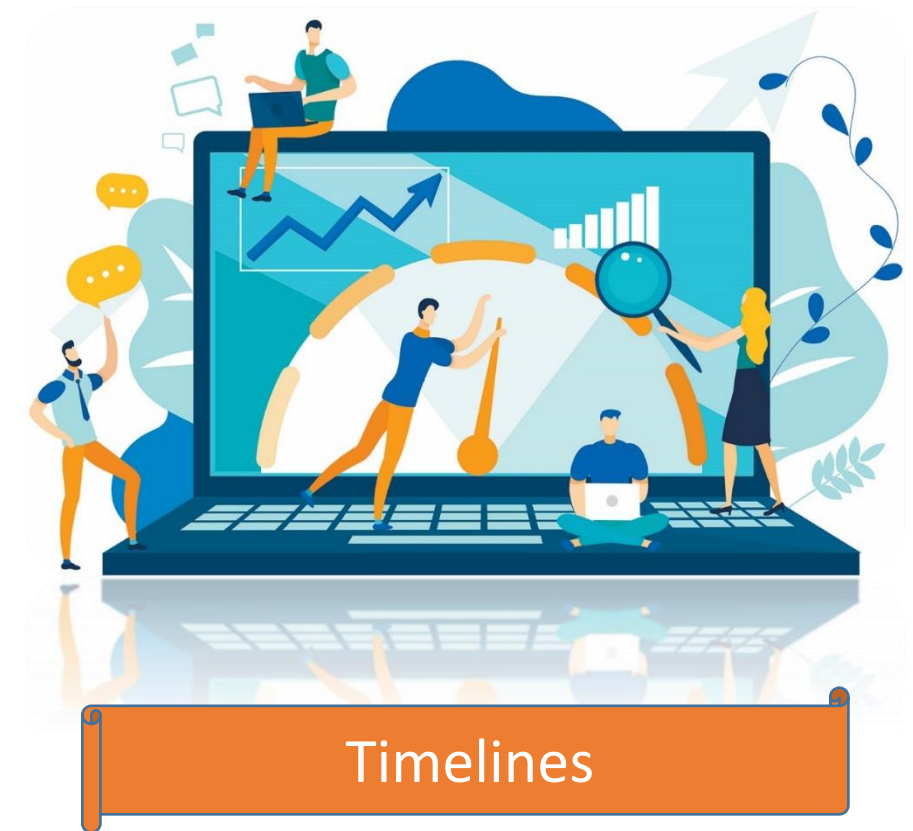
- Make sure that your results are disaggregated at the lowest level of population and geography possible, while respecting quality and confidentiality.
- What are the important population disaggregates for your country?



slido

# You are now ready to plan your analysis

- Put together a one page plan
  - ✓ Context
  - ✓ Analytical objectives
  - ✓ Data and methodology
  - ✓ How will you communicate your results?  
Detailed report, short report(s), fact sheet(s) or thematic report?
  - ✓ Template tables (tabulation plan)



# Tabulation plan example

Objective 1: Compare labour force participation of men versus women

Table template 1:

		Men	Women				
	Employed	%	%				
	Unemployed	%	%				
	Not in the labour force	%	%				



# Know your audience...

**Analysis and research products will differ in terms of their style and complexity of presentation**

- Many analytical publications are intended for **general** readers, while others are aimed at a more **specialized, technical** audience
- Certain publications focus more on descriptive methods while others place more emphasis on complex methods or multivariate modelling
- **Many products? Plan in advance!**

# Seek Feedback

- Can I benefit from feedback or expertise from colleagues? **Circulate your plan.**
  - ✓ Ongoing feedback, from the beginning
  - ✓ Subject matter
  - ✓ Data
  - ✓ Methodology
  - ✓ Terminology

Stakeholders, organizations?



# Quiz

slido

**You are tasked with analyzing data from the Census. How would you get started?**

- a) I would immediately start crunching data and produce a lot of tables on a lot of topics.
- ✓ I would search for previously published information to gather what we already know about a given topic and identify the data gaps.
- ✓ I would meet with stakeholders to discuss their information needs.
- ✓ I would prepare an analytical plan and circulate it for review.

# Exercise 1

- Familiarize yourself with the information available in your data source.
- Then, **generate three research questions** that can be answered with the available data.
- Post your questions in Slido.



slido



20 minutes  
Individual work



# Planning your analysis: Summary

- The analytical process can be viewed as a series of steps designed to address a **well-defined objective**.
- Once the topic has been defined and the value-added determined, the next step is to create an **analytical plan**.
- Incorporating **feedback** as early as you can is an important part of planning an analytical project.
- Make sure that your analysis involves a **purposeful**, objective and rigorous examination of data.



# Planning your analysis: Summary

- You are in the best position to understand the **context** surrounding your project/topic.
- Familiarizing yourself with findings from past Censuses is an important first step. This literature scan will ensure the relevance of your current analysis.
- Are there **trends** you need to monitor? **Comparisons** to be made against benchmarks? **Goals** to reach? **Differences** to understand?
- Are there important **stakeholders** to consult and engage with?

# Self-study suggestions for Module 1

Check out  
this free  
resource!

Self-study  
suggestions

The screenshot shows the Statistics Canada website at the URL <https://www.statcan.gc.ca/en/wtc/data-literacy>. The page title is "Data literacy training". The main content area features a banner with icons representing data literacy skills. Below the banner, there is a paragraph stating: "As Canada's national statistical organization, Statistics Canada is committed to sharing our knowledge and expertise to help all Canadians develop their data literacy skills. The goal is to provide learners with information on the basic concepts and skills with regard to a range of data literacy topics." This is followed by another paragraph: "The training is aimed at those who are new to data or those who have some experience with data but may need a refresher or want to expand their knowledge. We invite you to check out our [Learning catalogue](#) to learn more about our offerings including a great collection of short videos. Be sure to check back regularly as we will be continuing to release new training." Below this, there is a section titled "What is data literacy?" which explains that data literacy is the ability to derive meaningful information from data and focuses on competencies like reading, analyzing, interpreting, visualizing, and communicating data. It also mentions that data literacy includes the ability to assess the quality of data, protect and secure data, and use it responsibly and ethically. At the bottom of the page, there are three links: "Data literacy competencies", "Data journey", and "Learning catalogue". The "Learning catalogue" link is circled in red and has a tooltip that says "Check out the data literacy training available from Statistics Canada." The page also includes a "Report a problem or mistake on this page" button and a "Share this page" button.

## Analysis 101, part 1: Making an analytical plan

<https://www.statcan.gc.ca/en/wtc/data-literacy/catalogue/892000062020009>

## Data Quality in Six Dimensions

<https://www.statcan.gc.ca/en/wtc/data-literacy/catalogue/892000062020001>