



Producing Turnover Indices in Industry at INE Spain

Source of data

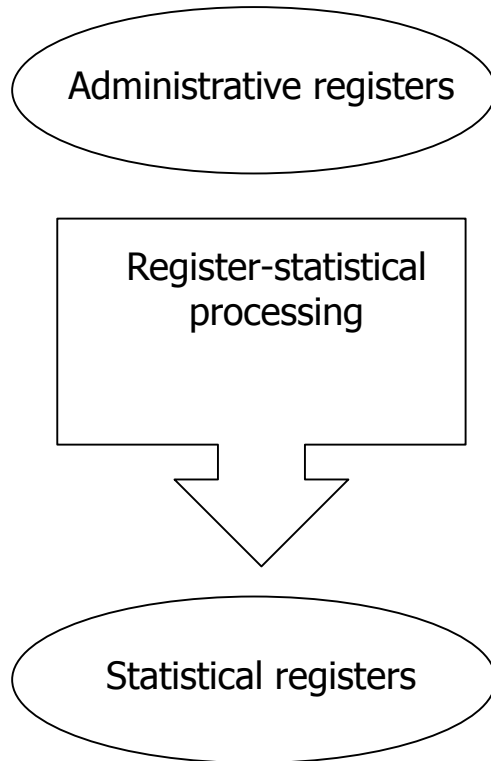
Workshop on Turnover Indices

Muscat, 22-24 May 2017

elena.rosa.perez@ine.es



1. Frame of the survey. Use of administrative registers and use of other surveys.
How are administrative data used? Which variables are used? How do we build the ITI frame?
2. Data collection. Once the sample has been selected, the data collection phase takes place.
Which data collection methods are used?
How is the data collection work organised?



At INE Spain the business register is called “Central Business Register” (CBR). Its basic objective is to be used as a frame for economic surveys.

The CBR is based on the use of administrative registers. Main sources the Tax Administration and Social Security.

Other sources: Private sources, Legal sources (Movements of the Mercantile Register), Statistical sources.

A single identifier for characterising the legal units (NIF-tax registration number) significantly facilitating the processes.



The main variables contained in the CBR are the following:

- Identification variables: trade name, postal address, telephone, fax, municipality, province, autonomous community.
- Economic variables: tax identification number, CNAE-2009 main activity code, number of employees, annual turnover.



Advantages of the Registers containing AARR data:

- Reduce the response burden on the respondent;
- Low costs;
- Almost complete coverage of population;

Disadvantages of the Registers containing AARR data:

- Dependent on the administrative system's population, object and variable definitions;
- The reporting of administrative data can be slow; the time between the reference period and when data are available for statistical purposes can be long;
- Changes in the administrative systems make comparisons difficult;



ITI statistical units are establishments, CBR contains information about businesses. So we need to look for another frame that contains regional and economic information about establishment: The Industrial Products Survey (PRODCOM).

- The Industrial Products Survey is an annual survey that collects information about a group of approx 5,000 industrial products that cover an important part of the Spanish industrial sector.
- The Prodcum frame is the CBR and the sampling unit is the company, but the information unit is the establishment.
- The collected variables are: manufactured product(s), the turnover generated by the product(s), validated identification information for each establishment of the company.



INE Spain carries out two STS-industry statistics: Index of Production in Industry (IPI) and Index of Turnover in Industry (ITI).

The initial sample was the same for IPI and ITI but differences may have arisen over time (an establishment that stops manufacturing products included in the basket of products selected in the IPI, a product included in the basket of IPI but the turnover of the establishment is too low).

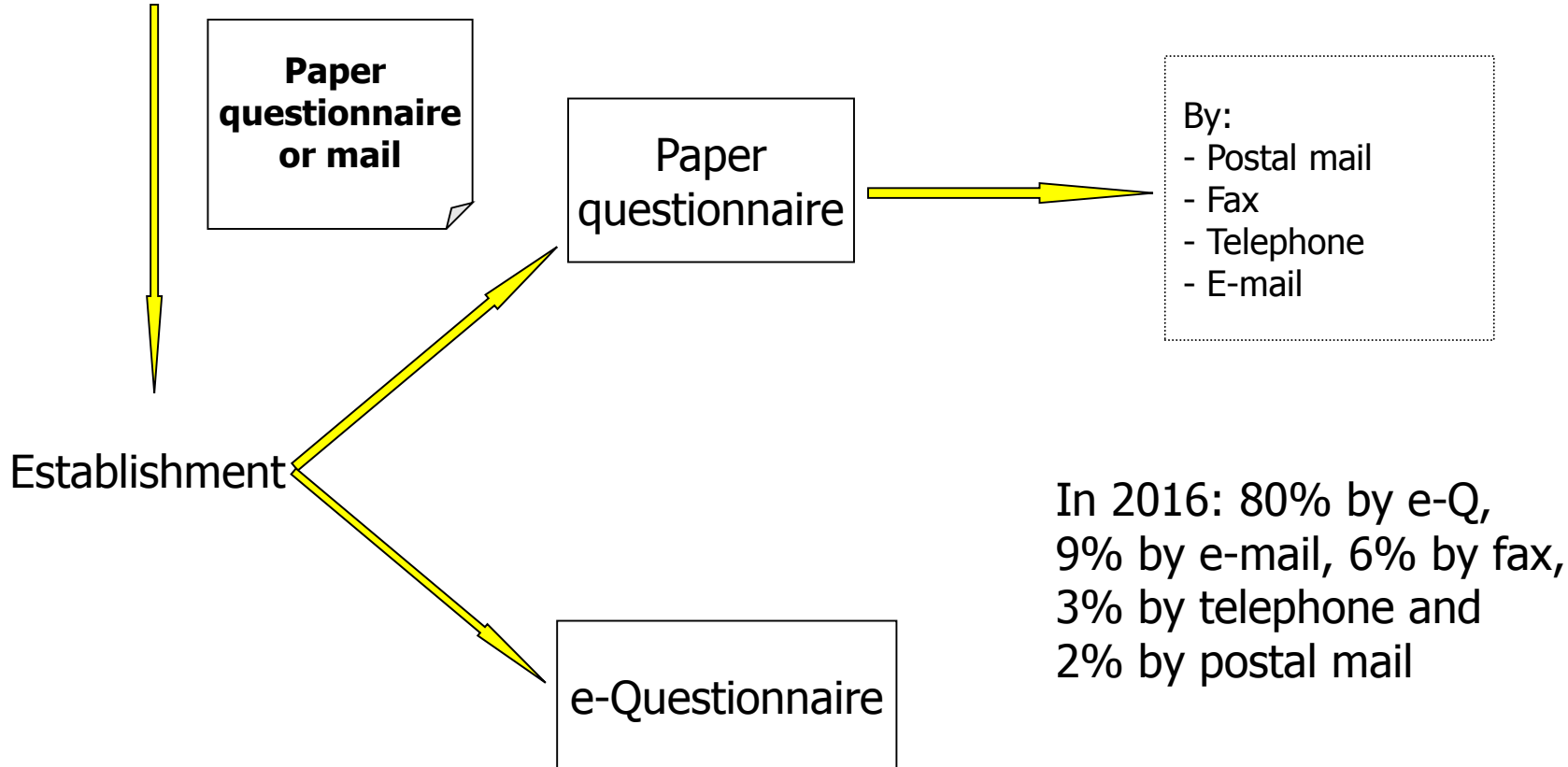
ITI and IPI sample frame is PRODCOM as it is the only frame containing information of both industrial products and the local units producing these products.

Annually a matrix is obtained indicating the % of the turnover of the sampling units in the IT Survey over the total turnover of the sampling units in the Industrial Products Survey. In the strata (NUTS2 * division/subdivision) where the value is under 80% new units are included in the ITI sample. The sampling method is cut-off, so the new units are those with a higher turnover.



INE
regional
offices

Data are required as from the last week of the reference month.
If a establishment doesn't send back the data in 10-15 days, then INE staff get in contact by mail or phone to recall the data.
If no data is provided 2 months after the reference month, then the establishment can be fined.





Instructions about
how to provide
the required data

Note including
the countries
of each market

Valor de la cifra de negocios y valor de los pedidos.

En euros sin decimales, sin incluir el IVA ni otros impuestos que gravan la operación.

NOTA IMPORTANTE PARA LOS CUESTIONARIOS DE ÍNDICES DE CIFRAS DE NEGOCIOS Y DE ENTRADAS DE PEDIDOS

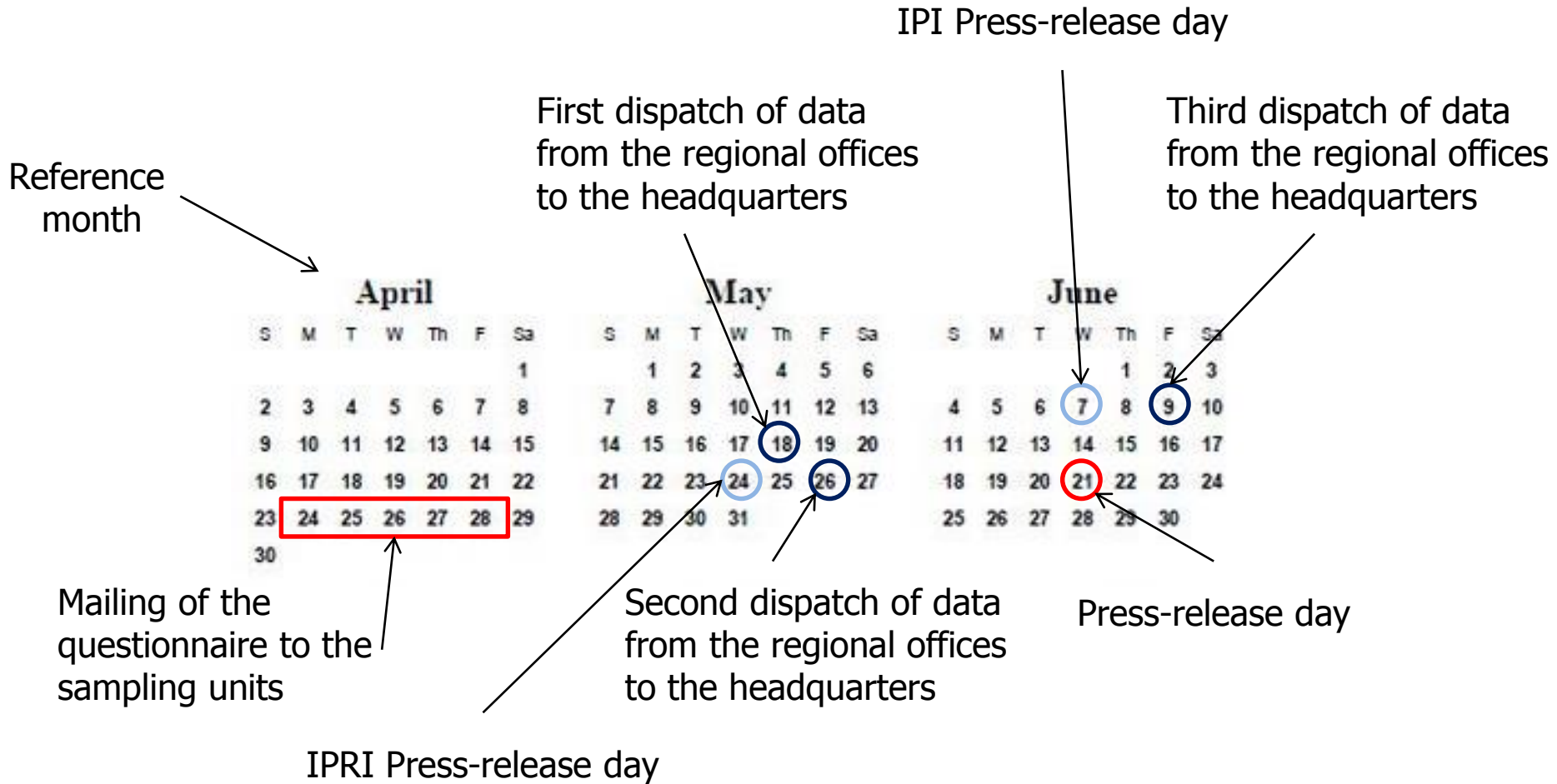
Valor de la cifra de negocios **	Valor de la cartera de pedidos al principio del mes (1)	Valor de los nuevos pedidos recibidos en el mes (2) **	Valor de los pedidos cancelados en el mes (3)	Valor de los pedidos satisfechos en el mes (4)	Valor de la cartera de pedidos al final del mes (1)+(2)-(3)-(4)
<input type="text"/> Total	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

** Desglose de la cifra de negocios por destino de las ventas y de los nuevos pedidos recibidos en el mes según su procedencia:

	Mercado interior		Mercado exterior	
			Unión Europea	Resto del mundo
		Zona Euro	Zona NO Euro	
Valor de la cifra de negocios	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Valor de los nuevos pedidos recibidos en el mes	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Breakdown
into the markets

Data collection dates



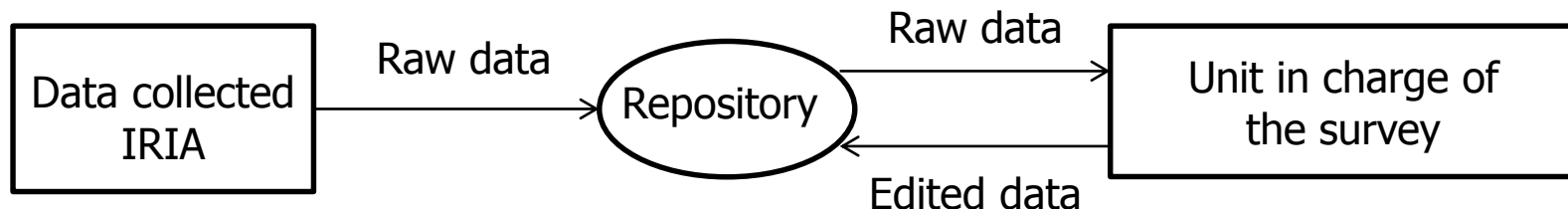


INE Spain has proposed a key-value pair data model to deal with the internal transmission of data sets in order to standardise the data warehousing.

It is a common data model for all statistical operations (standardized statistical methods and computer tools).

Key-value pair model: each value of each microdata set is univocally identified with a key.

This key contains codes of diverse standard classifications, making use of metadata. And it is valid for all statistical operations and for all kind of data (either quantitative or qualitative data, paradata, etc.).





The key is compound with the following components:

- An alphanumerical code to identify the statistical operation.
- An alphanumerical code to identify the time period of reference.
- An identifier to indicate whether they are raw or edited microdata.
- A version number to indicate whether they are provisional or definitive values.
- An identifier for the variable.
- A set of qualifiers specifying different attributes (statistical unit ID, geographical code, economic activity code,...).

A dictionary of data is created. This dictionary is kept in an XML file for machine consumption to allow for automatic reading and writing processes.



Identifier: name of the statistical variable

Qualifiers:

- In file name** : (StatSpain internal) code of the statistical operation – E30103
- In file name** : Validated value in *Final File* in data dictionary *Version 1* – FF_V1
- In file name** : Monthly time period (*Jan 2016*) – MM012016
- In file name** : Definitive (complete) set of values *Version 1* of the file/value – D_1
- ① : ID value of statistical unit – 289409300MM
- ② : NACE v2 code – 47.11
- ③ : (StatSpain internal) standardised NUTS2 code – 08

Value: value of the statistical variable – 9732

File name:

E30103.FF_V1.MM012016.D_1

Key-value pair:

Turnover	+	$\frac{\textcircled{1}}{289409300MM}$	$\frac{\textcircled{2}}{47.11}$	$\frac{\textcircled{3}}{08}$	+	9732
Identifier		Key				Value

Example for
retail trade
turnover



- Wallgren A., Wallgren B. (2007), *Register-based Statistics. Administrative Data for Statistical Purposes*, Willey, N.Y.
- Snijkers G., Haraldsen G., Jones J., Willimack D. (2013), *Designing and Conducting Business Surveys*, Willey, N.Y.
- Memobust, Eurostat: https://ec.europa.eu/eurostat/cros/content/handbook-methodology-modern-business-statistics_en
- INE Spain CBR:
<http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft37%2Fp201&file=inebase&L=1>
- INE Spain Industrial Products Survey:
http://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736149053&menu=ultiDatos&idp=1254735576715
- INE Spain Repository:
https://www.conference-service.com/NTTS2017/documents/agenda/data/abstracts/abstract_137.html



Any question??

