The Foundations of CountrySTAT E-Learning

EAC/FAO Advanced Training Workshop of CountrySTAT

Angela Piersante
Statistician

Lusaka, 12 – 16 November 2012

Background information

What is E-Learning?

E-learning is the computer and network-enabled transfer of skills and knowledge.

How it works?

- E-learning applications and processes include Web-based learning, computer-based learning, virtual education opportunities and digital collaboration.
- The content is delivered via the Internet, intranet/extranet, audio or video tape, satellite TV, and CD-ROM. It can be self-paced or instructor-led and includes media in the form of text, image, animation, streaming video and audio.

Why should we use E-Learning ?

The goal is to help our audience use effectively the key information included in the manual.

The course is learner-centered:

- The learning materials is relevant to the activities carried out by the audience
- Examples, self-assessments and interactive exercises play a key role
- All information is contextualized and put in relation to the task to be performed.

Background information

The course covers the statistical project topic areas:

1. The process of data collection, harmonization, standardization and validation
2. Data structure (main and sub-domains, main data set, connections between national and sub national)
3. Concepts and definitions common to CountrySTAT and FAOSTAT
4. The importance of adding metadata and how to build them.

A task analysis has been carried out in order to:

1. Analyze the features of the target audience
2. Identify the tasks that the user should be able to carry out
3. Identify the knowledge needed to carry out the identified tasks.

Main Objectives of the course

- Enhance users awareness of the CountrySTAT objectives and rationale
- Publish data that meet international standards and quality criteria
- Improve users capability in standardize local data so that it can fit into an international format
- Recognize the importance of accurate metadata and of inputting metadata into the system
- Enhance users knowledge of the international nomenclature, to develop/improve the local nomenclature
- Speed up the process of publication of official national data on the website.

Main Users of the course

All users who participate and contribute to the publication of statistics on agriculture and food on the CountrySTAT website, and in particular:

- Members of the Secretariat
- The National Coordinator
- Members of the Technical Working Group
- Members of Regional Organizations.

Content structure

Module I:

CountrySTAT knowledge base

All the "need-to-know" information to support the effective implementation of the CountrySTAT system an national level. It includes strong motivational elements:

1.1: Discovering CountrySTAT

1.2: The CountrySTAT website and the people who make it happen.

1.3: The CountrySTAT framework for data quality

Module II:

The data dissemination process

Practical information to support the key activities to ensure a successful publication process:

2.1: Phase 1: data and metadata collection.

2.2: Phase 2: harmonization and validation of national data.

2.3: Phase 3: the standardization of data.

Learning Objectives: Lesson 1

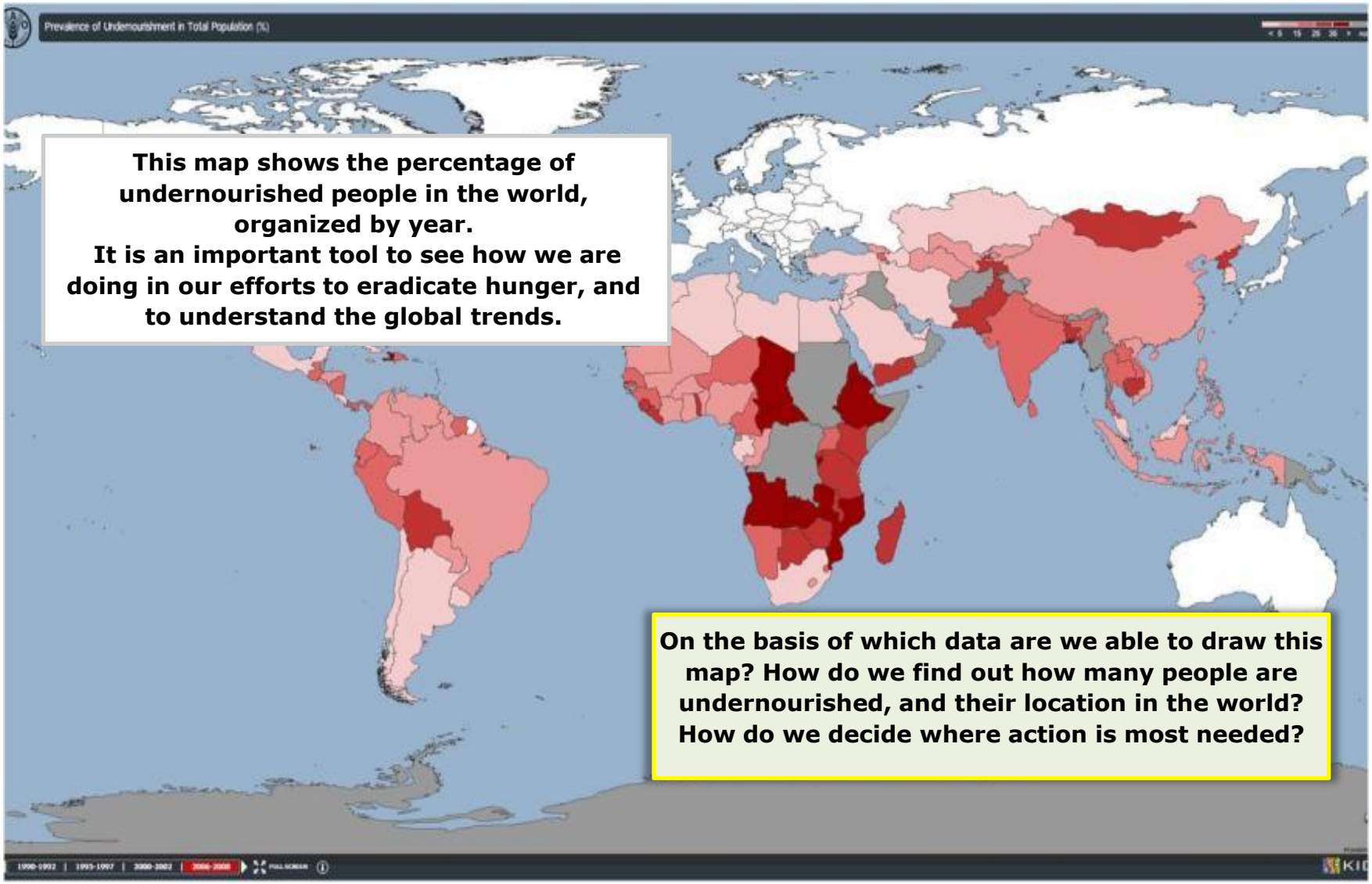
Discovering CountrySTAT

Would you like to use a one-stop-centre for national statistics on food and agriculture?

Would you like to provide reliable and timely data on your country, through a system that is easy to manage?

- Discover what is CountrySTAT, its purpose and benefits
- Find out how it can be useful to you.

Introduction



Introduction

The answer is: through reliable statistical data. In fact, in order to eradicate hunger in all countries and reduce the number of undernourished people, it is vital to design and implement:

Policies on agriculture development that are appropriate and efficient.

Efficient policies must be grounded into factual evidence about the sector, through a systematic and rigorous use of statistics for:

- **issue recognition;**
- **informing policy choice and design;**
- **forecast the future;**
- **monitor policy implementation;**
- **evaluate policy impact and make the adjustments needed.**

Agricultural investment projects that are sustainable and effective.

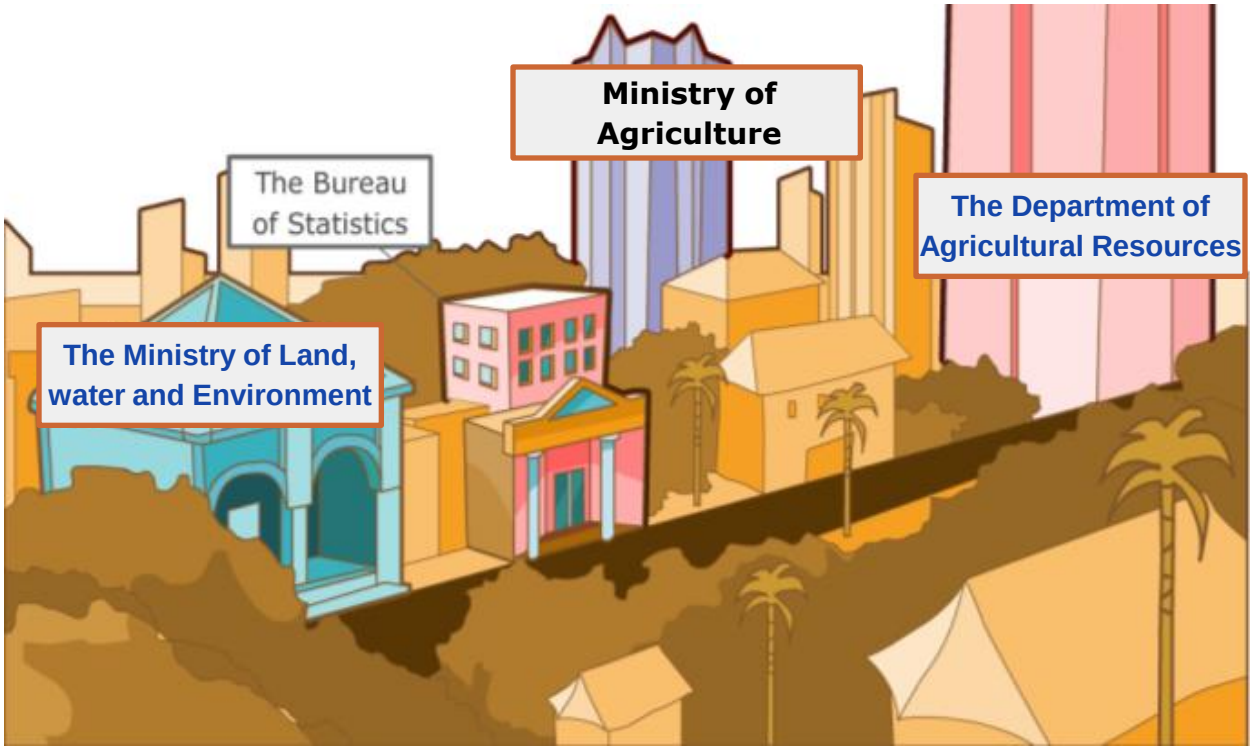
Effective decisions on these investment must be based on sound information about elements such as:

- **population characteristics;**
- **land and input use;**
- **productivity;**
- **prices;**
- **prevailing economic and social situations faced by producers.**

Life before CountrySTAT for the end user

What different does CountrySTAT make for you, as an end user?

Imagine that you are working for a national farmers association, and wish to improve the production of cereals in a certain area. You need as soon as possible the latest statistics on production and available land, but you are not using CountrySTAT. Where would you go to get your data?



It could take you the entire day to access just some of the data you need!

The users of CountrySTAT

Who are the main users of CountrySTAT?

End users

Those who need reliable data to make analysis on the Agricultural sector to make informed decisions. For example, policy and decision-makers, researchers, economists, analysts, national farmers associations.

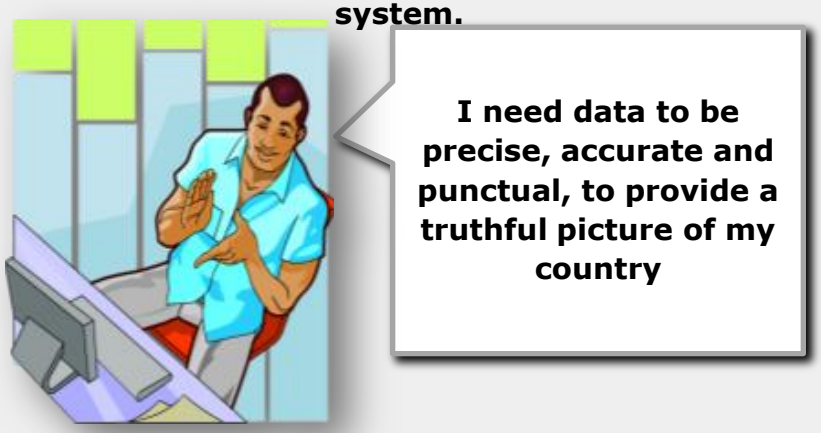
The end users connect to the CountrySTAT website and search the information they need.

Example of how the end user accesses information provided by CountrySTAT

Data providers

Those who gather the data about their country, and carry out the publication process of the data on the CountrySTAT website.

The data providers have a dedicated section on the website, where they can input the data in the system.



Example of how the data provider uses the CountrySTAT website

Life before CountrySTAT for the data provider

Once, I needed to provide trends in Cereal Production Estimates. I had to compare data from different agencies which produced surveys with different results . As you can see, the difference between the two data for 2005 was as high as 67.1%!
How could I harmonize them? What was the right decision?

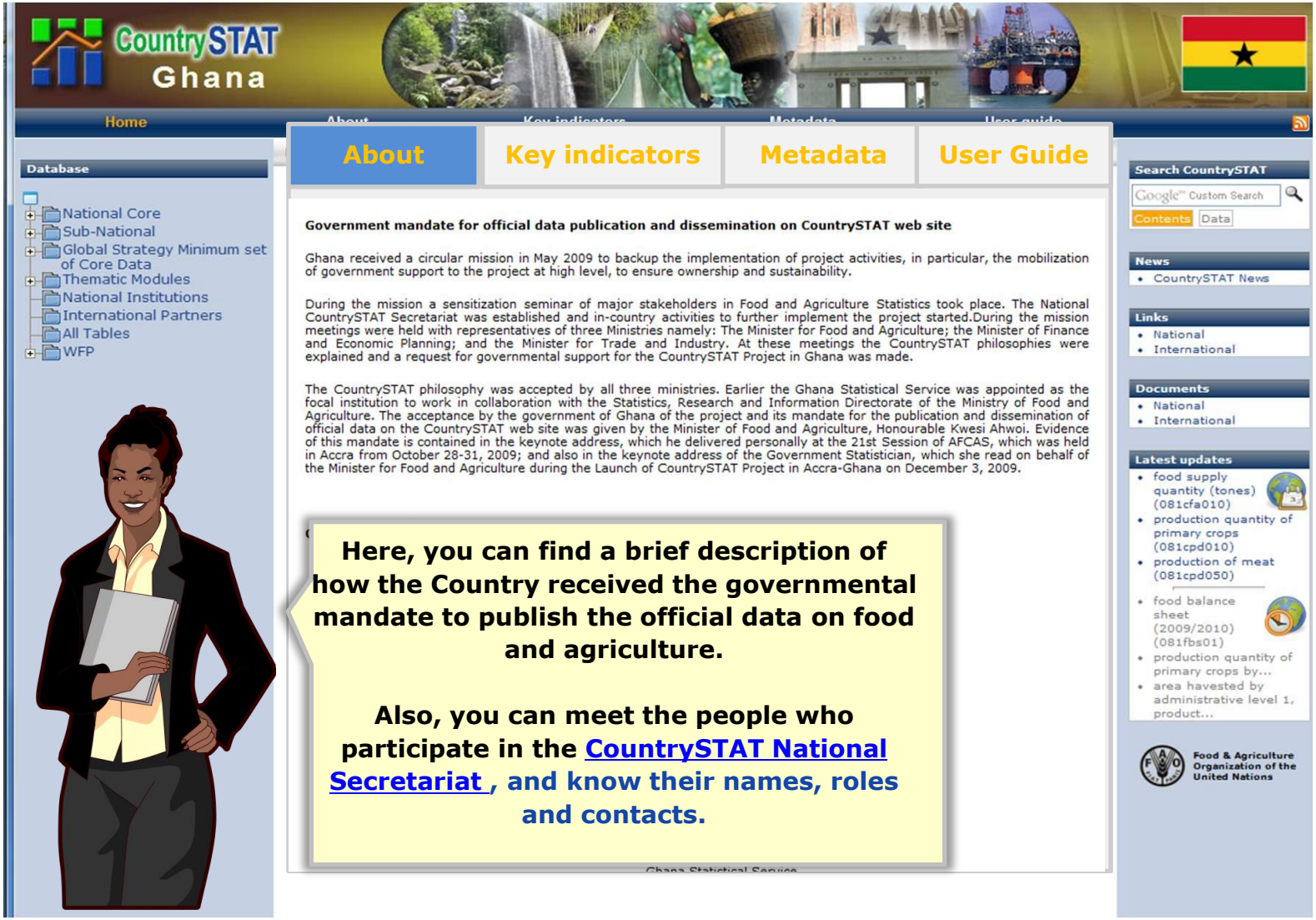
System	1996	2002	2003	2005	2006
Agency A *(early warning)*	1,377	1,767	1,811	1,900	2,098
Agency B production survey)	1,459	1,454	1,509	1,137	1,700
Difference	+5.6%	-21.5%	-20.0%	-67.1%	-23.4%

Learning Objectives: Lesson 1.2

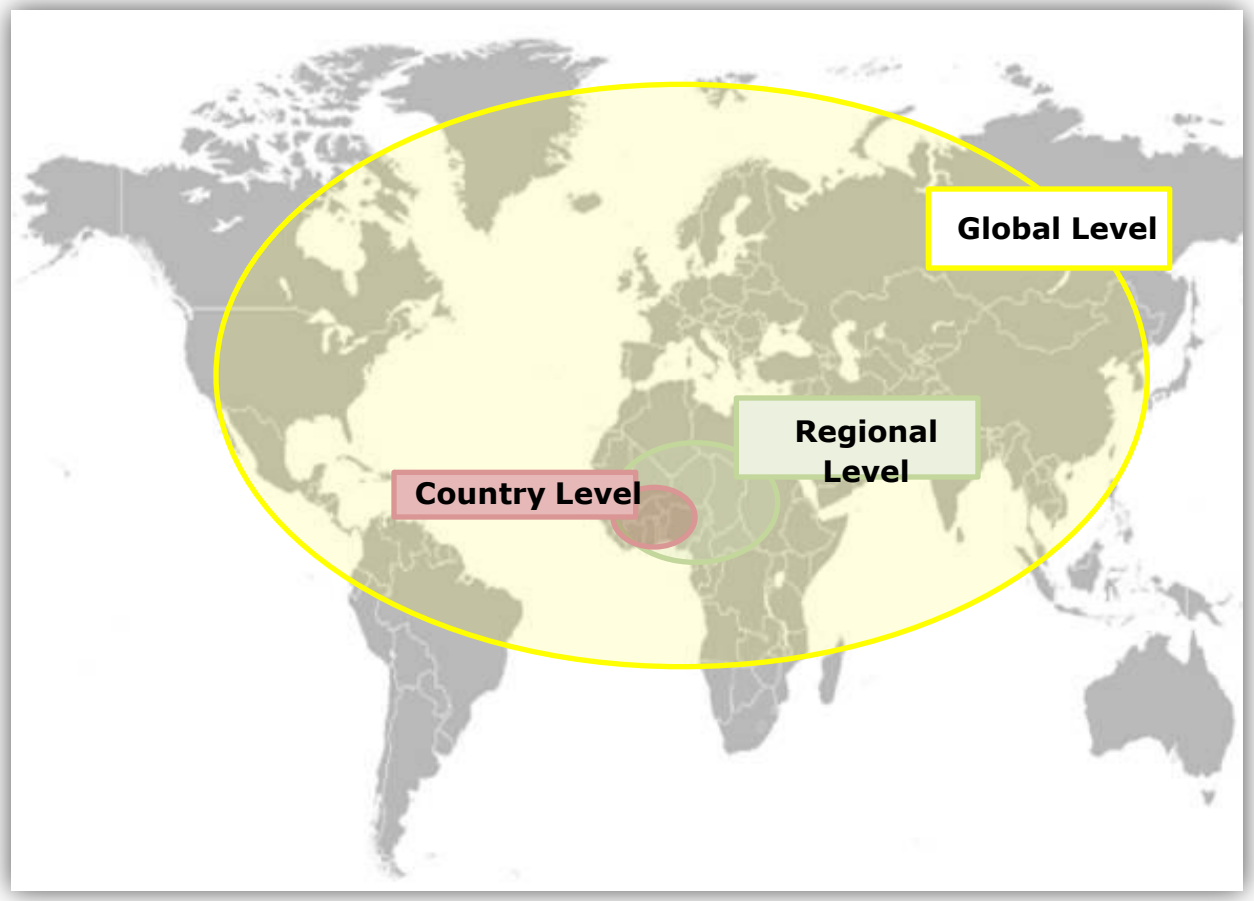
The CountrySTAT website, and the people who make it happen

- identify the key features of the CountrySTAT website;
- recognize specific roles and responsibilities in the CountrySTAT institutional framework;
- describe the interaction among the different groups of people involved in CountrySTAT.

The CountrySTAT website: the homepage of the participating countries



The people behind CountrySTAT



The people who ensure the implementation of CountrySTAT are part of CountrySTAT bodies, at three levels:

•Country,
•Regional, and
•Global.

Click on each tab to learn more about each level

The interaction among CountrySTAT members

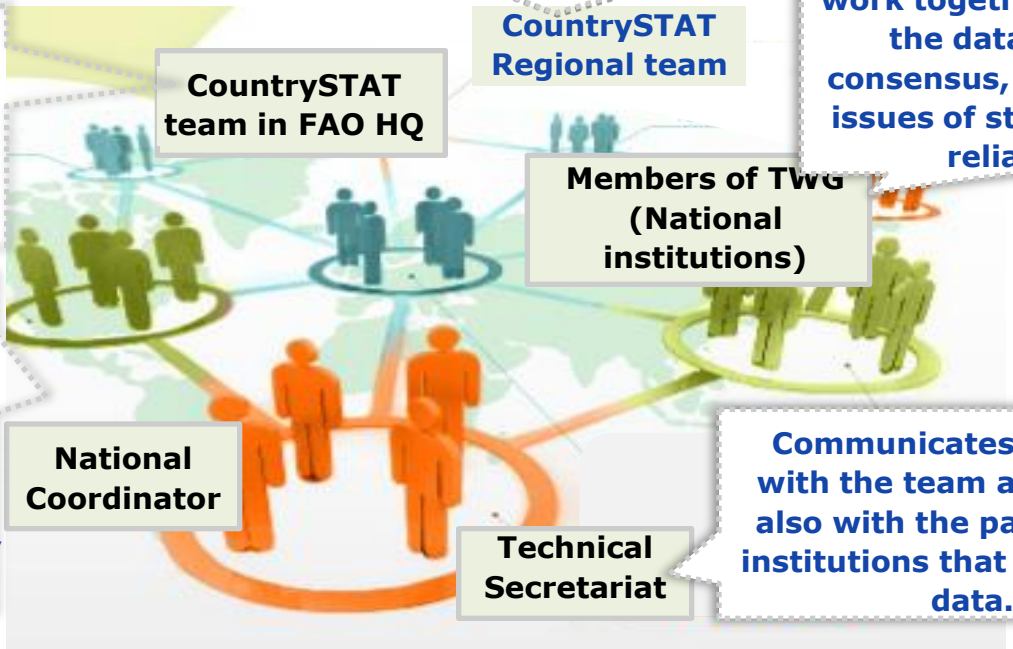
There is a dynamic process between these CountrySTAT bodies, facilitating national and institutional capacity development.

Dialogue and collaboration between CountrySTAT champions, participating Institutions and National organizations is critical to guarantee a successful implementation of CountrySTAT.

Roll the mouse on the groups of people to read some significative examples

The Regional team develops capacity building at regional level

The team at FAO communicates regularly with the National Secretariat, and provides support on the implementation of activities



The members the TWG work together to validate the data and find consensus, especially on issues of statistical data reliability.

Works in close collaboration with the Regional Organizations and the Team at FAO.S/he also coordinates the TWG meetings and is in close touch with all stakeholders, for advocacy purposes.

Communicates regularly with the team at FAO, and also with the participating institutions that provide the data.

Learning Objectives: Lesson 1.3

Promoting quality: international concepts, definitions and classification

- Recognize how the CountrySTAT quality framework addresses the needs of the users
- Describe how the CountrySTAT process unfolds, and who does what
- Recognize why international standards are important, and what is their use
- Know how to use the FAOSTAT product classification and definitions for comparability purposes.

Introduction

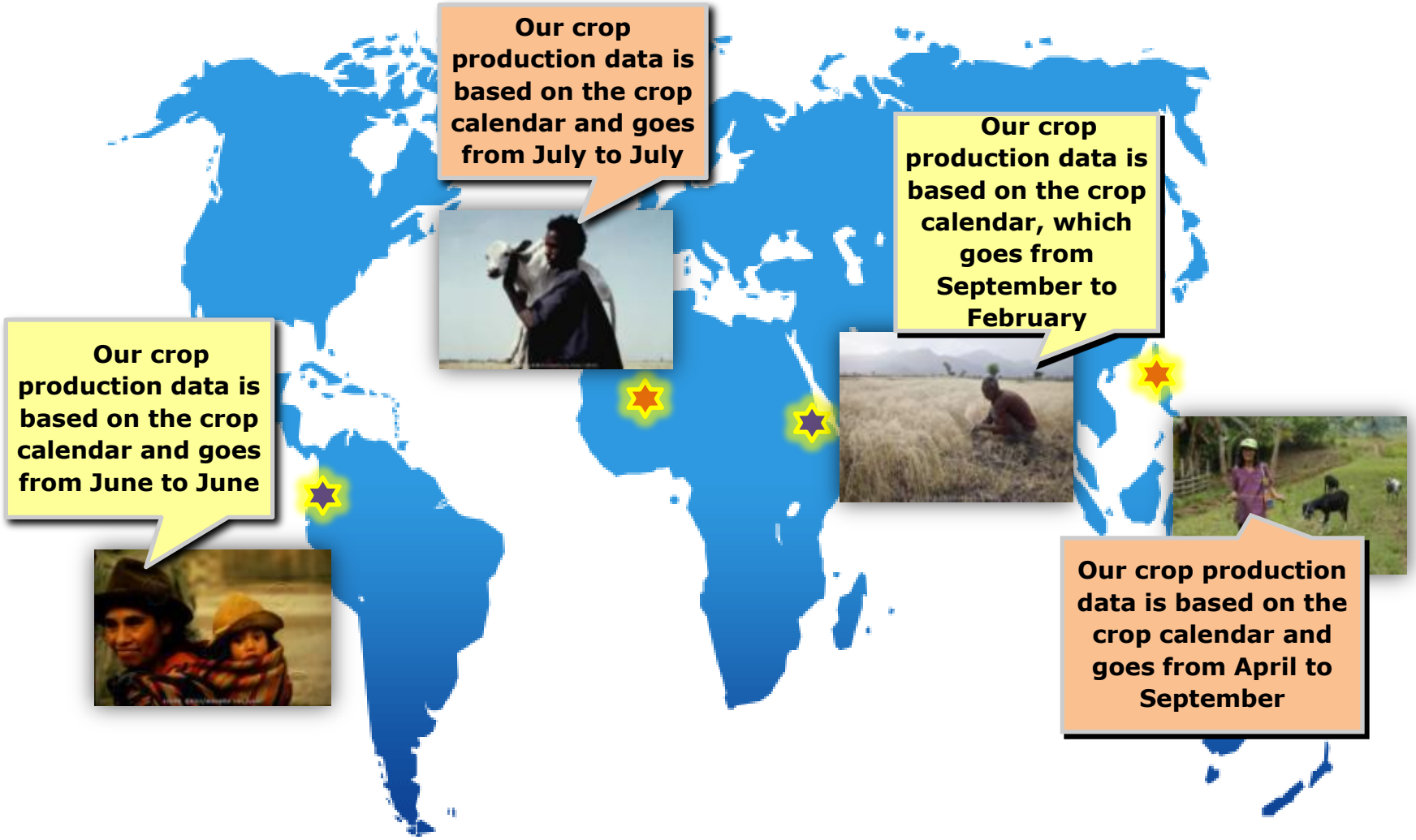
Comparable data is possible because the CountrySTAT process uses FAO *standards and concepts*, which are adopted by all those who implement CountrySTAT around the world.

We will discover how the people who implement CountrySTAT are able to make the local information understandable to all users, no matter where they are in the world.

In this lesson, we will learn more about this process, and in particular, we will consider:

- **How the CountrySTAT quality framework fulfils the user's needs in terms of data reliability**
- **How the CountrySTAT process unfolds, and who is involved**
- **What are the international standards and concepts on which the process is based**
- **What is the format for organizing the data (the standard data structure) used by data providers**
- **How to apply these standards, by using some real life examples.**

What would happen without standards?



What are the FAO international standards, concepts and definitions

Without standards

Konshu

"Big" beans

Azuki

With standards

Beans

We just mentioned that CountrySTAT is based on FAO international Standards, Concepts and Definitions.
Why are these standards so important?

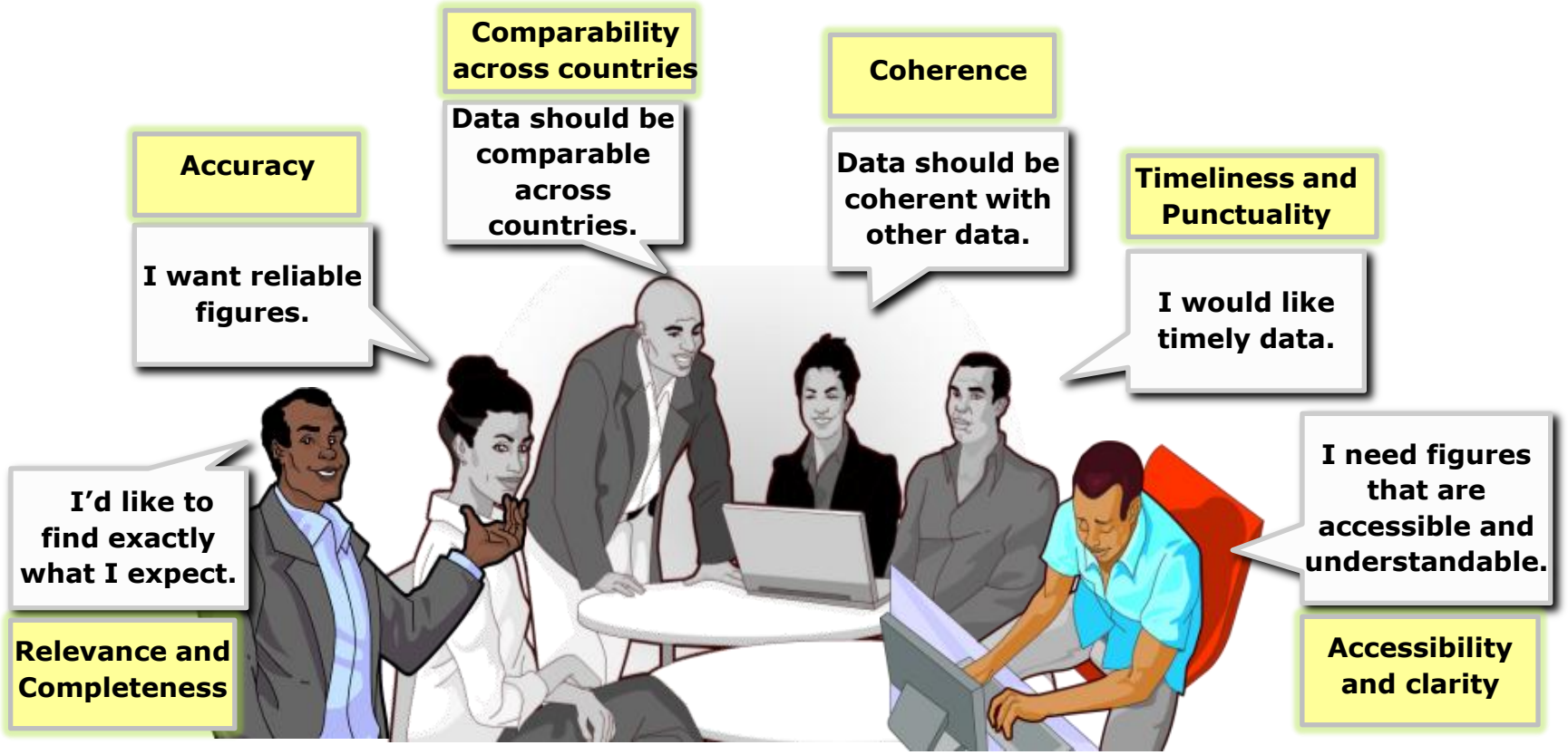
The CountrySTAT framework takes into account the existing national systems for food and agricultural statistics: however, national standards differ from country to country. Therefore, the data collected using national standards must be converted and harmonized with the international standards.
This facilitates comparison and integration of data, both within the country and across countries.

Example: comparability within the country

Example: comparability across countries

The CountrySTAT framework for Data Quality

Each of these identified needs clearly calls for specific characteristics, or dimensions that quality data must have:



These specific dimensions/characteristics are reported in data quality assessment common with Eurostat, IMF, and African Union.

Click to learn more

Learning Objectives: Lesson 2.1

Promoting reliability during data collection

- Recognize the features of a reliable primary data source
- Understand the use and importance of metadata
- Know how to create a Correspondence Table while collecting the data.

Identifying the appropriate source: use of metadata in CountrySTAT

We just considered the importance of metadata for the national statistical system. But how are metadata used in the CountrySTAT process? CountrySTAT uses metadata as:

1. Information **on the original primary data item** (e.g. its data source, focal point, geographical coverage. methodology), collected by the statisticians while gathering the primary data,
2. Information **describing the data processing, analysis and data revision methodology**, recorded by the TWG participants.

The CountrySTAT process

Collect data from various national sources

Harmonize and Validate local data at National level

Standardize data with international concepts

Input the data into website

At the data collection phase, statisticians gather the information, available directly from the source, such as the Focal point for reference, or the geographical coverage.

Then, during the TWG meeting, the participants report the information to describe the methodology used to harmonize and validate the data.

Finally, all metadata are inputted into a template for publication on the website, together with the data they refer to.

Identifying the appropriate source: use of metadata in CountrySTAT

In particular, these are the metadata accompanying data items published on CountrySTAT.

DATA ITEM
on CountrySTAT

At the time of data publication, this information will be inputted onto the website with the related data, using a specific software..

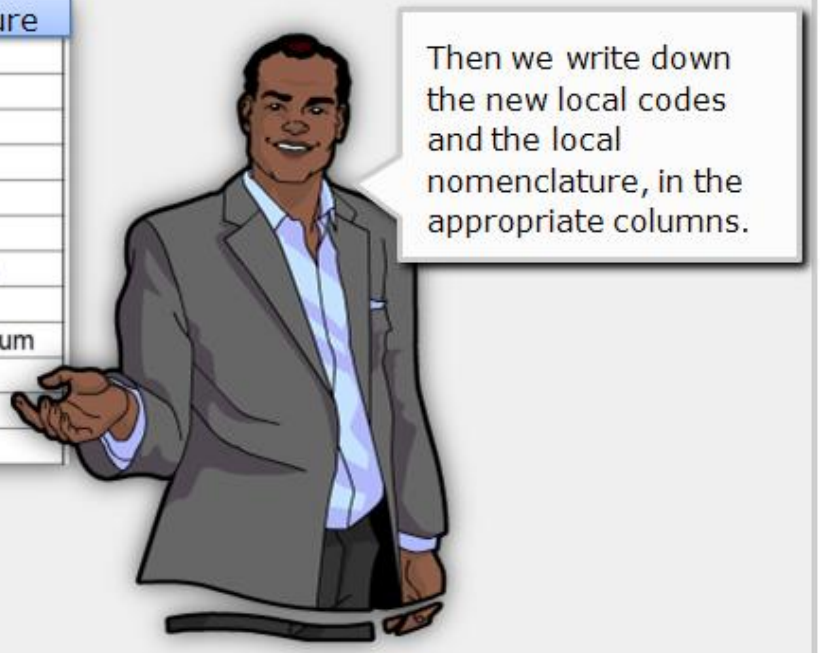
Sources	Focal point for contact	Concepts, definitions and classifications	Geographical coverage	Frequency of data collection	Dissemination method	Timeliness and punctuality

Census Surveys	Overview Design Conduct, operations and data quality control Statistical Report
Administrative records	Responsible Agency Description of the contained information Data source (the agency that supplies the information

Click each metadata to read what is the related information

Creating a Correspondence Table between national and international classification

FLC Item Code	FAOSTAT nomenclature	Local code	Local nomenclature
71	Rye		
72	Flour of Rye		
75	Oats		
76	Oats Rolled		
79	Millet		
80	Flour of Millet		
83	Sorghum	008301	White Sorghum
		008302	Red Sorghum
		008303	Improved Sorghum
84	Flour of Sorghum		
89	Buckwheat		
90	Flour of Buckwheat		



What if you believe that your local product does not belong to a category or you cannot find its category?

Learning Objectives: Lesson 2.2

Harmonizing and Validating the data: the TWG meeting

- Know how to standardize data according to the variables product, time, geographical levels and measurement units
- Identify seven different types of data analysis, to promote data coherence and reliability
- Know how to validate the data during the TWG meeting
- Identify some key activities of the TWG meeting.

Sub-National data

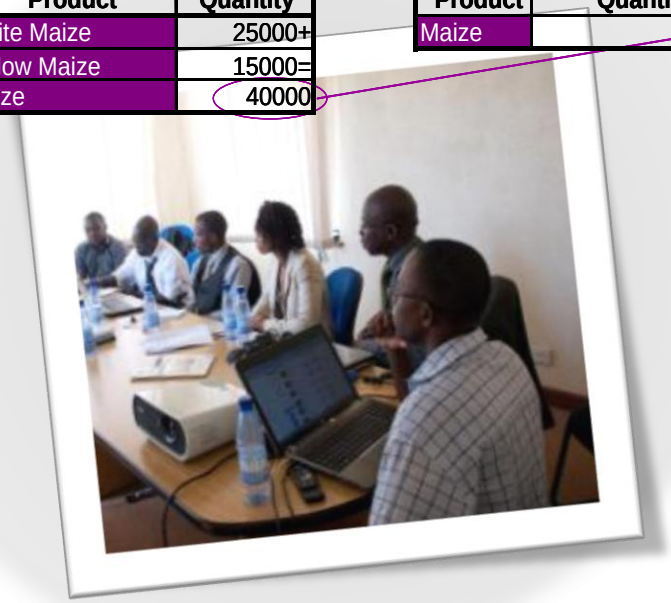
Production quantity of primary crops by local product (tonnes)

Product	Quantity
White Maize	25000+
Yellow Maize	15000=
Maize	40000

Core data

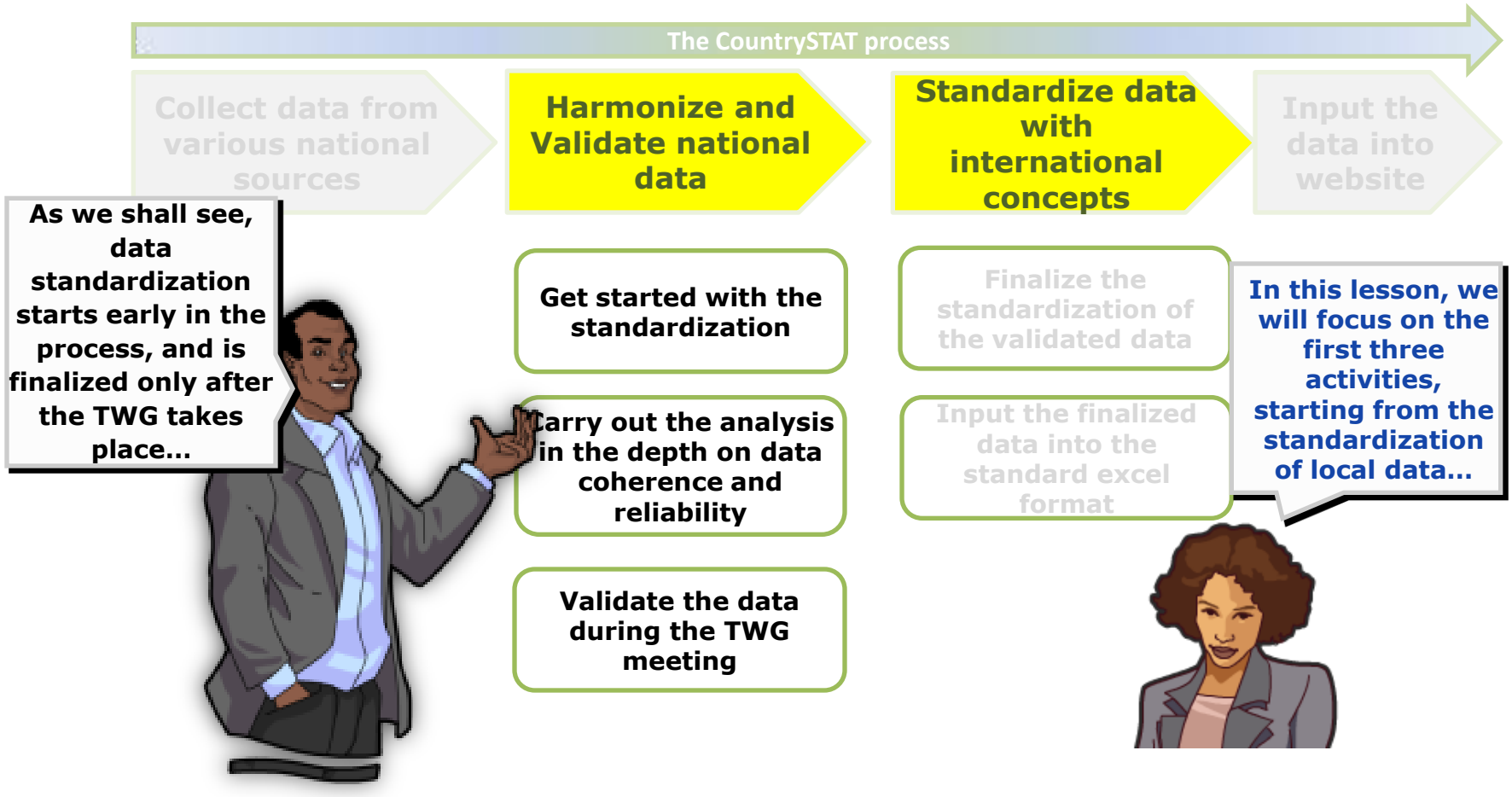
Production quantity of primary crops (tonnes)

Product	Quantity
Maize	40000



Introduction

Here are the next steps, which are at the heart of the CountrySTAT process. Amir and Paula, statisticians in their country's National Secretariat, will tell us more about these activities as the lesson unfolds:



The first activities in the standardization phase

The standardization activities are usually started out by the statisticians in the National Secretariat and then validated during the TWG meeting.
When standardizing local data, **three variables** are taken into account, plus the standard **measurement units**:

1. Products

Aggregation of local products variety to the international main category of **products**

2. Geographical coverage

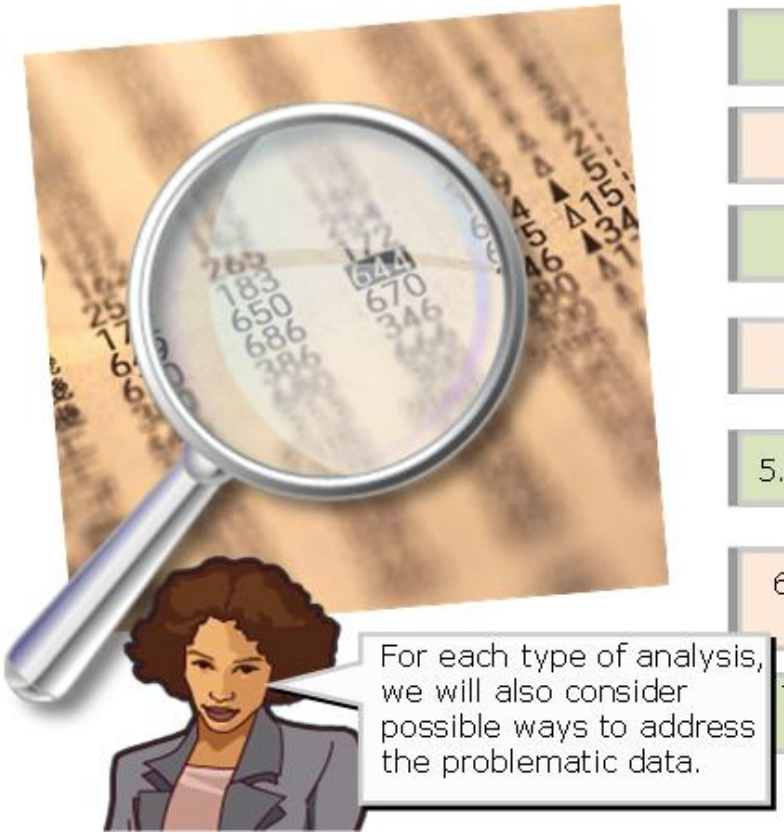
Aggregation of all geographical administrative levels to the **national level**

3. Time

Data should relate to the "**calendar year**" (instead of the various crop calendars)

Data must also be aligned to the **standard international measurement units**, as we shall see later in this lesson.

Data analysis for data reliability



We shall now consider seven types of analysis:

1. **Coherence between national sources**
2. **Missing and completeness data**
3. Anomalies in the **historical trends**
4. Incoherence **between related indicators**
5. Coherence between **Core** and **Sub-national** data
6. Consistency between the local and international **concepts and definitions**
7. **Correspondence** between **national and international classification**

The preparation of a TWG meeting

Amir and Paula have just organized a successful TWG meeting in their country.
To start, let's ask them some key questions:

The National Secretariat, and appointed representatives of **all** the national participating institutions. In addition, there is at least one representative of the FAO HQ CountrySTAT team, and sometimes other UN agencies (such as WFP).

The National Coordinator has a key role in the TWG meeting. Her/his activities include: preparing the documentation and the data issues for the meeting (including the agenda and the introductory document); coordinating the meeting activities and organizing the tasks to be carried out; preparing the final report.

Who should participate in the TWG?

What is the role of the National Coordinator?

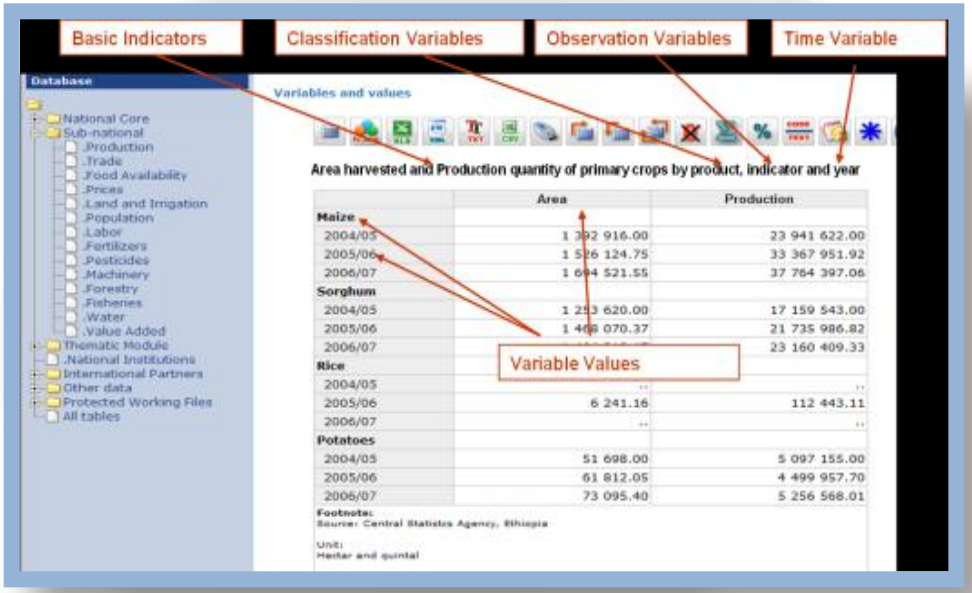
How many days should the TWG last?

Click the text boxes above to ask the questions

Learning Objectives: Lesson 2.3

Finalizing the data for the CountrySTAT website

- Recognize the requirements of the statistical tables used to publish data on the CountrySTAT website,
- Identify the difference between the tables used in the CORE and those in the Sub-national modules,
- Recognize the key elements of a table title,
- Identify the structure of the matrix codes in the CountrySTAT statistical tables.



Introduction

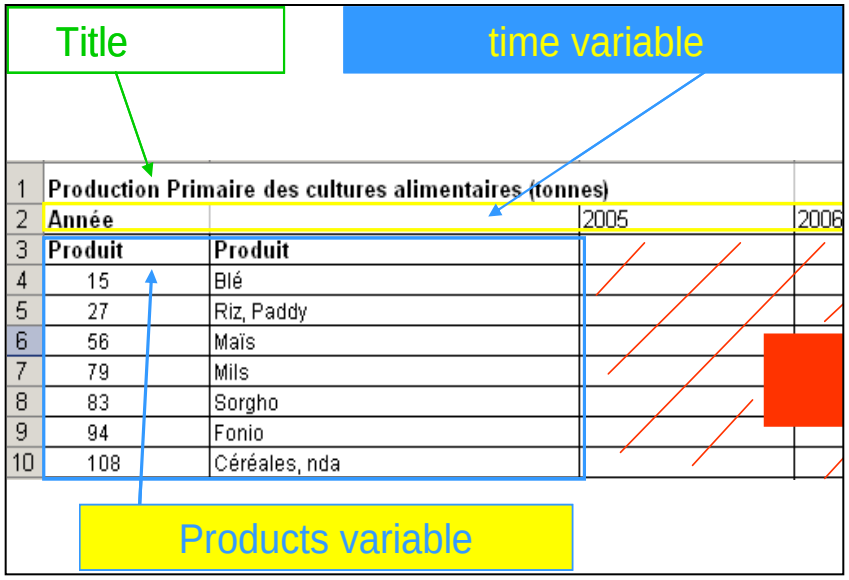
When we make the national data available on the CountrySTAT website, we reach a very vast audience: access to information must be simple and immediate.

How to organize these data in a way that is compatible with all the national data management systems? CountrySTAT uses a standard data structure in order to make the system easy to use, and to allow national and international users to get information in a meaningful, synthetic, simple and consistent way.

In this last lesson, we will learn about the CountrySTAT standard data structure and the rules for the publication of statistical data on the web, according to the International Standards.

This is the last step we take to finalize our work and see the tangible results on the CountrySTAT website!

Requirements of a statistical table: the structure



1	Production Primaire des cultures alimentaires (tonnes)			
2	Année		2005	2006
3	Produit	Produit		
4	15	Blé		
5	27	Riz, Paddy		
6	56	Maïs		
7	79	Mils		
8	83	Sorgho		
9	94	Fonio		
10	108	Céréales, nda		

Example of table with two variables: products and time.

Each statistical table represents an indicator such as "utilization of pesticides", "land use", etc.
It is represented by a multidimensional matrix called cube. Each dimension is a variable. A simple table represents one indicator, while a complex table can have one or more indicators.

When statisticians in the National Secretariat organize the data for the National CORE or Sub national sections, they can use as templates the tables available on the CountrySTAT website.
They all include the variable Time.

Click to read the related section on the manual

On the other hand, when we publish data in the Thematic Modules, we could build our own statistical tables, as we shall see later.

The final published table with the essential information

Production quantity of Primary Crops
by Year, Product (tonnes)

Methodological metadata

ID	130CPD010
Title	Production quantity
Domain	Production
Unit	tonnes
Contact	Planning Department, Ministry of Agriculture and Food Security
Source	Annual Production Estimates Survey, Ministry of Agriculture and Food Security
Creation date	5/8/2012
Last update date	5/8/2012
Next update date	5/8/2012
Update frequency	Semi-Annually
Description	
Notes	2012 figures are for second round and that the variance with the third and final round is usually plus or minus 5%
Attached files	

Matrix code

Unit of measurement

Contacts for additional information

Source of data (agency compiling the data)

Last update

Description of variable values

CONFIGURATION

Columns: Year

Rows: Product

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Cassava	713,876	829,821	895,420	2,757,186	3,313,126	1,512,792	1,703,355	2,532,079	2,197,640	2,832,141	3,238,943	3,491,183	3,823,236	4,000,986	4,316,373	4,692,202
Castor	453	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]	[..]

Focal point for contact

Yasin Tabit Ahmed
tabitays@yahoo.com

Concepts, definitions and classifications

Other tef_commodity%20description.doc

Concepts and definitions adopted at international level are those of FAOSTAT except for the classification of one product: Teff. Teff crop is one of the major staple food crops in Ethiopia and it cannot be classified under any other cereal crop. It should stand alone as one of major cereal grain crop when it comes to Ethiopia. The price of teff has been collected and reported in three verities, (i.e. white, mixed and brown) since they have significantly different prices in the market.

Geographical coverage

Country

Administrative Lev 1

- Region

Surveys

Annual Agricultural Sample Survey

The CSA has been collecting data on cropland area, production and productivity for about three decades using probability based sample surveys from agricultural households under the umbrella of Rural Integrated Household Survey Program (RIHSP).

The agriculture statistics data collection on area, production, yield and farm management practice is conducted twice annually in order to capture both seasons agricultural practice of Ethiopia. These are: main season (Meher) and short rain season (Belg) survey.

Sample Design

- The primary agricultural statistics data are collected from two sources: small holding farmers and commercial farm enterprises.
- Farm enterprises are used as sampling units for commercial farm category and agricultural households (holders) are the ultimate sampling units in the case of small holding farmers.
- A stratified two-stage cluster sample design has been used to select samples of agricultural households. The agro-ecological zones and administrative divisions of the country is used to form the strata.
- Within each stratum the enumeration areas (EAs) are used as primary sampling units (PSUs) and agricultural households are secondary sampling units.
- The AgSS data is collected from more than 65,000 households in 2,280 sample areas and from over 1,700 large and medium scale commercial farms.
- The survey findings reports are produced for each administrative zones of the regional states and also aggregated at regional and country levels.

Data Collection methods

In Ethiopia, farmers are not able to provide their land size and their product in standard units and it is also difficult to convert from local units to standard units since these local units size are very much vary from area to another.
Therefore, the CSA uses objective methods to collect data on cropland acreage by applying rope and compass method and hand held GPS for area measurement and crop cutting exercises on sample plots of each crop is also done for crop yield estimation.
Direct interview method is used for collection of information on crop yield conditions, farm inputs, practices, land use, product utilization, livestock characteristics, etc.

Contents of AgSS

The annual agricultural sample survey is comprised of several questionnaires/modules administered to households at several times:

- Crop production forecast survey (cropland area , condition of crop yield , etc.). September - October.
- Agricultural practices and general information (fertilizer, pesticides, irrigation, soil conservation, etc. September -October.
- Cropland area measurement. October - November.
- Livestock survey (number, type, age, purpose, milk, egg, honey, animal, veterinary information). November
- Crop cutting exercise (yield). October- January.
- Utilization of crop and livestock product (percent used for consumption, sale, stock, etc.): January- February.
- Commercial farm survey (area, production, farm management practices, diary and fattening): February - March).
- Short rain season (Belg) agriculture survey (area, production, yield, farm input utilization, practice, etc.) April-June.

Statistical Reports of Annual Agricultural Sample Surveys

Based on annual agricultural sample survey, the nine statistical reports have been produced and disseminated annually. They are:

- Crop Production Forecast: November-December
- Area and production: April
- Farm management practices: May
- Land utilization: May
- Crop and livestock product utilization: May
- Livestock and Livestock characteristics: January-February
- Large and Medium scale commercial farms: June-July
- Belg season, area and production: August
- Belg season, farm management practices: August.

These survey findings reports are published in hard (bulletins) and electronic copies.

An example is the "Agricultural Sample Survey 2007-2008 (2000 E.C.)"

Reports are also available on the CSA website (http://www.csa.gov.et/).

Frequency of data collection

- Annual

Dissemination method

- Bulletins
- CD-ROM
  Using Micro data management toolkit
- WEB
  CountrySTAT, ENADA

Thank you !!!