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

Overview of Business Surveys

Project for the Regional Advancement of Statistics in the Caribbean (PRASC)

Delivering insight through data for a better Canada

Serge Godbout
February 17, 2022

Outline

- Why conduct business surveys
  - Introduction | Official economic statistics | Gross domestic product
- Infrastructure supporting the statistical framework
  - Standards | Administrative data | Statistical registers | Business surveys
- Some quality indicators
  - Frame coverage | Response rates | Coefficient of variation
- Process model for business surveys
  - From initialization to evaluation

Why do we need economic statistics?

Good question, isn't it…

- Statistics on the economy are essential for making informed decisions related to social, economic and environmental issues
- **Evidence-based decision making**

To answer it, nothing better than looking at some examples

Why do we need economic statistics?

(Some examples may not apply to all countries)

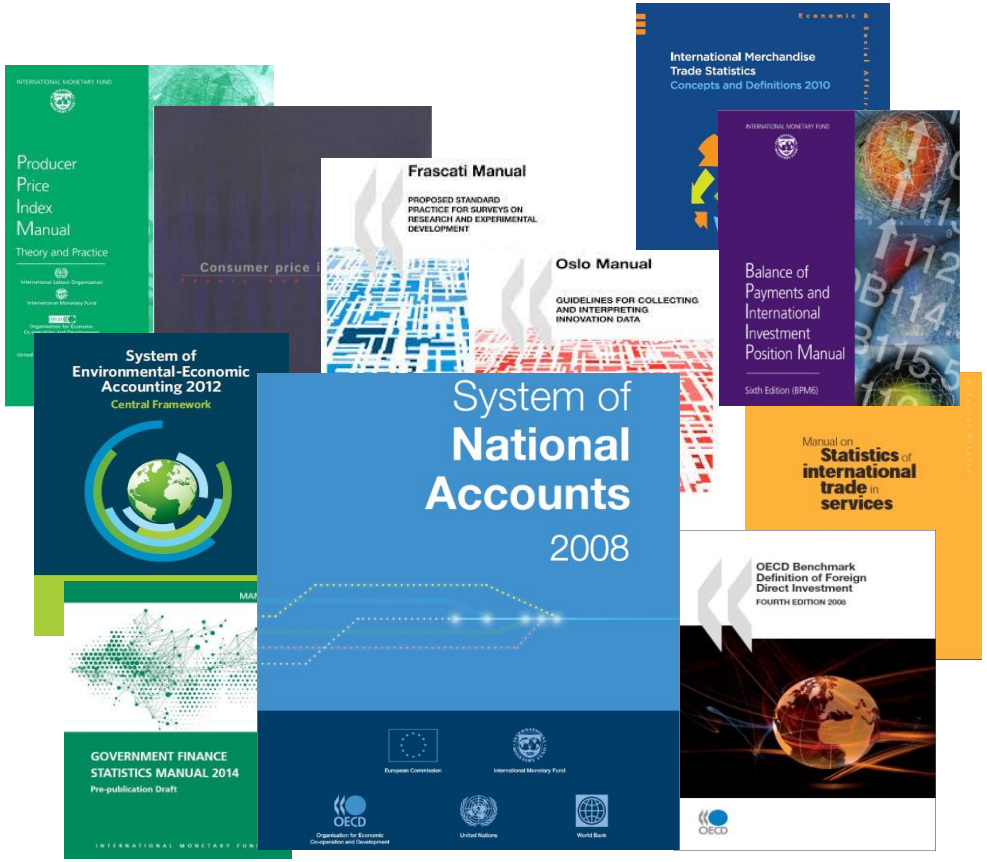
- If you had a business, how would you compare your performance to others?
- If you were to establish the government budget, what kind of information would you need?
- If you wanted to compare two countries' economies?
- If you were to lose your job – what are you getting?
- What transfers should be made to regions, what information is needed?
- How would you justify the salary increase?
- What are the impacts on economy of specific situations (natural disasters, COVID-19, etc.)
- Is the economy growing? What is driving the growth?
- Are firms investing (capital) in the country's economy? If so, where?
- Do firms control global value chains, are they participants or are they left out?
- Do businesses invest in: R&D, environment, skills development, etc.
- What are the sectors of the economy in which our country is strong/weak?

Why do we need economic statistics?

- Statistics Acts
  - Typically mandates the National Statistical Office (NSO) to collect data, compile them, produce estimates, as well as analyse and disseminate them
  - To do so, typically grants the NSO the right to collect data from citizens, businesses as well as organizations
  - Outlines the responsibilities of the NSO to protect the data
    - Confidentiality
    - Security

Official Economic Statistics

- Statistical frameworks (manuals) define:
  - The rules to be followed by National accountants
- Internationally recognized and adopted allowing for comparability
- These frameworks are concerned with:
  - The players
  - What they do, how they do it and where they do it
  - How to properly record their actions and interactions and the result of their actions in a meaningful way.



Official Economic Statistics

- Many key statistics are released monthly, with others released quarterly and annually
- Some examples
  - Consumer Price Index: Measure of inflation
  - Area specific estimates: Measure the health of different business sectors
  - Employment statistics: Measured from both the employer and employee side
  - Gross Domestic Product: Pulse of the overall economy

Gross Domestic Product (GDP)

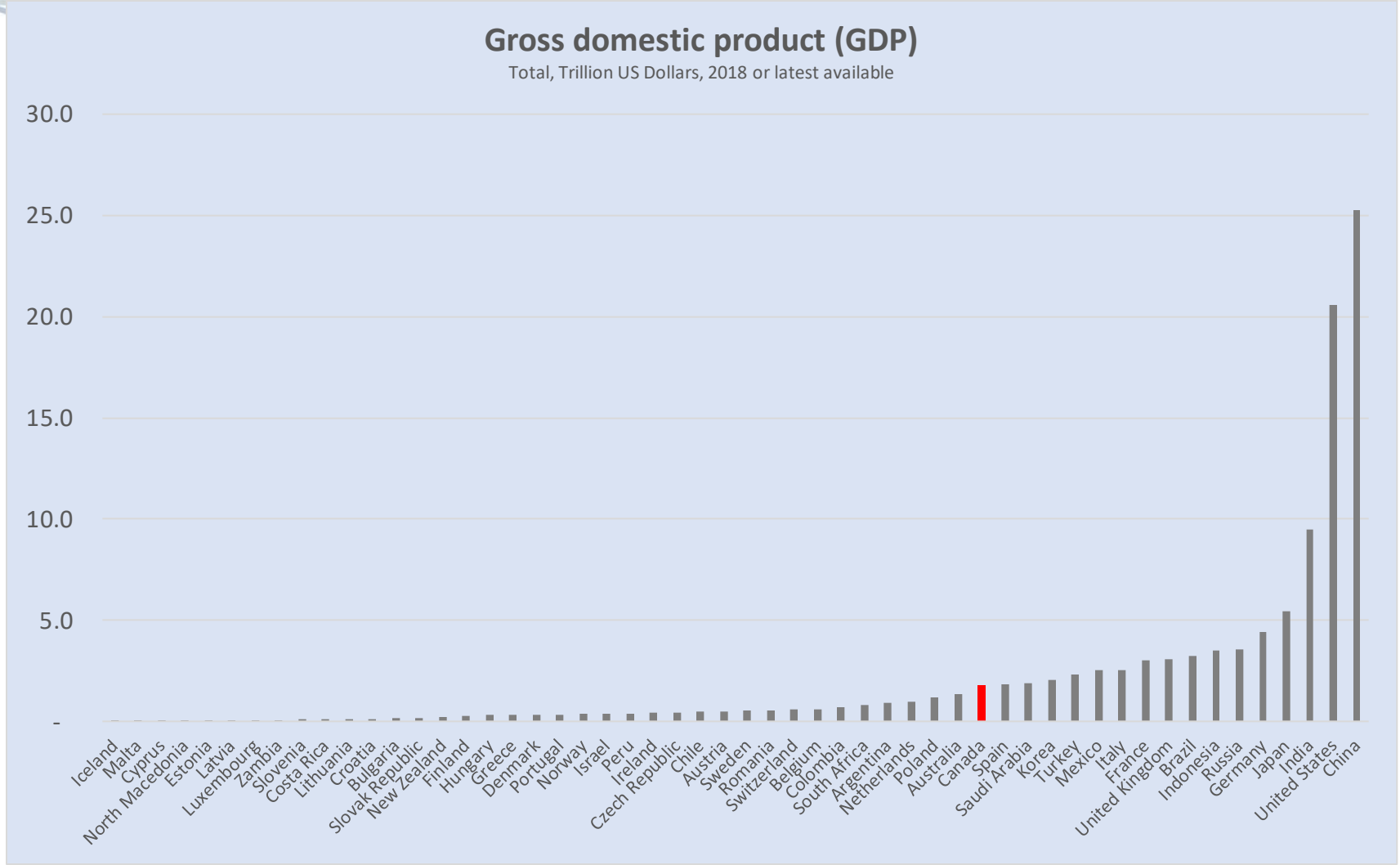
GDP is the total unduplicated value of the goods and services produced in a country or region during a given period

- Determining GDP
  - Value-added approach: gross output less intermediate inputs
  - Income approach: sum of incomes originating from productive activity
  - **Expenditure approach: sum of final expenditures**

GDP = C + I + G + X - M

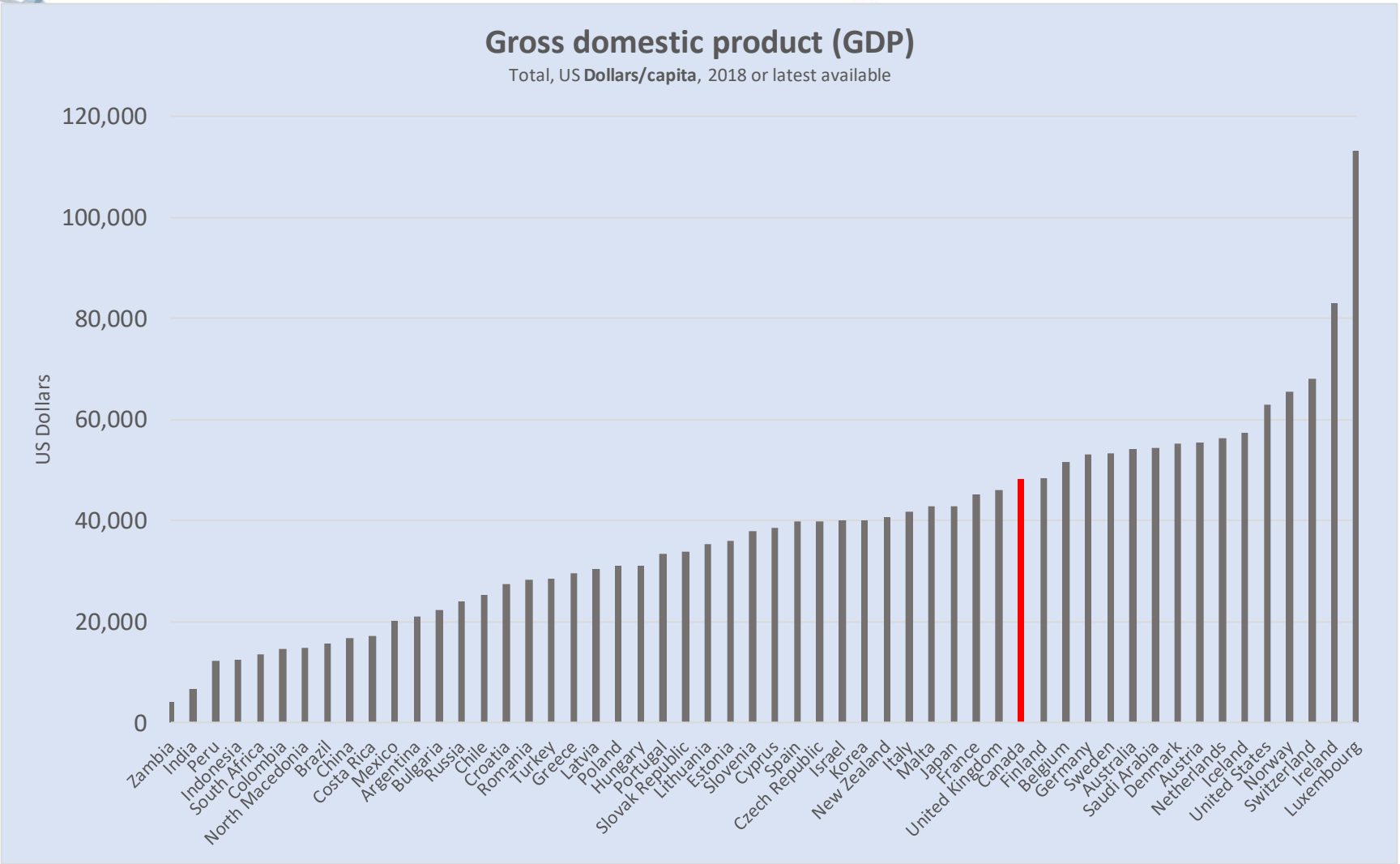
C	• **Consumption** • Private expenditures in the economy (including Manufacturing, Retail and Services)
I	• **Investment** • Compile expenditures by business on machinery, equipment, buildings and other goods (admin data)
G	• **Government spending** • Compile the financial information of all levels of government in the country
X	• **Exports** • Gross exports of goods and services
M	• **Imports** • Gross imports of goods and services

OECD countries
As an example



Source: OECD, https://data.oecd.org/gdp/gross-domestic-product-gdp.htm

OECD countries $/Capita



Source: OECD, https://data.oecd.org/gdp/gross-domestic-product-gdp.htm

Summary – Why economic statistics?

- Statistics acts
- Statistical framework
- Official statistics
- Gross Domestic Product

QUESTIONS?

Infrastructure supporting frameworks

- Economic statistics frameworks need a supporting infrastructure which includes:
  - Statistical processes
  - Classifications (industries, products, geography)
  - Administrative data
  - Registers – business, buildings and dwellings, geography
  - Methodology
  - Collection Instruments

Infrastructure – Standards

- International comparability – all countries have to follow the same ‘standards’ to allow countries and governments to compare their performance
  - **In North America**
    - North American Industry Coding System (NAICS)
    - North American Product Coding System (NAPCS)
  - **United Nations Statistics Division (UNSD)**
    - International Standard Industrial Classifications (ISIC)
    - Classification Of Individual Consumption According to Purpose (COICOP)
  - **International Standard Organisation (ISO)**
    - Countries, dependent territories and their principal subdivisions (ISO-3166)

Infrastructure – Administrative Data

- Admin data could be used to reduce response burden in business surveys
- Canada Revenue Agency (CRA) is an important provider of administrative data
  - Some examples (non exhaustive list)
    - Data from incorporated (T2) and unincorporated (T1) business tax returns
    - Data on Goods and Services Tax / Harmonized Sales Tax
    - The Business Number (BN) File
    - Data on personal income tax (T1)
    - Data on salary and wages paid to employees (T4 & PD7)
    - Data on charities (T1030) and non-profit organizations (T1044)
    - Data on partnerships (T5013)
    - Data on some tax credit programs (e.g. R&D via T661) Etc.

Infrastructure – Statistical Registers

- In order to gather information on the economy, you need to have a current and complete frame of all units being observed
  - Businesses, dwellings and geographies
- Structure of the frame aligns with definitions and classifications.
  - A Statistical Business Register (SBR) is an inventory that essentially includes all businesses operating in your areas with some of their related information.
  - Continuously updated from administrative data and survey feedback
    - Coverage (updates on active/inactive enterprises)
    - Classifications (industry, region, country of control, etc.)
    - Contact information

Infrastructure – Business Surveys

When do we lead Business Surveys to produce economic statistics?

- Surveys are carried out when there is a need for information and no admin data exists to meet this data need
  - i.e. when concepts are not collected by revenue agencies, *e.g. e-commerce, acres of crop, employee types, truck travelling distance, etc*.
- Admin data reduce surveys' costs by:
  - reducing the number of enterprises to be sampled and still get a fixed and good quality target for the official estimates (methodology)
  - being used as auxiliary information in many survey steps

Summary – Infrastructure supporting frameworks

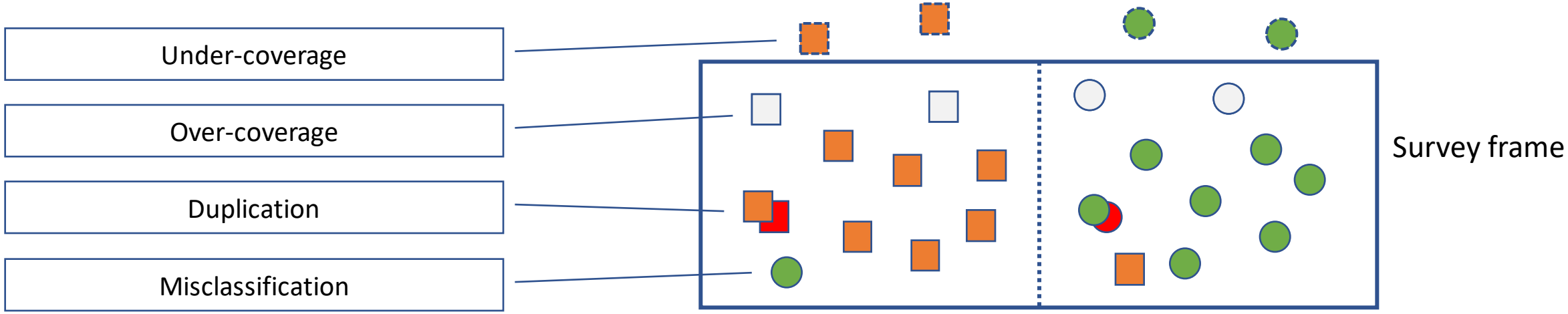
- Standards
- Administrative data
- Statistical registers
- Business surveys

QUESTIONS?

Some Quality Indicators

Frame coverage: Impact on bias, variance and efficiency

- Under-coverage: Excluding some units that belong to the population
- Over-coverage: Including some units that don’t belong to the population
- Duplication of the same unit
- Misclassification: Inaccurate values given to some variables



Some Quality Indicators

Response rates

- Number of questionnaires completed and received divided by the total number of questionnaires
- May be weighted or not, using sampling weights and/or a size variable (revenue, employment)
- Can be measured for different subgroups

- A low response rate may indicate a high risk of bias
  - But there is no direct relationship between response rate and quality

Some Quality Indicators

Coefficient of variation (CV)

- Square root of the standard deviation (square-root of the variance) divided by the estimate

- Often used to inform users about the quality of the estimates
  - Gives a size of the confidence interval around the estimates

Example of quality ratings derived from CV

CV	Quality Rating	Description
0-5%	A	Excellent
5-10%	B	Very good
10-15%	C	Good
15-25%	D	Acceptable
25-35%	E	Use with caution
More than 35%	F	Too unreliable to be published

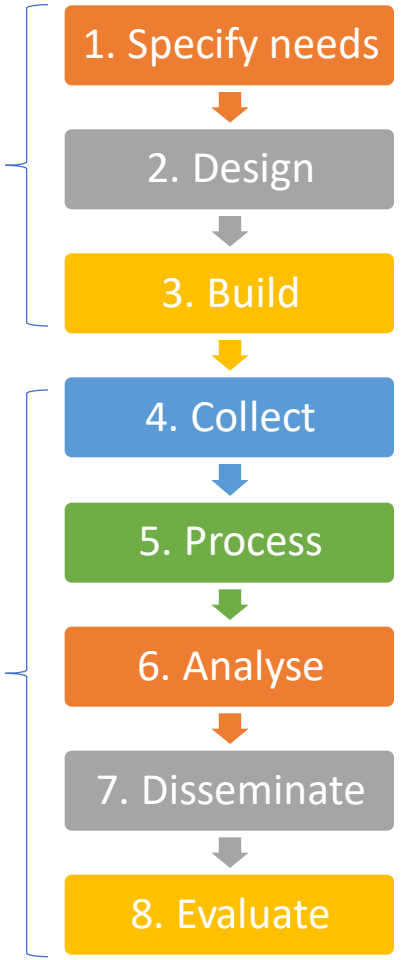
Summary – Some Quality Indicators

- Frame coverage
- Response rate
- Coefficient of variation (CV)

QUESTIONS?

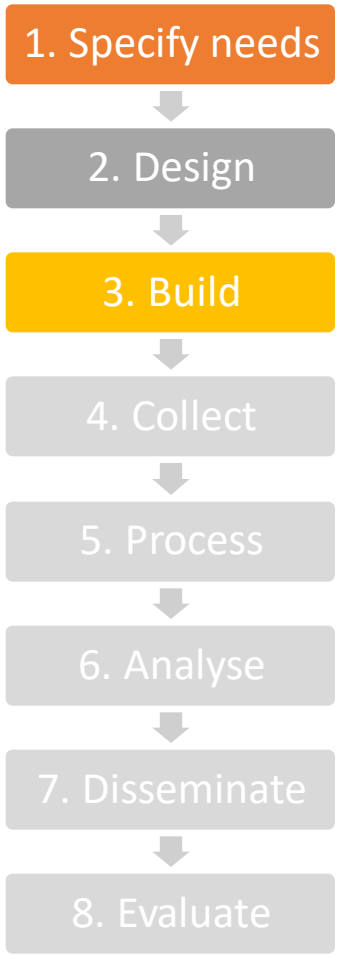
Process Model for Business Surveys

- Generic Statistical Business Process Model (GSBPM) https://statswiki.unece.org/display/GSBPM/GSBPM+v5.1
  - Applies to surveys, administrative data, or hybrid models
- Initialization (phases 1-3)
  - New survey or in review mode
- Production cycle (phases 4-8)
  - Repeated surveys: monthly, quarterly, annually, etc.



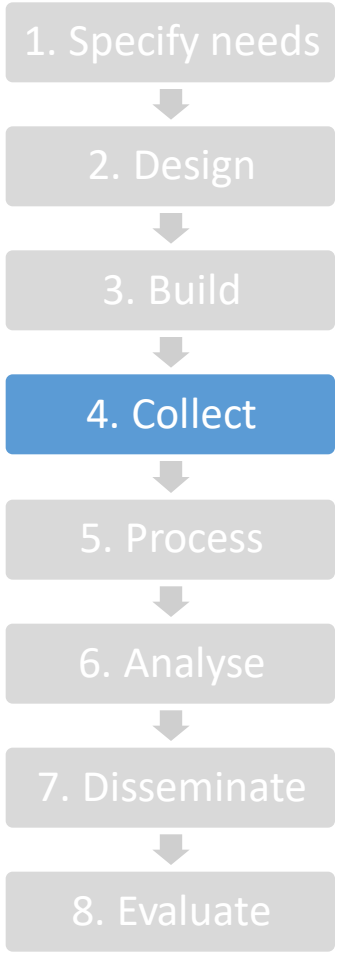
Business Surveys – Initialization

- Specify needs
  - Identify statistical needs, consult, establish objectives and identify concept, check data availability
  - Prepare a business case
- Design
  - Outputs, variables, collection, frame and sample, processing and analysis, systems and flow
- Build
  - Build or reuse or all system components
  - Test systems and process
  - Finalize production systems



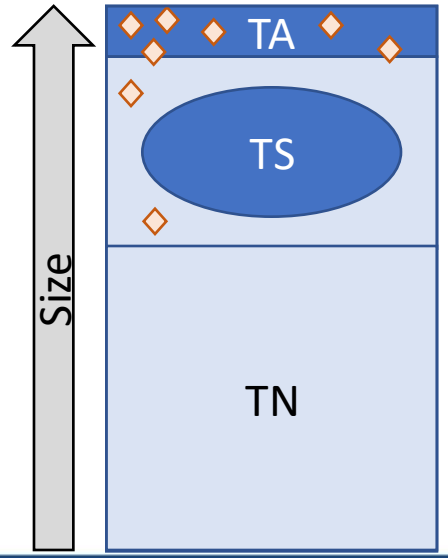
Business Surveys – Collect

- Frame
  - Starting point for sampling, data collection, and inference
  - Extraction from the business register using population definition
    - Typical criteria = Active units + Geography groupings + Industry groupings + Employment/revenue minimum thresholds
  - Key Information
    - Unit identification
    - Contact information
    - Classification variables
    - Information on frame updates
    - Linkage information



Business Surveys – Collect

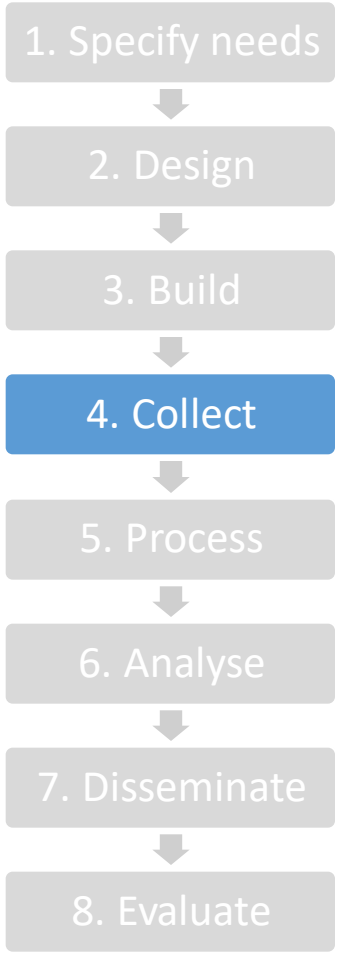
- Sampling
  - Capitalize on the skewed nature of the business population
  - Consider publication domains, typically industry and geography
- Typical strata (could differ depending on the type of survey)
  - Must take (MT): Very large units, or with unique characteristics, forced in sample
  - Take all (TA): Class of large units, selected with certainty
  - Take some (TS): Class of mid-sized units, sampled based on sample allocation
  - Take none (TN): Class of very small units, with cumulative small contribution



1. Specify needs
2. Design
3. Build
4. Collect
5. Process
6. Analyse
7. Disseminate
8. Evaluate

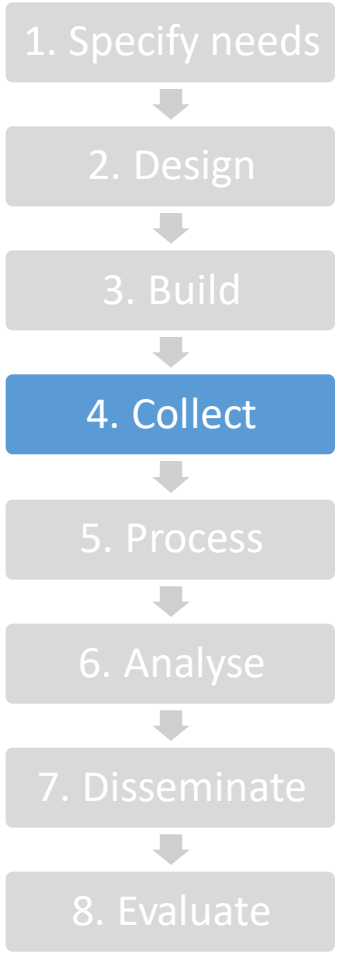
Business Surveys – Collect

- Data collection
  - Collection tools: paper or electronic questionnaire
  - Collection modes
    - Mailout/mail-back, In-person (CAPI), telephone (CATI) or internet
    - Electronic file transfers
  - Follow-up for nonresponse and/or failed-edits
    - Phone, fax, email
- High impact on collection costs, timeline, quality and response burden
  - Active collection management, or adaptive design, is a good strategy to efficiently increase response rates and improving quality while controlling costs



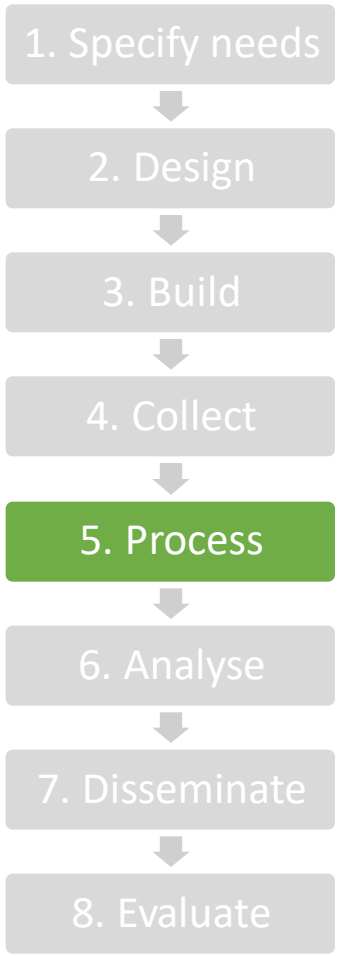
Business Surveys – Collect

- What about administrative or alternative data?
- Only phase 4. Collect changes significantly
  - Sampling and collection are replaced by a data acquisition process
  - A frame is still relevant
    - Record linkage
    - Adding classification variables
    - Assessing the coverage
    - Reduce potential bias



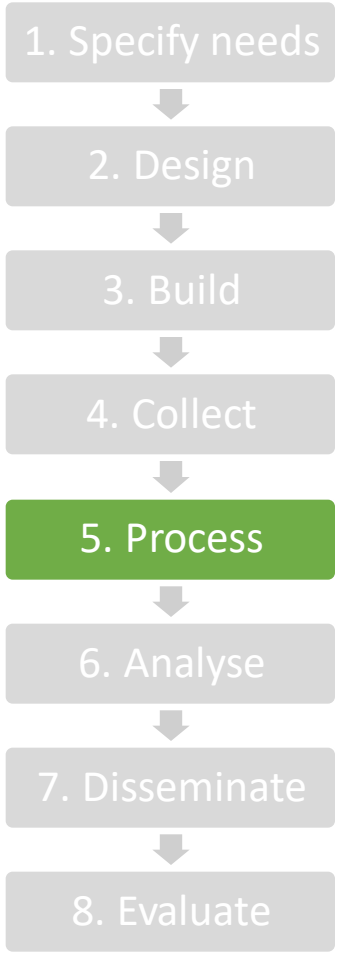
Business Surveys – Process

- Data integration
  - Combine data from collection, administrative or other sources
- Classify and code
- Editing and imputation
  - Identify errors and inconsistencies from defined rules
  - Replace missing values using various methods
    - Deterministic rules
    - Historical data, with or without trends
    - Modeling: mean, ratio, regression, etc.
    - Secondary sources, like tax data
    - Manual edits



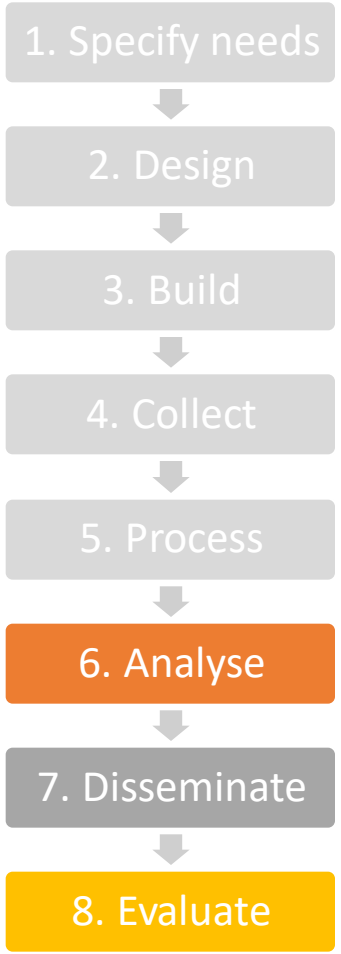
Business Surveys – Process

- Nonresponse strategy
  - Partial nonresponse: Often addressed through imputation
  - Complete nonresponse: Imputation, reweighting or calibration
- Modeling and estimation
  - Outlier detection: Units with large weighted contribution
  - Estimation weights
    - Adjustments for outliers, nonresponse and take-none portion
    - Calibration: Adjust the weights to better reflect known information about the population
  - Calculate aggregates
    - Variables of interests, domain estimates
    - Variance estimation and other quality indicators



Business Surveys – Post-Processing

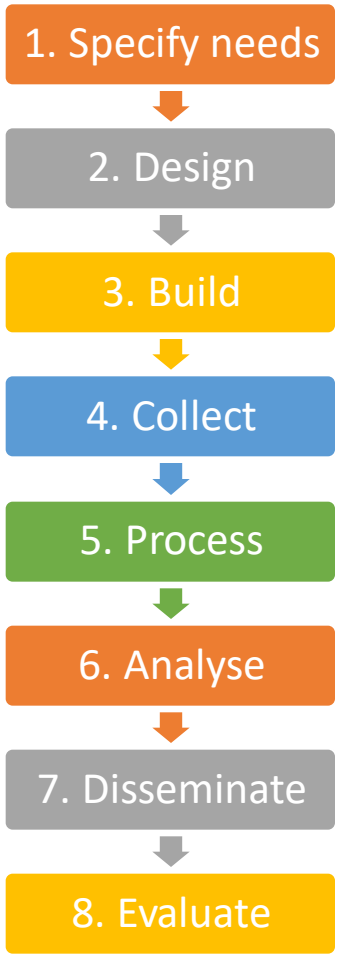
- Analyse
  - Validate, interpret and explain outputs
  - Apply disclosure control
    - Avoid disseminating domains having too few respondents
    - Avoid disseminating estimates based on dominant records
    - Admin data should be protected as they are provided at the micro level to the NSO
    - In case data is shared externally at the micro level (called microdata files), methods to protect admin and respondent data should be used (permutation, noise, etc.)
- Disseminate
- Evaluate



Summary – Process Model for Business Surveys

- Initialization
- Collect: Frame, sampling, data collection/acquisition
- Process: Data integration, classify and code, editing and imputation, nonresponse strategy, modeling and imputation
- Post-processing, including disclosure control

QUESTIONS?



Conclusion

- Economic statistics are essential to make evidence-based decisions for policies, GDP calculation, studies on society/population/enterprises, etc.
- Solid infrastructure is key to support a statistical framework
- Surveys are reliable vehicles for the production of statistics
  - Various steps, from initialization to production cycles
  - Well-tested methods and well-adapted tools

Thank you all! Questions?

You can contact the PRASC team at:

prasc@statcan.gc.ca

Delivering insight through data for a better Canada