

Understanding data and statistics better – for more effective SDG decision making

Massive open online course for policy-makers worldwide

Course Syllabus

Course duration: 9 April – 18 December 2020



Table of contents

Table of contents	2
Background	3
Target audience	3
Learning objectives	3
Structure	3
Learning methodology	4
Module 1. Statistical Literacy	4
Module 2. Interpreting and assessing the fitness for purpose of data	5
Module 3. Communicating with data	6
Module 4. Data sources	7
Module 5. Policy formulation, monitoring and evaluation	7
Certification and feedback	8
Technical requirements	9

Background

This course¹ aims to offer a double opportunity: i) to help improve the understanding of how better data can help design more effective public policies and ii) to brush up on your core data skills. The course covers aspects related to the interpretation, analysis, and presentation of data of direct relevance to the SDGs and more broadly national and sectoral policies aimed at ensuring sustainable and inclusive development. It introduces these topics using concrete examples of data use and public policy context from around the world. The course reviews both traditional and non-traditional data sources and discusses their uses, as well as the opportunities and challenges presented by the latter. It uses a combination of story-telling, examples, videos and hands-on exercises to tailor to a variety of learners.

Target audience

The course is intended for public officials in charge of SDG policies, as well as data producers and anyone interested to learn more about the role of data in public policy design. National Statistical Offices may also find it a useful resource in their work aimed at strengthening data skills and data demand from their countries' decision-makers from across line ministries as well as in their efforts to further the understanding of data producers from across the National Statistical System and beyond when it comes to data use and data demand.

Participants can follow the course at their own pace, and can earn a certificate provided they have met the certification requirements described in the “Certification and feedback” section.

Learning objectives

It is expected that, by the end of the course, participants will be better positioned to:

- 1) Explain the importance of timely, quality and disaggregated data for informing policies and decisions in support of the SDGs;
- 2) Interpret and assess the fitness-for-purpose of different sources of data depending on the objective and task at hand;
- 3) Effectively and correctly communicate data to tell a story using tables, graphs and maps;
- 4) Identify potential data sources from among official, scientific and new types of data for monitoring SDG indicators using a set of criteria;
- 5) Discuss how data should be used to formulate evidence-based policies or conduct M&E activities.

Structure

Before you start the course, you will be invited to take a short survey and read this syllabus. The course is organized in 5 modules. Module 1 serves as an introduction, discussing the role of data in evidence-based policy-making. It is followed by Modules 2 and 3 that aim to strengthen learner's competences related to interpreting and assessing the fitness for purpose of data and communicating with data telling stories with tables and graphs. Module 4 provides a comprehensive coverage of various data sources and discusses how they can be leveraged for policy design. Module 5 covers the aspects related the use of data in policy formulation and monitoring and evaluation from policy perspective. Each module is followed by a module test. After you have completed 5 modules, you will be invited to take the final assessment and complete the feedback questionnaire.

¹ Among persons having provided substantive contributions to the development of this course are Sharleen Forbes, John Harraway, Helen McGivallray, Aniket Bhushan, Vibeke Nielsen, Leandre Ngogang Wandji and Elena Proden.

Short introductory survey

Module 1. Statistical literacy

Module 2. Interpreting and assessing the fitness for purpose of data

Module 3. Communicating with data

Module 4. Data sources for SDG indicators

Module 5. Policy formulation, monitoring and evaluation

Final assessment

Feedback survey

Learning methodology

This self-paced course consists of five modules developed around specific learning objectives, and illustrated through multiple case studies. The modules will be rolled out progressively. The estimated learning time per module ranges from 2 to 5 hours. In addition to case studies and learning videos, modules also contain a number of study questions and hands-on exercises that aim to enable a learner better retain the presented information and practice a number of key skills. Each module is followed by an objective assessment consisting of 10 randomly selected questions.

Module 1. Statistical Literacy

Module 1 serves as an introduction, discussing the role of data in evidence-based policy-making. The module will illustrate the importance of data in informing policies in selected SDG areas providing both historical and contemporary examples. It will also offer a data refresher to those who need to brush up on basic data skills ranging from distinguishing between different data sources, using basic measurements, and understanding the notion of probability. It will also focus on practical skills related to the reading of tables and graphs summarizing data, understanding summary statistics describing your variable of interest – a set of data values describing a given feature - and measures of association between two variables.

Learning objectives

It is expected that, by the end of this module, participants will be better positioned to:

1. Provide concrete examples of how data, including disaggregated data, can help inform better policies for SDGs;
2. Correctly interpret and use key data measurements and notions relevant in policy contexts;
3. Explain what types of tables and graphs should be used for what purposes when summarizing data;
4. Use summary statistics to describe a data set and measures of association to describe the relationship between two variables.

Structure

This module alternates stories with theory accompanied by exercises:

Moodle:

Ex-ante test 1

Lesson 1:

Story 1. Moving the world from 'sick and poor' to 'healthy and wealthy'

Brush up on your data skills

- a. Data types
- b. Calculating measures
- c. Working with percentage change

Story 2. Estimating the Global Burden of Disease: Rotovirus

- d. Probability
- e. Summarizing data
- f. Using statistics to summarize data

Story 3. What's wrong with smoking? One country's 180° turn in 50 years of smoking policies

- e. Relationships between two variables

Moodle:

Module 1 test

Module 2. Interpreting and assessing the fitness for purpose of data

This module first introduces key concepts related to quality assurance and key principles guiding the production of official statistics. It then describes different data sources distinguishing between those expected to describe the entire population of interest (e.g., country population), those that are based on sample surveys and those that can be referred to as other data sources, including integrated data sets and geospatial imagery but also non-traditional data sources such as Big Data and other. The module highlights the advantages and disadvantages of these broad categories of data used or considered for future use in public policy design. It also provides an overview of sample survey techniques, estimation and hypothesis testing to help improve a learner's understanding of the meaning and limitations of a sample survey data, as well as correctly interpret statistics from published data, reports and studies in the context of public policy analysis and policy choice. A more detailed typology of data used by National Statistical Offices and the examples of use on non-traditional data sources by National Statistical Offices to fill in data gaps, address timeliness and granularity requirements of the SDGs is then reviewed in more detail in Module 4.

Learning objectives

It is expected that, by the end of this module, participants will be better positioned to:

1. Assess the fitness for purpose of census, administrative and sample survey data;
2. Define key ideas in sample survey techniques, estimation and hypothesis testing;
3. Use statistics from published data, reports and studies to answer a given public sector question.

Structure

This module comprises two parts with case studies built inside each.

Moodle:

Ex-ante test 2

Lesson 2:

- a. Using metadata to understand key characteristics of a dataset
- b. Understanding data from census and government administrative sources
- c. Understanding sampling techniques to assess survey data
- d. Applying InFocus: Practicing data assessment and interpretation of results

Moodle:

Module 2 test

Module 3. Communicating with data

The module aims to look at ways of communicating the results of data explorations and statistical analyses with a variety of audiences, primarily policy and other decision makers but also the media and the general public who can be powerful influencers of public policy. The first provides an introduction into the use of story-telling techniques, while the second and the third deal with the use of table and graphs respectively. The last part looks at maps as effective communication means. The module covers the “dos” and the “don'ts” of data communication providing a number of very practical tips and concrete examples. It aims to help strengthen learners' capacities to use data in a way that is scientifically correct but also highly relevant to respective audiences' needs and easy to understand for those with little data or statistical background.

Learning objectives

1. Use story-telling techniques to communicate data in a way that is clear to different audiences, impactful and accurate;
2. Describe key results from reading tables, graphs and maps in an effective manner;
3. Build impactful and correct tables and graphs to illustrate data and findings.

Structure

This module comprises two parts:

Moodle:

Ex-ante test 3

Lesson 3:

- a. Telling a story with words
- b. Using tables
- c. Visualising data with graphs
- d. Visualising data with maps

Moodle:

Module 3 test

Module 4. Data sources for SDG indicators

Module 4 aims to investigate possible sources of data that could be used to inform indicators developed for the Sustainable Development Goals. The key issues that need to be taken into account are the quality of the data, and level of access to the data. Part 1 covers National Statistical Offices (NSOs) and most of the International Organizations as the major providers of long-term data about people and their social and economic activities. NSOs have little data about physical attributes of the planet (e.g. climate) or animal and plant specific data, and there are only a few specific international agencies that deal with this type of data so we are often reliant on research and scientific sources covered in Part 2. Part 3 covers new or non-traditional data sources that present emerging opportunities, as well as challenges for monitoring progress on SDGs and public policies.

Learning objectives

1. Apply a set of criteria to examine what data source can be used for monitoring a given SDG indicator;
2. Recognize main data providers and data sources, including those from National Statistical Offices and International Organizations, on demographic and socio-economic characteristics;
3. Identify and evaluate the fitness for purpose of relevant research and scientific data for informing selected policy decisions;
4. Discuss the opportunities and risks associated with the use of non-traditional data sources for monitoring selected SDGs using concrete examples.

Structure

This module comprises three parts:

Moodle:

Ex-ante test 4

Lesson 4:

- a. Understanding data sources used for various SDG indicators
- b. Research and Scientific data
- c. Non-traditional data sources

Moodle:

Module 4 test

Module 5. Policy formulation, monitoring and evaluation

Module 5 covers data use for policy formulation and the evaluation of the impact of interventions for the SDGs 1-12. Part 1 of this module discusses the dynamics of population (how they grow or decline) and their importance in policy formulation, i.e., the age groups future policies should be targeted to. It reviews how to take different age structures into account when comparing populations (either sub-groups within countries or between countries), and how to evaluate the relative risk incurred in

different groups of the population. Part 2 gives a brief look at price indices used to measure inflation, that needs to be taken into account when costing policies or interventions, and provides a brief overview of the System of National Accounts. Part 3 gives an overview of data use for monitoring and evaluation, including for monitoring changes over time in key outcomes (to see if an intervention is having an overall impact) and for investigating the relationships between variables (e.g. demographic, ethnic or regional variables and social or economic variables) as evidence base for designing new or amending existing policy interventions. It is often the role of the research, scientific and evaluation communities to analyse public policies and programmes and examine the short- (and sometimes longer-) term impacts.

Learning objectives

1. Assess implications of a given population structure for designing various policies;
2. Describe how inflation is measured and can be factored into the costing of policy interventions;
3. Explain how data are and can be used in the context of monitoring and evaluating the impacts of public policies and programmes with the use of specific examples.

Structure

This module comprises three parts:

Moodle:

Ex-ante test 5

- a. Using data to understand and compare populations and develop policy interventions
- b. Measuring inflation and other economic measures
- c. Using data in national monitoring and evaluation

Moodle:

Module 5 test

Certification and feedback

Each of the five modules is preceded by an ex-ante assessment to evaluate your baseline knowledge and skills and is followed by an objective assessment or a quiz. Three attempts are allowed per quiz. The attempt with the highest score is retained as the grade for the module. Each attempt proposes 10 questions drawn at random from the respective question bank. For each question, the correct responses add up to 1 point. Each incorrect response is penalized individually by -0.1 point. Participants would need to obtain a minimum of 7 points out 10 points for the module quiz to complete the module assessment requirements.

Those participants who complete the requirements for at least 4 out of 5 modules are entitled to a certificate of participation. Those participants who meet the above requirement for all of the 5 modules and pass the final test can obtain a certificate of completion and special badge “Foundational data skills for the SDGs”.

Participants will be requested to provide feedback on the course by filling in a feedback form which can be accessed in the ‘Course Evaluation’ section on the course home page.

Technical requirements

Browser:

- The course works best with Firefox 3.6 or higher (download for free at <http://www.mozilla-europe.org/en/firefox>)
- The course is also compatible with Google Chrome (download for free at <https://www.google.com/intl/en/chrome/browser/>)
- It is not recommended to use Internet Explorer for technical reasons
- Note JavaScript & Cookies must be enabled

Software:

- Adobe Acrobat Reader (download for free at <http://www.adobe.com/products/acrobat/readstep2.html>).
- Adobe Flash Player (download for free at <http://get.adobe.com/flashplayer>)
- Microsoft Office (Windows or Apple version) or Open Office (download for free at <http://www.openoffice.org>)

Platform: Windows 95, 98, 2000, NT, ME, XP or superior; MacOS 9 or MacOS X

Hardware: 64 MB of RAM, 1 GB of free disk space

Modem: 56 K