MINISTERE DE L'ECONOMIE NUMERIQUE ET DE LA POSTE

ARTCI

Connecting the World

International Telecommunication Union

ITU Regional Workshop on ICT Statistics for Africa

Abidjan, Côte d'Ivoire

6-8 May 2019

Overview of household ICT indicators, survey design, data processing and data dissemination

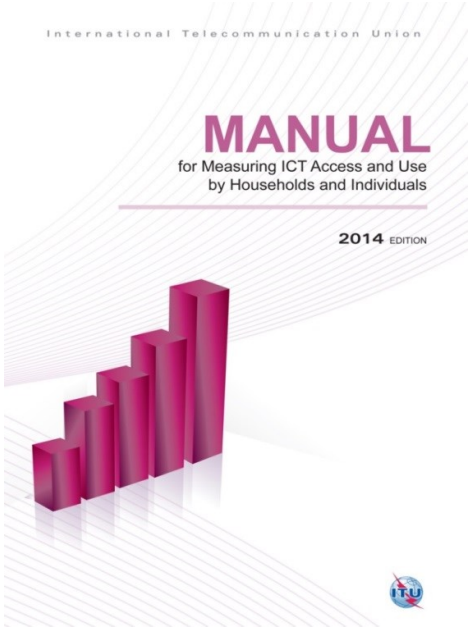
Martin Schaaper
Senior ICT Analyst
ICT Data and Statistics Division
Telecommunication Development Bureau
International Telecommunication Union

OVERVIEW OF HOUSEHOLD ICT INDICATORS

Tashkent, 12-14 February 2019

ITU Manual (2014)

- Chapter 1. Introduction
- Chapter 2. **Coordination** among national stakeholders in ICT measurement
- Chapter 3. **Planning and preparation** for ICT household surveys
- Chapter 4. **Statistical standards** and measurement topics for ICT household statistics
- Chapter 5**. Data sources** and **collection techniques** for ICT household statistics
- Chapter 6. Question and **questionnaire** design for ICT household surveys
- Chapter 7. **Designing** ICT household surveys
- Chapter 8. **Data processing** for ICT household statistics
- Chapter 9. **Data quality and evaluation** for ICT household statistics
- Chapter 10. **Dissemination** of ICT household data and metadata



Preparation and revision process

- First release in 2009
- Version 2 launched at WTIS 2013 (December 2013, Mexico)
- Revision of indicators in 2014-2015:
  - added HH16
  - HH17, HH18, HH19 **not yet in the Manual**
- More revisions made in subsequent years: all published on the ITU website

ITU statistical standards: ICT household statistics

- Statistical standards associated with the **core ICT indicators** for household **access** to, and individual **use** of, ICT:
  - concepts
  - definitions of terms
  - model questions
  - classificatory variables (breakdowns)
  - scope
  - units (households and individuals)
- Formula of calculation
- Use (policy relevance)

Core household indicators, main concepts

- The indicators consist of those:
  - Referring to household access to ICT equipment and services
  - Referring to individuals' use/ownership of ICT equipment and services

Concept of access

- ICT device/service should be available for use of any member of the household at any time
- Device can be owned or not by the household
- Applies to all indicators referring to household ICT access
- Device should be in a working condition

Core ICT HH indicators (2016 rev.)

HH1	Proportion of households with a radio
HH2	Proportion of households with a television
HH3	Proportion of households with telephone
HH4	Proportion of households with a computer
HH5	Proportion of individuals using a computer
HH6	Proportion of households with Internet
HH7	Proportion of individuals using the Internet
HH8	Proportion of individuals using the Internet, by location
HH9	Proportion of individuals using the Internet, by type of activity
HH10	Proportion of individuals using a mobile cellular telephone
HH11	Proportion of households with Internet, by type of service
HH12	Proportion of individuals using the Internet, by frequency
HH13	Proportion of households with multichannel television, by type
HH14	Barriers to household Internet access
HH15	Individuals with ICT skills, by type of skills
HH16	Household expenditure on ICT
HH17	Proportion of individuals using the Internet, by type of portable device and network used to access the Internet
HH18	Proportion of individuals who own a mobile phone
HH19	Proportion of individuals not using the Internet, by type of reason

Core ICT HH indicators (2016 rev.)

HH1	Proportion of households with a radio
HH2	Proportion of households with a television
HH3	Proportion of households with telephone
HH4	Proportion of households with a computer
HH5	Proportion of individuals using a computer
HH6	Proportion of households with Internet
HH7	Proportion of individuals using the Internet
HH8	Proportion of individuals using the Internet, by location
HH9	Proportion of individuals using the Internet, by type of activity
HH10	Proportion of individuals using a mobile cellular telephone
HH11	Proportion of households with Internet, by type of service
HH12	Proportion of individuals using the Internet, by frequency
HH13	Proportion of households with multichannel television, by type
HH14	Barriers to household Internet access
HH15	Individuals with ICT skills, by type of skills
HH16	Household expenditure on ICT
HH17	Proportion of individuals using the Internet, by type of portable device and network used to access the Internet
HH18	Proportion of individuals who own a mobile phone
HH19	Proportion of individuals not using the Internet, by type of reason

Indicator HH1: Proportion of households with a radio

Definitions:

This is the proportion of households that have a radio.

A *radio* is defined as a device capable of receiving broadcast radio signals, using common frequencies, such as FM, AM, LW and SW. A radio may be a stand-alone device, or it may be integrated with another device, such as an alarm clock, an audio player, a mobile telephone or a computer.

Indicator HH2: Proportion of households with a television

Definitions:

This is the proportion of households that have a television (TV).

A *television (TV)* is a device capable of receiving broadcast television signals, using popular access means such as over-the-air, cable and satellite. A television set is typically a stand-alone device, but it may also be integrated with another device, such as a computer or a mobile telephone.

Indicator HH3: Proportion of households with telephone

Definitions:

This is the proportion of households that have a telephone.

A *fixed telephone line* refers to a telephone line connecting a customer's terminal equipment (e.g. telephone set, facsimile machine) to the public switched telephone network (PSTN) and which has a dedicated port on a telephone exchange. This term is synonymous with the terms *main station* or *Direct Exchange Line* (DEL) that are commonly used in telecommunication documents. It may not be the same as an access line or a subscription.

A *mobile (cellular) telephone* refers to a portable telephone subscribing to a public mobile telephone service using cellular technology, which provides access to the PSTN. This includes analogue and digital cellular systems and technologies such as IMT-2000 (3G) and IMT-Advanced. Users of both postpaid subscriptions and prepaid accounts are included.

Indicator HH4: Proportion of households with a computer

Definitions:

This is the proportion of households that have a computer.

A *computer* refers to a desktop computer, a laptop (portable) computer or a tablet (or similar handheld computer).

- Desktop: a computer that usually remains fixed in one place; normally the user is placed in front of it, behind the keyboard.
- Laptop (portable) computer: a computer that is small enough to carry and usually enables the same tasks as a desktop computer; it includes notebooks and netbooks but does not include tablets and similar handheld computers.
- Tablet (or similar handheld computer): a tablet is a computer that is integrated into a flat touch screen, operated by touching the screen rather than (or as well as) using a physical keyboard.

It does not include equipment with some embedded computing abilities, such as smart TV sets, and devices with telephony as their primary function, such as smartphones.

Indicator HH5: Proportion of individuals using a computer

Definitions:

This is the proportion of individuals who used a computer from any location in the last three months.

A *computer* refers to a desktop computer, a laptop (portable) computer or a tablet (or similar handheld computer).

- Desktop: a computer that usually remains fixed in one place; normally the user is placed in front of it, behind the keyboard.
- Laptop (portable) computer: a computer that is small enough to carry and usually enables the same tasks as a desktop computer; it includes notebooks and netbooks but does not include tablets and similar handheld computers.
- Tablet (or similar handheld computer): a tablet is a computer that is integrated into a flat touch screen, operated by touching the screen rather than (or as well as) using a physical keyboard.

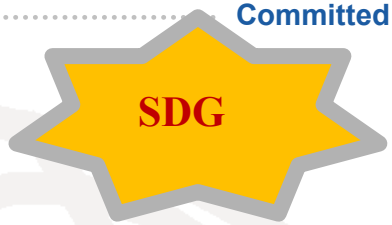
It does not include equipment with some embedded computing abilities, such as smart TV sets, and devices with telephony as their primary function, such as smartphones.

Indicator HH6: Proportion of households with Internet

Definitions:

This is the proportion of households with Internet access at home.

The *Internet* is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer – it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.). Access can be via a fixed or mobile network.



Indicator HH7: Proportion of individuals using the Internet

Definitions:

This is the proportion of individuals who used the Internet from any location in the last three months.

The *Internet* is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer – it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.). Access can be via a fixed or mobile network.

Indicator HH8: Proportion of individuals using the Internet, by location

Definitions:

This is the proportion of individuals who used the Internet from specified locations in the last three months.

The *Internet* is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer – it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.). Access can be via a fixed or mobile network, including wireless access at a WiFi 'hotspot'.

Access via a mobile device should be classified to the appropriate location or to 'while commuting, in transport or walking', that is while mobile.

Locations of Internet use are defined as follows:

- Home
- Work
- Place of education
- Another person's home
- Facility open to the public
  - *Of which: Community Internet access facility*
- While commuting, in transport or walking

Revised by EGH in 2018

Indicator HH9: Proportion of individuals using the Internet, by type of activity

Definitions:

This is the proportion of individuals who undertook one or more activities using the Internet for private (defined as non-work) purposes from any location in the last three months. Internet activities are classified in groups of similar activities, and are defined as follows:

Access to information

- Getting information about goods or services
- Seeking health-related information (on injury, disease, nutrition etc.).
- Getting information from general government organizations
- *For more activities, see the ITU website with the revised indicator*

Communication and collaboration

- Telephoning over the Internet/VoIP (using Skype, iTalk, etc.; includes video calls via webcam)
- *For more activities, see the ITU website with the revised indicator*

Electronic commerce, trade, and transactions

- *For the detailed activities, see the ITU website with the revised indicator*

Learning

- *For the detailed activities, see the ITU website with the revised indicator*

Professional life

- *For the detailed activities, see the ITU website with the revised indicator*

Digital content consumption

- *For the detailed activities, see the ITU website with the revised indicator*

Digital content creation

- *For the detailed activities, see the ITU website with the revised indicator*

Revised by EGH in 2018

Indicator HH10: Proportion of individuals using a mobile cellular telephone

Definitions:

This is the proportion of individuals who used a mobile telephone in the last three months.

A *mobile (cellular) telephone* refers to a portable telephone subscribing to a public mobile telephone service using cellular technology, which provides access to the PSTN. This includes analogue and digital cellular systems and technologies such as IMT-2000 (3G) and IMT-Advanced. Users of both postpaid subscriptions and prepaid accounts are included.

Indicator HH11: Proportion of households with Internet, by type of service

Definitions:

This is the proportion of households with access to the Internet, by type of service.

The *Internet* is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer – it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.). Access can be via a fixed or mobile network.

The broad types of Internet services to be identified are the following:

- Fixed (wired) narrowband network: includes analogue modem (dial-up via standard telephone line), ISDN (Integrated Services Digital Network), DSL (Digital Subscriber Line) at advertised download speeds below 256 kbit/s, and other forms of access with an advertised download speed of less than 256 kbit/s
- Fixed (wired) broadband network: refers to technologies at advertised download speeds of at least 256 kbit/s, such as DSL, cable modem, high speed leased lines, fibre-to-the-home/building, powerline and other fixed (wired) broadband
- Terrestrial fixed (wireless) broadband network: refers to technologies at advertised download speeds of at least 256 kbit/s, such as WiMAX, fixed CDMA
- Satellite broadband network (via a satellite connection), at advertised download speeds of at least 256 kbit/s
- Mobile broadband network (at least 3G, e.g. UMTS) via a handset
- Mobile broadband network (at least 3G, e.g. UMTS) via a card (e.g. integrated SIM card in a computer) or USB modem

Indicator HH12: Proportion of individuals using the Internet, by frequency

Definitions and notes:

This is the frequency of Internet use by individuals who used the Internet from any location in the last three months.

The *Internet* is a worldwide public computer network. It provides access to a number of communication services including the World Wide Web and carries e-mail, news, entertainment and data files, irrespective of the device used (not assumed to be only via a computer – it may also be by mobile telephone, tablet, PDA, games machine, digital TV etc.). Access can be via a fixed or mobile network.

Frequency of use categories are as follows:

- At least once a day: once a working day for respondents who only (or most frequently) use the Internet from work or school etc.
- At least once a week but not every day
- Less than once a week.

Indicator HH13: Proportion of households with multichannel television, by type

Definitions:

This is the proportion of households with multichannel television (TV) and by type of multichannel service. Multichannel TV services are as follows:

- Cable TV (CATV): multichannel programming delivered over a coaxial cable for viewing on television sets
- Direct-to-home (DTH) satellite services: TV services received via a satellite dish capable of receiving satellite television broadcasts
- Internet-protocol TV (IPTV): multimedia services such as television/video/audio/text/graphics/data delivered over an IP-based network managed to support the required level of quality of service, quality of experience, security, interactivity and reliability; it does not include video accessed over the public Internet, for example, by streaming. IPTV services are also generally aimed at viewing over a television set rather than a personal computer.
- Digital terrestrial TV (DTT): the technological evolution from analogue terrestrial television, providing capability for significantly more channels

Indicator HH14: Barriers to household Internet access

Definitions:

This measures the barriers to Internet access for households without Internet access. It is expressed as a proportion of households without Internet access.

Barriers (that is, reasons for not having Internet) are:

- Do not need the Internet (not useful, not interesting)
- Do not know how to use it
- Cost of Internet use is too high (service charges, etc.)
- Privacy or security concerns
- Internet service is not available in the area
- Cultural reasons (e.g. exposure to harmful content)
- Don’t know what Internet is
- Not allowed to use the Internet
- Lack of local content
- Other reason, specify

Revised by EGH in 2015

Indicator HH15: Individuals with ICT skills, by type of skills

Definitions:

This refers to ICT skills, defined for the purpose of this indicator as individuals having undertaken certain activities in the last three months, independent of the device(s) used.

Activities to measure ICT skills are as follows:

- Using copy and paste tools to duplicate or move data, information and content in digital environments (e.g. within a document, between devices, on the cloud)
- Sending messages (e.g. e-mail, messaging service, SMS) with attached files (e.g. document, picture, video)
- Using basic arithmetic formulae in a spreadsheet
- Connecting and installing new devices (e.g. a modem, camera, printer) through wired or wireless technologies
- Finding, downloading, installing and configuring software and apps
- Creating electronic presentations with presentation software (including text, images, sound, video or charts)
- Transferring files or applications between devices (including via cloud-storage)
- Setting up effective security measures (e.g. strong passwords, log-in attempt notification) to protect devices and online accounts
- Changing privacy settings on your device, account or app to limit the sharing of personal data and information (e.g. name, contact information, photos)
- Programming or coding in digital environments (e.g. computer software, app development)

Revised by EGH in 2018

Indicator HH16: Household expenditure on ICT

Definitions:

This measures the percentage of total household expenditure that is expended on ICT goods and services as follows:

- Telephone and telefax equipment (COICOP 08.2.0): Purchases of telephones, radio-telephones, telefax machines, telephone-answering machines and telephone loudspeakers; repair of such equipment.
- Telephone and telefax services (COICOP 08.3.0): Installation and subscription costs of personal telephone equipment; includes telephone calls (from any location), information transmission services, Internet connection services, hire of telephones.
- Equipment for the reception, recording and reproduction of sound and picture (COICOP 09.1.1): Television sets, video cassette players and recorders, television aerials of all types; radio sets, car radios, radio clocks, two-way radios, amateur radio receivers and transmitters; gramophones, tape players and recorders, cassette players and recorders, CD-players, personal stereos, stereo systems and their constituent units (turntables, tuners, amplifiers, speakers, etc.), microphones and earphones.
- Information processing equipment (COICOP 09.1.3): Personal computers, visual display units, printers and miscellaneous accessories accompanying them; computer software packages such as operating systems, applications, languages, etc.; calculators, including pocket calculators; typewriters and word processors. (Also includes laptops, tablets, e-book readers.)
- Repair of audio-visual, photographic and information processing equipment (COICOP 09.1.5)

HH17: Individuals using the Internet, by type of portable device and network used to access the Internet

a. Mobile phone

a1) via mobile cellular network

a2) via other wireless networks (e.g. WiFi)

b. Tablet

b1) via mobile cellular network, using USB key/dongle or integrated data SIM card

b2) via other wireless networks (e.g. WiFi)

c. Portable computer (laptop, notebook, netbook)

c1) via mobile cellular network, using USB key/dongle or integrated data SIM card or mobile cellular telephone as modem

c2) via other wireless networks (e.g. WiFi)

d. Other portable devices (e.g. portable games consoles, watches, e-book readers etc.)

First collected in 2015 (agreed by EGH in 2014)

HH18: Proportion of individuals who own a mobile phone

This is the proportion of individuals who own a mobile phone. An individual owns a mobile cellular phone if he/she has a mobile cellular phone device with at least one active SIM card for personal use. It includes mobile cellular phones supplied by employers that can be used for personal reasons (to make personal calls, access the Internet, etc.) and those who have a mobile phone for personal use that is not registered under his/her name. It excludes individuals who have only active SIM card(s) and not a mobile phone device.

First collected in 2015 (agreed by EGH in 2014)

HH19: Proportion of individuals not using the Internet, by type of reasons

Response categories:

- *Do not need the Internet (not useful, not interesting)*
- *Do not know how to use it*
- *Cost of Internet use is too high (service charges, etc.)*
- *Privacy or security concerns*
- *Internet service is not available in the area*
- *Cultural reasons (e.g. exposure to harmful content)*
- *Don't know what Internet is*
- *Not allowed to use the Internet*
- *Lack of local content*
- *Other reason, specify*

First collected in 2016 (agreed by EGH in 2015)

Age scope

- Countries should report ICT usage information for the three main core indicators on individuals' use of ICTs (computer, mobile phone and Internet) for the entire population of the country, i.e. there is no minimum age scope any more for these indicators

Reference period

- Information on ICT usage should be collected and reported with a reference period of the last 3 months

DATA DISAGGREGATIONS

Disaggregating the data by socio-demographics: why and how

- Important to policy-makers
- Disaggregation shows socio-economic problems that create barriers to use of ICT by individuals. These problems are diverse and broadly cover lack of opportunity and lack of ability. They include illiteracy and other linguistic limitations, socio-cultural barriers, lack of ICT and other skills, lack of confidence or awareness and low income.
- Gives more information i.e. who is using the ICTs i.e. male/ female, age, location (urban/ rural) etc

Individual characteristics

Sex:

- Sex disaggregation of data is a fundamental requirement for gender statistics and in particular for the analysis of the gender gap in the use of ICT. A MUST HAVE FOR ALL CORE INDICATORS

Age:

- Age is a strong determinant of ICT use so a common age cut-off and categories are important

- *Recommended ranges: under 5; 5–9; 10–14; 15–24; 25–34; 35–44; 45–54; 55–64; 65–74 and 75 and over*

Education levels:

For international comparisons, countries required to classify education as International Standards Classification of Education follows:

- primary education or lower (ISCED levels 0, 1),
- lower secondary education (ISCED level 2),
- upper secondary education or post-secondary non-tertiary education (ISCED
- levels 3,4),
- tertiary education (ISCED levels 5, 6), and
- post-tertiary education (ISCED levels 7, 8).

Labour Force:

Based on the International Labour Organization (ILO) International Classification of Status in Employment (ICSE-93), with additional categories for those who are unemployed or outside the labour force.

- Employee;
- Self-employed (includes the four categories: employers, own-account workers, members of producers' cooperatives, and contributing family workers);
- Workers not classifiable by status (for whom insufficient relevant information is available, and/or who cannot be included in the preceding categories);
- Unemployed; and
- Outside the labour force. i.e student, retired.

Further classification may be given as per occupation.

Disability status:

Because of a Health problem:

1) Do you have difficulty seeing even if wearing glasses?
2) Do you have difficulty hearing even if using a hearing aid?
3) Do you have difficulty walking or climbing stairs?
4) Do you have difficulty remembering or concentrating?
5) Do you have difficulty with (self-care such as) washing all over or dressing?
6) Using your usual language, do you have difficulty communicating (for example understanding or being understood by others)?

Response categories:

- No difficulty; Some difficulty; A lot of difficulty; Cannot do at all

Other classifications at individual level are:

- level of literacy, ethnicity, languages spoken, language skills.
- The revised ICT household indicators include HH15, *Individuals with ICT skills, by type of skills. It will therefore be possible to cross-classify the individual use* indicators by ICT skill level.

Household Characteristics

- Household composition (*households with children under 15 and households without children under 15).* Household composition is relevant to measuring the digital divide in households with children
- Household size (number of household members, including those outside any age scope imposed).
- Geographical disaggregation such as urban/ rural. Countries use their own definition for the urban/ rural and include it in the metadata. Countries can disaggregate this to towns, districts, counties to match their local needs.
- Household with electricity can be used especially for the household ICT access indicators
- Household income

Cross-classification of data

Can produce information that is very useful for analytical purposes as is more detailed

- example: Internet use by young women (data are cross-classified by age and gender).

ITU proposes the following cross-classification:

- household composition by rural/urban,
- rural/urban by sex,
- age by sex,
- educational attainment by sex,
- status in the labour force by sex, and
- occupation by sex.

ICT Access by urban/rural and household composition

Committed to Connecting the World

International Telecommunication Union

No.	Indicator	All households	Urban/Rural		Household composition					
			Urban	Rural	has children under 15			does not have children under 15		
					Total	Urban	Rural	Total	Urban	Rural
HH1	Number of households with a radio									
HH2	Number of households with a television									
HH3	Number of households with any telephone (fixed and/or mobile)									
	Number of households with fixed telephone only									
	Number of households with mobile cellular telephone only									
	Number of households with both fixed and mobile telephone									
HH4	Number of households with a computer (all types of computer)									
	Desktop									
	Laptop (portable) computer									
	Tablet (or similar handheld computer)									
HH6	Number of households with Internet									
HH11	Number of households with Internet, by type of service									
	Fixed (wired) narrowband network									
	Fixed (wired) broadband network									
	Terrestrial fixed (wireless) broadband network									
	Satellite broadband network									
	Mobile broadband network via a handset									
	Mobile broadband network via a card or USB modem									
HH13	Number of households with multichannel television by type									
	Cable TV (CATV)									
	Direct-to-home (DTH) satellite services									
	Internet-protocol TV (IPTV)									

SURVEY DESIGN

Data sources

- Household ICT data are collected via a survey of households (including population censuses and mail surveys).
  - No examples are known where this is not the case
- Administrative data (or data collected from ISPs) are useful for measuring e.g. numbers of Internet subscriptions
  - but this is not the role of the household core ICT indicators

Types of household surveys

- An important consideration is whether questions are added to an existing survey or a stand-alone survey is conducted
- In the case of an existing survey, there may be more than one survey vehicle available

Types of surveys (continued)

- *Stand-alone household surveys* deal with a single topic (such as ICT access and use).
- *Multi-purpose household surveys* collect data on more than one subject via a single survey.
- *Household budget expenditure surveys* measure household expenditure and can be used to identify household access to ICT equipment and services.
- *Population censuses* can be used to collect ICT access and/or use data.
  - Principles and Recommendations for Population and Housing Censuses Rev.2 included eight (8) questions on availability of ICT in households

Question and questionnaire design – introduction

- Why is questionnaire design important?
- Poor questionnaire design can be a significant source of non sampling error
  - therefore design should be done carefully…
  - …and sufficient time should be allowed for thorough testing of the questionnaire.
- A questionnaire design that is optimal for all respondents is unlikely to be achieved.
- Ideally, questionnaires should be administered by interviewers.

General principles of questionnaire design

- Questionnaires should be as short as possible.
- Questionnaire structure should:
  - maintain respondents' motivation and interest
  - have a logical flow with appropriate skips
  - enable grouping of questions for a particular respondent.
- Question wording should be:
  - clear and simple, with short questions
  - not double-barreled (a question that is actually two questions)
  - unbiased e.g. not lead the respondent.
- Avoid order and memory errors in list (response item) questions.
- Avoid recall error problems.

General principles (continued)

- Build trust with the respondent:
  - avoid sensitive questions if possible
  - assure confidentiality of responses.
- Where several languages are spoken, there should be a version for at least each major language.
- Questions and whole questionnaires should be thoroughly tested before use
  - test that respondents can answer questions accurately
  - …and have a common understanding of the questions.
  - Testing can be qualitative or quantitative.
  - Testing can be done in stages.
  - Early testing can assist with planning e.g. time for completion.
- Take care with non-question elements, e.g.
  - include check digits with identifiers if possible, and
  - ensure that boxes are large enough for the response.

ICT model questions

- Need to be adapted by countries, taking into account:
  - language and culture
  - the form of data collection.
- It is important to:
  - retain question meaning and
  - preserve the logic, by using the same populations of households or individuals for each question.
- Variations to model questions:
  - can add or split categories.

Reference period

- Reference period refers to the recall period the respondent is reporting on.
  - Recall error tends to be worse with a longer reference period.
  - The problem will be worse if there is a long lag between the end of the reference period and the time of data collection.
- The length of the recall period for the ICT *individual use* core indicators is 3 months.*

* *Previously this was 12 months.*

Internet access services

- The response categories in the model question for core indicator HH11 may be too technical for some respondents.
- Questions on this topic should use categories that are relevant to services existing within the country.
  - This could mean specifying the brand names of widely available services.
- Interviewers should be familiar with the names of common Internet access services.

Survey scope and coverage

- The *scope* of a survey refers to the target population to be represented by the survey.
- *Coverage* is the degree to which the in-scope units are present on the survey frame
  - and are therefore represented in the sample.

Scope for ICT household surveys

- Countries should report ICT usage information for the entire population of the country, i.e. there is no minimum age scope any more for these indicators

- Individual countries may restrict the household scope for practical reasons
  - for example, to households occupying private dwellings
  - or households in urban areas (though this is not recommended).

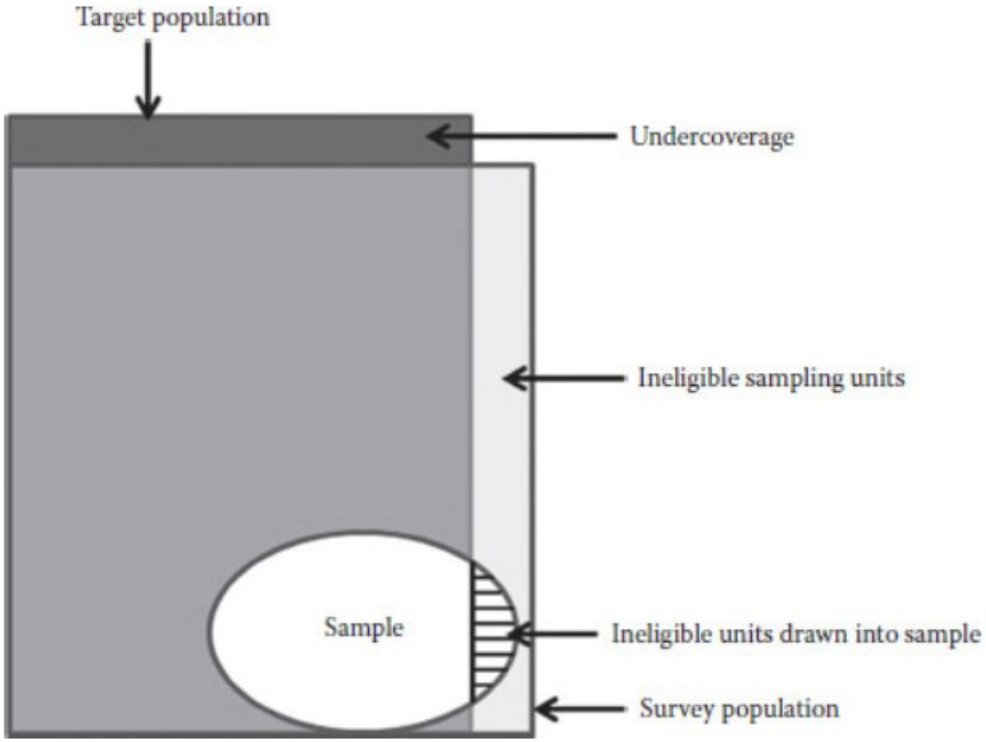
Target populations and survey frames

- The target population is the survey scope.
  - It is the population about which survey estimates will be produced.
- The survey frame (also called *sampling frame*) is a list from which units of a survey are selected.
  - Ideally it contains all members of the target population.

Survey frames

- Examples of survey frames used for household surveys are: a register of persons, an electoral roll, a population census data file, a list of dwellings used for property valuation purposes or a master sampling frame.
  - A master sampling frame (or master sample) is a large sample which is used for several surveys.
- Frequently, there is more than one stage of sampling.
  - and this requires more than one frame.

- Target population / survey population



Complex Survey Data Analysis with SAS, Lewis

Statistical units

- For ICT use measurement, there are typically two statistical units:
  - *households* and *individuals*.
- The *household unit*
  - refers to information about ICT access in the household.
- The *individual* unit
  - provides information on use of ICT.
- The core indicators require both units.

How are the units defined?

- Both households and individuals are defined in terms of the survey scope
  - for instance, all individuals aged over 14, or
  - only households occupying private dwellings.
- Households need additional definition.
  - They may be difficult to define in a consistent way…
  - because living arrangements vary across countries.

Definition of *household*

- For the purposes of the *Manual*, a h*ousehold* is defined as a unit:
  - consisting of *one or more* people, who
    - may or may not be related to each other
    - share accommodation
    - make common provision for food.

Sample design and selection

- The *Manual* does not specify a particular sampling methodology.
  - A variety of methodologies is used by countries.

Some preliminaries

Sampling

- Simple Random Sampling (SRS)
  - Each statistical unit has the same chance of selection
- Complex sampling
  - Primary sampling Units
  - Strata (eg urban /rural)
  - Cluster
- Sample size and allocation

Sample design – general principles

- For a given sampling methodology, sample sizes need to be higher where a higher degree of reliability or confidence is required.
- For a given sampling methodology, a greater level of detail in output requires a higher sample size for a constant degree of reliability.
- Oversampling to compensate for non-response is commonly practised.

Sample selection – general principles

- Households, and individuals within those households, should be selected in an unbiased manner.
- For instance, individuals within households should be randomly chosen.
  - Options include selecting all individuals, selecting one at random, or asking one individual to report information about others.
  - Where a selected individual is not available at the time of interview, they should be contacted later.

DATA COLLECTION AND DATA PROCESSING

Data collection methods

- Direct observation
- Postal
- F2F
- Telephone
- CATI
- CAPI
- CAWI
- Online

Execution of surveys

- Implementation of fieldwork
  - Need for a well-organised and effective field organisation
  - Equipment and materials
  - Management of survey operations, clear well defined line of command
  - Publicity
  - Selection of interviewers
  - Training of interviewers
  - Field supervisors
  - Follow-up of non-respondents
  - Reducing non-response

Data processing

- By 'data processing', we are referring to:
  - data entry
  - editing household ICT data
  - imputation for missing data and non-response
  - weighting of data
  - derivation and reporting of ICT household indicators.

Data entry

- Data entry may occur:
  - during the interview (using CAPI or CATI) or
  - as a separate step, which may be in a specialized data entry unit.
- Use of CAPI/CATI may reduce data entry errors.
- Good training and procedures are necessary to ensure accuracy of data entry.
- *Check digits* are useful to check that record identifiers or codes have been entered correctly.
- Data entry errors may also be found during data editing.

Editing household ICT data

- There are two main types of data editing:
  - Micro-editing
    - is applied to individual records
    - includes range, skip, consistency, logical and typographical checks
    - Edits can be *fatal* or *non-fatal.*
    - Micro-edits may be performed during an interview.
  - Macro-editing
    - is a check of aggregated data for coherence
    - includes consistency, relationship and logical checks.
- Both types are relevant to ICT data.

Imputation for non-response

- Non-response can occur for:
  - the whole response (unit non-response) or
  - part of the questionnaire (item non-response).
- Imputation: estimate for unit or item non-response.
- In household surveys, other units may substitute non-responding units.
- Both non-response and estimates for non-response estimates can be a significant source of error (bias)
  - .. and therefore need to be handled carefully.
  - The best option is to try to avoid non-response!

Weighting of data

- Data from a sample need to be weighted to represent the target population.
- The initial 'design weight' of a unit is the inverse of its probability of selection
  - for example, if the population size is 1000 households and 200 are randomly sampled, then the probability of selection is 200/1000 and the weight is 1000/200 (that is, 5).
- Design weights incorporate the possibility of selection at each stage of sampling.

Weighting of data (continued)

- Design weights may be adjusted as the survey progresses to reflect:
  - non-response
  - out-of-scope units and/or
  - frame problems.
- Important to weight responses according to independent estimated distributions of the population.
  - This compensates for non-representativeness of the effective sample (that is, the population of respondents to the survey).
  - It may be referred to as 'post-stratification' or 'benchmarking'.

Derivation and reporting of ICT household indicators

- Derivation of the ICT household indicators is not complex
  - but care needs to be taken with denominators
  - and aggregation of categories where relevant.
- For international reporting, countries are asked to provide total numbers (instead of percentages)
  - for example, the number of adult Internet users or the number of male computer users.
- These numbers need to be population estimates (i.e. weighted sampling values).

Derivation and reporting (continued)

- Numbers for the total population, and each relevant sub-population, also need to be provided so that proportions can be derived.
  - Populations are the total number of in-scope households and individuals.
  - Sub-populations follow the classificatory variables and include the number of males, the number of households with children under 15, etc.
- Population (and sub-population) numbers should represent the target population…
  - Weighted values.

Data quality and evaluation

- As we have seen, it is important to control and minimize error.
  - To do this, we need to understand sources of error.
- It is also important to work within a data quality framework …
- … and to evaluate survey procedures.

Survey error

- Survey error can be *sampling* or *non-sampling* error.
  - There may be tradeoffs between sampling and non-sampling error.
- Countries should reduce survey error as much as possible by:
  - using well-designed samples of sufficient size
  - carefully designing and testing questions and questionnaires
  - intensive training and checking of interviewers
  - reducing non-response rates as far as possible
  - minimizing data processing errors.

Sampling error

- Sampling error occurs because only a fraction of the total population is sampled.
  - It does not apply to censuses.
- The sampling error of an estimate can be expressed as:
  - the standard error of the estimate
  - a ratio of the standard error of the estimate to the value of the estimate (percentage)
  - as a confidence interval around the estimate.

Survey evaluation

- Documentation of survey processes and procedures as input to evaluation
  - detailed plans and budgets
  - changes as the survey progresses, including
    - changes in sample design during the field phase
    - changes in budgeted costs
    - changes in procedures.
- Staff input at all levels can be very useful for evaluating the survey.
- Documentation which is publicly disseminated is covered in Module 5.

DATA DISSEMINATION

Data dissemination

- What are the different ways that data can be released?
- Formats include:
  - hardcopy (publications, summaries or press releases etc)
  - electronic (including pdf files of hardcopy publications, web publications)
  - data tables on websites, spreadsheets or CD-ROMs.

Tabulation plan

- Tabulation and analysis plan
  - tabulation plans
  - dummy tables
  - titles
  - stubs and captions
  - substantive variables
  - background variables
  - population groups
  - categories of classification

Metadata reporting and dissemination

- Data on their own are not very useful.
- They need to be described and explained
  - for instance, how they were collected and
  - how accurate they are.
- This type of information is referred to as 'metadata' – 'data about data'.
  - includes information about the survey and about individual data observations.
- ICT statistics metadata repositories can be useful.

Metadata on data quality

- An important element of metadata is information relating to data quality.
- In a previous presentation, we discussed the dimensions of data quality
  - relevance
  - accuracy
  - timeliness and punctuality
  - accessibility and clarity
  - comparability
  - coherence.
- We also discussed quality assurance frameworks based on these dimensions.

Metadata on limitations of survey data

- Very important elements of data quality.
- Include information on limitations due to:
  - sampling error
  - non-sampling error
  - lack of comparability (e.g. with past data).

Data collection and dissemination of ICT household statistics by ITU

- ITU started collecting ICT household statistics from NSOs in 2003.
- From 2005, ITU has sent an annual questionnaire to NSOs.
- Data compiled from these collections have mainly been distributed through publications and the ITU World Telecommunication/ICT Indicators (WTI) Database.

Derivation and reporting of data to ITU

- When reporting data to ITU, countries should provide:
  - numbers of units with a particular 'ICT characteristic', for example, Internet users, plus
  - numbers for the total population, for example the number of in-scope adults, and each relevant sub-population.
  - Numbers should represent the target population and not the sample.
- Any categories that have been split need to be re-aggregated (at the unit record level).

THANK YOU

For more information:

http://www.itu.int/ict

and

indicators@itu.int