

## EXPERIENCE OF NATIONAL STATISTIC INSTITUTE COLLECTING DATA USING XML-FILES

9<sup>TH</sup> INTERNATIONAL FORUM ON TOURISM STATISTICS  
Paris, 19-21 November 2008

Experience / application paper

Session: New approaches to data collection from the supply side

### Authors:

- **Mr. Fernando Cortina**

Sub director of Services Statistics Directorate.  
Instituto Nacional de Estadística (Spain)  
E-mail: [fcortina@ine.es](mailto:fcortina@ine.es)

- **Mrs. Rafaela Mayo**

Sub director Assistant of Services Statistics Directorate  
Instituto Nacional de Estadística (Spain)  
E-mail: [ramayo@ine.es](mailto:ramayo@ine.es)

- **Mrs. María Velasco**

Head of Unit of Tourism Statistics  
Instituto Nacional de Estadística (Spain)  
E-mail: [mrvelasco@ine.es](mailto:mrvelasco@ine.es)

## EXPERIENCE OF NATIONAL STATISTIC INSTITUTE COLLECTING DATA USING XML-FILES

### INTRODUCTION

In statistic, for any kind of variable, as much detail is available, as much useful the information is. As it is well-known, tourism is a local phenomenon, therefore users of tourism statistics want figures with a very detailed level of breakdown.

One of the most important variables in the tourism statistics from the supply side is the number of overnight stays. There are an important demand of this variable by origin and destination of guests. The main consequence of this need of information of users is an increase of respond burden for accommodation establishments.

The best source to obtain the breakdown by origin of travellers and overnights is the accommodation establishment, but that implies that it had to fill a huge questionnaire.

Nowadays, almost all hotels and similar establishments have installed management softwares in their check desks, where the main data of guests are recorded when they check in. Those databases have the necessary information about origin of travellers, but it is too much amount of data to be provided using the normal procedures of collecting data.

On the other hand, the European Statistics Code of Practice includes as principle 9: Non-excessive burden on respondents *The reporting burden should be proportionate to the needs of the users and should not be excessive for respondents. The statistical authority monitors the respond burden and sets targets for its reduction over time.* One of the indicators of that principle points out that the information sought from business is, as far as possible, readily available from their accounts and electronic means are used where possible to facilitate its return.

Spanish Statistic Institute has participated in a working group with some regional statistic institutes in order to agree the design of a XML-file. This XML-file contains, among other variables, data of in/out/overnight stays for all the days of the month of reference, by country of residence (or region of residence, if the guest lives in Spain). With an unique XML-file, establishments can answer occupancy surveys carried out by different statistic institutes.

Therefore, the respond burden is reduced, not only for the easiness and rapidity on preparing and sending the requested information, but also, with only one file, the establishment is answering more than one official survey.

Besides, with this system of collecting data, new projects and studies could be developed. The spread of the system will enable to carry out more precise studies in small areas and by punctual events as fairs, sport events,... that take place during two or three days.

## PREPARATORY WORKS

The XML file project for Hotel Occupancy Survey was developed in the following phases.

### 1- Market research

During 2004-2005, INE Spain carried out an analysis of different management software packages installed in hotel establishment. This analysis evaluated both the structure of the sector of enterprises that offered those softwares (market share of enterprises) and the information contained in each one of them, useful for statistics purposes.

### 2- Discussion and design of the structure of the XML file

In June 2006, a working group made up of some regional statistics institutes of Spain<sup>1</sup> and National Statistics Institute (INE). The main objective of this working group was agreed the design of a XML file that could be used to collect the variables of the different hotel occupancy surveys carried out by the members of the group. This phase lasted over 15 months.

After months of meetings, the aim was reached and the file has the following structure<sup>2</sup>:

- Identification variables of the establishment (not only location variables, but also category, capacity, no. of rooms...)
- In, out and overnight stays for each day of the month of reference, broken-down by residence of guests (country, if the guest resides out of Spain, or province/island, if the guest resides in Spain).
- No. of occupied rooms each day of the reference months, distinguishing the type of room.
- Information on ADR (Average Daily Rate) depending on type of client.
- Data on employed personnel.

Different associations and hotel enterprises were consulted to check that all this information was recorded in the management softwares or was easy to obtain from other softwares used in the establishments.

During the meetings of the working group, it was discussed the problematic issue of asking variables as ADR (Average Daily Rate), RevPAR (Revenue Per Available Room) and employed personnel, because that data could be recorded in databases not included in management softwares and, moreover establishments could be reluctant of providing these specific information by this innovate way. It was decided then to divide the file in two parts. The first part includes everything except variables on prices and employment, that make up the second part of the file.

Then, if one establishment decides to send the XML file, it has to contain compulsorily the first part, having the chance of sending the second part also by

---

<sup>1</sup> In Spain, some regions carry out their own statistics for their geographical scope

<sup>2</sup> The detailed structure of the XML file can be downloaded in the following website:  
<https://arce.ine.es/ARCE/ficheros/SchemaHoteles.xsd>

XML file, or using the rest of method of transmitting the information to INE. It is important to remark that all the variables are obligatory, but what is optional is the way of sending the information (for your information, all the XML files received until now are complete, containing both parts of the file).

By other hand, some validation rules were agreed by the working group in order to ensure the coherence of the information contained in the file. If the file doesn't satisfy these rules, it can not be sent by the respondent. There aren't too many rules to make easier the process of sending the XML file.

### **3- Diffusion of the new collecting data system**

Different ways and means were used to communicate to hotel establishments and software enterprises that collecting data by XML files was going to be possible.

- A) The agreed design of XML file, the rules and some classifications were published in the INE web site ([www.ine.es](http://www.ine.es)) and in the web site of the other regional statistics institutes.
- B) Meetings with main hotel chains in Spain, that have their own management softwares, to explain the project and motivate them to used XML file. Meetings also with hotel associations to inform their members about the project.
- C) Contacts with softwares business to give them an incentive to install a new tool in the management softwares to extract the necessary data and develop the XML file for the Hotel Occupancy Survey.
- D) And letters explaining the new way of transmitting the questionnaire of Hotel Occupancy Surveys were sent to the sample units.

These meetings and contacts were organized during five months.

### **4- Test phase**

Once this new system was announced, it started a test period in order to give time to hotels and enterprises that design the management softwares to adapt them to get the XML file.

Moreover, they had the chance of testing if the files were correctly designed, the validation rules were well programmed, because INE put the application that receives the files on. During at least one month, hotels, chains and software companies were sending testing XML files, checking where there were problems (in the files or in the application).

### **5- Introduction of the new alternative to collect date**

Since May 2008, the hotel establishments in the Hotel Occupancy Survey sample can send the questionnaire through XML file.

INE send the paper questionnaire to the units in the sample, where the login and password of each establishment is included. These two data are necessary to transmit the XML-file to INE.

But the hotel establishments can choose the way of sending the questionnaire to the National Statistic Institute: postal, e-mail, Internet (connecting to a website

where the questionnaire can be fill in manually), fax and now also sending the XML-file telematically.

*A priori*, there was no information about how many establishments were going to use the new way of sending questionnaires, although based on similar experiences, it was believed that the process was going to be slow but progressive.

The methodology of the survey was also adjusted. The collection data software and the estimation program had to be ready to receive the XML files and merge these data with the ones sent by respondents by 'traditional' means.

If the establishment choose any way except XML file, it send information on a week of reference –number of travellers and overnight stays per day broken down by place of residence (see questionnaire in the annex) and if the establishment is 3, 4 or 5 gold star, it also provides information of number of travellers and overnight stays of the month of reference.

But the data contained in the XML file are much more. Hotels provide information on travellers and overnight stays for each days of the month of reference, and the list of countries has also increased. In the paper questionnaire there are only the main countries of origin for Spain, but the XML file is designed to catch any country of the world.

In the next chapter of the paper, the methodology of the survey and the new formulas are shown.

## **METHODOLOGY**

### **Presentation of the Hotel Occupancy Survey**

The results of the Hotel Occupancy Survey give notice of two aspects considered when analysing tourist trends. With regards to demand, there is information on travellers, overnight stays and average stay distributed by country of residence of the travellers and category of the establishment they are staying in or by Autonomous Community of origin in the case of travellers that live in Spain. As regard the supply, information includes the estimated number of establishments open for the season, estimated number of bedplaces, level of occupancy and information on employment in the sector, in terms of the category of the establishment.

This information is offered monthly, on a national, Autonomous Community and provincial level. Results on the most important tourist areas and sites are also disseminated.

The main objective of this survey is to measure the evolution of a set of variables which allows to describe the basic characteristics of the hotel sector.

### **Statistical unit**

This statistic is monthly and the data refer to the activity carried out by the hotels and similar establishments located within the national territory.

The population object of the study is constituted by all the hotel and similar establishments existing in Spain that are open during the month of reference. A hotel or similar accommodation is defined as establishment that provides accommodation, typically on a daily or weekly basis, principally for short stays of visitors. This includes the provision of furnished accommodation in guest rooms and suites. Services include daily cleaning and bed-making. A range of additional services may be provided such as food and beverage services, parking, laundry services, swimming pools and exercise rooms, recreational facilities as well as conference and convention facilities.

Hotel establishments are classified according to their category, gold and silver, and within these by number of stars.

### **Framework of the survey and sample design**

As framework for the selection of the sample, a national register of hotels and similar accommodation is built being the main sources of information the administrative registers of the Autonomous Communities (they have the competence to approve the opening modifications and closing licences). In the register appear, among others, the following variables for each establishment: name, address, category, normal opening period, number of bedplaces and number of rooms.

The framework is divided in strata, defined by category-province (or island). The survey is exhaustive in all provinces, except in some strata where the population is big enough to select a sample. In Balears Islands, the sample design distinguishes between high and low season due to the fact that tourism is a very seasonal phenomenon.

### Collection of the information

The enquiry pertaining to basic data refers to one full week of the month, randomly selected, in such a way that all establishments cover the whole month. A second enquiry is performed regarding the total number of travellers checked in and overnight stays caused during the month the data refer to, in the strata of three, four and five gold stars.

And, as it has been mentioned before, if the sample units send the XML file, they provide very detailed data for each day of the month of reference.

### Formulas

The formulas were re-written according to the rule that if there is no data from XML files, there is no change compared to the original ones.

The following cases had been distinguishing:

- A) Weekly information plus information from XML file (for all strata).
- B) Weekly information, monthly information and information from XML file (only for establishments whose category is 3, 4 or 5 gold stars).

*A) Weekly information plus information from XML file (for all strata)*

No of estimated open establishments in the month of reference

$$\hat{E}_{jk} = E_{jk} \frac{\left( e_{jk} + \sum_{i=1}^{e''_{jk}} \frac{dm_{ijk}}{D} \right)}{\left( e_{jk} + e''_{jk} \right) + c_{jk}}$$

whereas:

*i*, establishment;

*j*, province where the establishment is located;

*k*, establishment category;

*E*, number of open establishment in the month of reference according to directory information;

*dm*, number of days that the establishment has been opened during the month of reference;

*D*, number of days of the month of reference (28, 29, 30, 31);

*e*, number of establishments that have filled in the weekly questionnaire;

*e''*, number of establishments that have filled in the XML file;

*c*, number of sample units that are closed during its week of reference;

No. of estimated bedplaces

$$\widehat{P}_{jk} = \sum_{i=1}^{E_{jk}} P_{ijk} \cdot \frac{\left( \sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)}{\left( \sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \right) + \sum_{i=1}^{c_{jk}} P_{ijk}}$$

whereas:

$P$ , number of bedplaces according to the directory information;

No. of travellers in the month of reference

$$\widehat{V}_{jkm} = \left[ \left( \sum_{i=1}^{e_{jk}} V_{ijkm} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} V_{ijkm} \right] \cdot \frac{\widehat{P}_{jk}}{\left( \sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left[ \left( \sum_{i=1}^{e_{jk}} V_{ijkm} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} V_{ijkm} \right] \cdot \rho_{jk}$$

whereas:

$m$ , place of residence of traveller;

$V$ , number of travellers that have checked in the establishment;

$\rho_{jk}$ , elevation factor by bedplaces of province  $j$  and category  $k$

No. of overnight stays

$$\widehat{N}_{jkm} = \left[ \left( \sum_{i=1}^{e_{jk}} N_{ijkm} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} N_{ijkm} \right] \cdot \frac{\widehat{P}_{jk}}{\left( \sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left[ \left( \sum_{i=1}^{e_{jk}} N_{ijkm} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} N_{ijkm} \right] \cdot \rho_{jk}$$

whereas:

$N$ , number of overnight stays;

No of overnight stays in extra bedplaces

$$\widehat{N}'_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} N'_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} N'_{ijk} \right] \cdot \rho_{jk}$$

whereas:

$N'$ , number of overnight stays in extra bedplaces:

### Average stay by category

$$ES_{jk} = \frac{\sum_m \widehat{N}_{jkm}}{\sum_m \widehat{V}_{jkm}}$$

### Average stay by country of residence

$$ES_{jm} = \frac{\sum_k \widehat{N}_{jkm}}{\sum_k \widehat{V}_{jkm}}$$

### Number of estimated bedrooms

$$\widehat{H}_{jk} = \sum_{i=1}^{E_{jk}} H_{ijk} \cdot \frac{\left( \sum_{i=1}^{e_{jk}} H_{ijk} + \sum_{i=1}^{e''_{jk}} H_{ijk} \cdot \frac{dm_{ijk}}{D} \right)}{\left( \sum_{i=1}^{e_{jk}} H_{ijk} + \sum_{i=1}^{e''_{jk}} H_{ijk} \right) + \sum_{i=1}^{c_{jk}} H_{ijk}}$$

whereas:

$H$ , number of bedrooms according to the directory;

### Occupied bedrooms

$$\widehat{B}_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} B_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} B_{ijk} \right] \cdot \frac{\widehat{H}_{jk}}{\sum_{i=1}^{e_{jk}} H_{ijk} + \sum_{i=1}^{e''_{jk}} H_{ijk} \cdot \frac{dm_{ijk}}{D}} = \left[ \left( \sum_{i=1}^{e_{jk}} B_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} B_{ijk} \right] \cdot \beta_{jk}$$

whereas:

$B$ , number of occupied bedrooms;

$\beta_{jk}$ , elevation factor by bedrooms of province  $j$  and category  $k$

### Level of occupancy per bedplaces by category

$$\widehat{GP}_{jk} = \frac{\widehat{N}_{jk}}{D \cdot \widehat{P}_{jk} + \widehat{N}'_{jk}} \cdot 100$$

And the total:

$$\hat{GP}_j = \frac{\sum_k \hat{GP}_{jk} \cdot \hat{P}_{jk}}{\sum_k \hat{P}_{jk}}$$

Level of occupancy per bedplaces at weekend

$$GP^{fs}_{jk} = \frac{\hat{N}^{fs}_{jk}}{D^{fs} \cdot \hat{P}_{jk} + \hat{N}'^{fs}_{jk}} \cdot 100$$

whereas:

$$\hat{N}^{fs}_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} N^{fs}_{ijk} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N^{fs}_{ijk} \right] \cdot \rho_{jk}, \text{ No. of overnight stays at weekend}$$

$$\hat{N}'^{fs}_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} N'^{fs}_{ijk} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N'^{fs}_{ijk} \right] \cdot \rho_{jk}, \text{ No of overnight stays in extra}$$

bedplaces at weekend

$N^{fs}$ , number overnight stays at weekend;

$N'^{fs}$ , number of overnight stays in extra bedplaces at weekend;

$D^{fs}$ , number of Fridays and Saturdays in the month of reference;

And the total

$$\hat{GP}^{fs}_j = \frac{\sum_k \hat{GP}^{fs}_{jk} \cdot \hat{P}_{jk}}{\sum_k \hat{P}_{jk}}$$

Level of occupancy per bedroom

$$\hat{GH}_{jk} = \frac{\hat{B}_{jk}}{D \cdot \hat{H}_{jk}} \cdot 100$$

And the total

$$\hat{GH}_j = \frac{\sum_k \hat{GH}_{jk} \cdot \hat{H}_{jk}}{\sum_k \hat{H}_{jk}}$$

B) Weekly information, monthly information plus information from XML file (for all strata)

No of estimated open establishments in the month of reference

$$\widehat{EM}_{jk} = E_{jk} \frac{\left( \sum_{i=1}^{e'_{jk}} \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} \frac{dm_{ijk}}{D} \right)}{(e'_{jk} + e''_{jk}) + c'_{jk}}$$

whereas:

$i$ , establishment;

$j$ , province where the establishment is located;

$k$ , establishment category;

$E$ , number of open establishment in the month of reference according to directory information;

$dm$ , number of days that the establishment has been opened during the month of reference;

$D$ , number of days of the month of reference (28, 29, 30, 31);

$e'$ , number of establishments that have filled in the monthly questionnaire;

$e''$ , number of establishments that have filled in the XML file;

$c'$ , number of sample units that are closed in the month of reference;

No of estimated bedplaces

$$\widehat{PM}_{jk} = \sum_{i=1}^{E_{jk}} P_{ijk} \cdot \frac{\left( \sum_{i=1}^{e'_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)}{\left( \sum_{i=1}^{e'_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \right) + \sum_{i=1}^{c'_{jk}} P_{ijk}}$$

whereas:

$P$ , number of bedplaces according to directory information;

No of travellers in the month of reference

$$\widehat{VM}_{jk} = \left( \sum_{i=1}^{e'_{jk}} VM_{ijk} + \sum_{i=1}^{e''_{jk}} V_{ijk} \right) \cdot \frac{\widehat{PM}_{jk}}{\left( \sum_{i=1}^{e'_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left( \sum_{i=1}^{e'_{jk}} VM_{ijk} + \sum_{i=1}^{e''_{jk}} V_{ijk} \right) \cdot \alpha_{jk}$$

whereas:

$VM$ , total number of travellers that have checked in the establishment in the month of reference (information of the monthly questionnaire);

$\alpha_{jk}$ , elevation factor by bedplaces of province  $j$  and category  $k$

$$\widehat{VM}_{jkm} = \widehat{VM}_{jk} \cdot \frac{\widehat{V}_{jkm}}{\sum_m \widehat{V}_{jkm}}, \text{ (} m: \text{ place of residence of travellers)}$$

#### No of overnight stays

$$\widehat{NM}_{jk} = \left( \sum_{i=1}^{e'_{jk}} NM_{ijk} + \sum_{i=1}^{e''_{jk}} N_{ijk} \right) \cdot \frac{\widehat{PM}_{jk}}{\left( \sum_{i=1}^{e'_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left( \sum_{i=1}^{e'_{jk}} NM_{ijk} + \sum_{i=1}^{e''_{jk}} N_{ijk} \right) \cdot \alpha_{jk}$$

whereas:

$NM$ , total number of overnight stays in the month of reference (information of the monthly questionnaire);

$$\widehat{NM}_{jkm} = \widehat{NM}_{jk} \cdot \frac{\widehat{N}_{jkm}}{\sum_m \widehat{N}_{jkm}}, \text{ (} m: \text{ place of residence of travellers)}$$

#### No of overnight stays in extra bedplaces

$$\widehat{NM}'_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} N'_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} N'_{ijk} \right] \cdot \frac{\widehat{PM}_{jk}}{\left( \sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left[ \left( \sum_{i=1}^{e_{jk}} N'_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} N'_{ijk} \right] \cdot \varpi_{jk}$$

whereas:

$N'$ , number of overnight stays in extra bedplaces;

$\varpi_{jk}$ , elevation factor by bedplaces of province  $j$  and category  $k$

#### Average stay by category

$$ES\widehat{M}_{jk} = \frac{N\widehat{M}_{jk}}{V\widehat{M}_{jk}}$$

#### Average stay by country of residence

$$ES\widehat{M}_{jm} = \frac{\sum_k N\widehat{M}_{jkm}}{\sum_k V\widehat{M}_{jkm}}$$

Number of estimated bedrooms

$$\widehat{HM}_{jk} = \sum_{i=1}^{E_{jk}} H_{ijk} \cdot \frac{\left( \sum_{i=1}^{e'_{jk}} H_{ijk} \cdot \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} H_{ijk} \cdot \frac{dm_{ijk}}{D} \right)}{\left( \sum_{i=1}^{e'_{jk}} H_{ijk} + \sum_{i=1}^{e''_{jk}} H_{ijk} \right) + \sum_{i=1}^{c'_{jk}} H_{ijk}}$$

whereas

$H$ , number of bedrooms according to directory information;

Occupied bedrooms

$$\widehat{BM}_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} B_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} B_{ijk} \right] \cdot \frac{\widehat{HM}_{jk}}{\left( \sum_{i=1}^{e_{jk}} H_{ijk} + \sum_{i=1}^{e''_{jk}} H_{ijk} \cdot \frac{dm_{ijk}}{D} \right)} = \left[ \left( \sum_{i=1}^{e_{jk}} B_{ijk} \cdot \frac{D}{7} \right) + \sum_{i=1}^{e''_{jk}} B_{ijk} \right] \cdot \delta_{jk}$$

whereas

$B$ , number of occupied bedrooms;

$\delta_{jk}$ , elevation factor by bedrooms of province  $j$  and category  $k$

Level of occupancy per bedplaces by category

$$\widehat{GPM}_{jk} = \frac{\widehat{NM}_{jk}}{D \cdot \widehat{PM}_{jk} + \widehat{NM}'_{jk}} \cdot 100$$

And the total

$$\widehat{GPM}_j = \frac{\sum_k \widehat{GPM}_{jk} \cdot \widehat{PM}_{jk}}{\sum_k \widehat{PM}_{jk}}$$

Level of occupancy per bedplaces at weekend

$$\widehat{GPM}^{fs}_{jk} = \frac{\widehat{NM}^{fs}_{jk}}{D^{fs} \cdot \widehat{PM}_{jk} + \widehat{NM}'^{fs}_{jk}} \cdot 100$$

whereas:

$$\widehat{NM}^{fs}_{jk} = \left[ \left( \sum_{i=1}^{e_{jk}} N_{ijk}^{fs} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N_{ijk}^{fs} \right] \cdot \varpi_{jk}, \text{ No. of overnight stays at weekend;}$$

$$\widehat{NM}'_{jk}{}^{fs} = \left[ \left( \sum_{i=1}^{e_{jk}} N'_{ijk}{}^{fs} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N'_{ijk}{}^{fs} \right] \cdot \varpi_{jk}, \text{ No. of overnight stays in extra}$$

bedplaces at weekend;

$N^{fs}$ , number of overnight stays at weekend;

$N'{}^{fs}$ , number of overnight stays in extra bedplaces at weekend;

$D^{fs}$ , number of Fridays and Saturdays in the month of reference;

And the total

$$\widehat{GPM}'_j{}^{fs} = \frac{\sum_k \widehat{GPM}'_{jk}{}^{fs} \cdot \widehat{PM}_{jk}}{\sum_k \widehat{PM}_{jk}}$$

Level of occupancy per bedroom

$$\widehat{GHM}'_{jk} = \frac{\widehat{BM}_{jk}}{D \cdot \widehat{HM}_{jk}} \cdot 100$$

And the total

$$\widehat{GHM}'_j = \frac{\sum_k \widehat{GHM}'_{jk} \cdot \widehat{HM}_{jk}}{\sum_k \widehat{HM}_{jk}}$$

Analysing the formulas used to calculate the results of the survey, it can be stated that if only establishments which category is three, four or five gold star send XML files, the total number of travellers and overnight stays don't change because they have to provide anyway the monthly information. There are then changes in the breakdown by place of residence of travellers.

But if establishments which other category send XML files, the total estimations will vary, because it is introduced real monthly information in the formulas and with the 'traditional questionnaire' the monthly information was estimated based on the weekly data.

## FIRST RESULTS

The first XML files were received with data of June 2008. Ten establishments decided to send the requested information using this new method. In July, there were 13 establishments and in August, 17. It is known that the process is going to be slow, but the number is increasing and there are many hotel establishments and hotel chains (with many establishments) interested in using this way of sending data.

As it was mentioned, to have information on guests and overnight stay of each day of the month of reference, affects to the breakdown by residence of guests, but also to the total number of guests and overnight stays when establishments of low categories send the XML, because they don't have to send monthly information if they use the paper questionnaire.

We can compare the estimations obtained with the whole information with the one calculated using only the information of the seven days asked in the paper questionnaire. The published results for these months are the followings ones (XML):

		No. of guests			No. of overnight stays			Average stay
		Total	Residents in Spain	Residents abroad	Total	Residents in Spain	Residents abroad	
<b>June</b>	XML	8.329.961	4.457.989	3.871.973	27.817.756	10.462.696	17.355.059	3,34
	Week	8.329.832	4.458.524	3.871.308	27.817.563	10.463.051	17.354.512	3,34
	Difference	129	-535	665	193	-355	547	0,00
<b>July</b>	XML	9.235.192	4.978.226	4.256.966	34.614.439	13.604.690	21.009.749	3,75
	Week	9.235.299	4.977.789	4.257.510	34.614.707	13.604.369	21.010.338	3,75
	Difference	-107	437	-544	-268	321	-589	0,00
<b>August</b>	XML	9.954.128	5.578.487	4.375.641	37.922.603	16.813.251	21.109.352	3,81
	Week	9.954.097	5.578.736	4.375.361	37.922.525	16.811.323	21.111.202	3,81
	Difference	31	-249	280	78	1928	-1850	0,00

Although there are slight differences in the total estimations, the real differences are in the breakdown by residence. Having available the information about all the days of the month of reference, and (this is the most important) taking this information directly from the management systems of hotels, the quality of these estimations increases considerably.

If these differences, in future, are significant, then a system of link coefficient should be calculate in order of not having broken series.

# ANNEX



### Identificación del establecimiento

#### 1. Modificaciones en la identificación (Cumplimente sólo los apartados sujetos a variación)

Nombre o razón social del establecimiento		NIF	
Signatura	Dirección		
Código Postal	Municipio	Provincia	
Habitaciones	Plazas	Categoría	Teléfono
			Fax
Cadena hotelera a la que pertenece			

**Este cuestionario se enviará, una vez cumplimentado, en los CINCO DÍAS SIGUIENTES a los que se refieren los datos, a la Delegación Provincial del Instituto Nacional de Estadística, a la que podrá dirigirse para cualquier aclaración.**

**Persona de contacto a quien dirigirse, en caso necesario, para consultas, aclaraciones o modificaciones sobre este cuestionario.**

Sr. D.: \_\_\_\_\_

Cargo o puesto que ocupa: \_\_\_\_\_

Teléfono: \_\_\_\_\_ Fax: \_\_\_\_\_

E-mail: \_\_\_\_\_

FIRMA O SELLO DEL ESTABLECIMIENTO

#### Naturaleza, características y finalidad

Esta estadística se enmarca dentro de las Encuestas de Ocupación en Alojamientos Turísticos. Los objetivos son ampliar y actualizar la información de los establecimientos del sector según directrices de la Unión Europea.

#### Legislación cumplimentación obligatoria

#### Estadística de

##### Secreto Estadístico

Serán objeto de protección y quedarán amparados por el **secreto estadístico**, los datos personales que obtengan los servicios estadísticos tanto directamente de los informantes como a través de fuentes administrativas (art. 13.1 de la Ley de la Función Estadística Pública de 9 de mayo de 1989, (LFEP)). Todo el personal estadístico tendrá la obligación de preservar el secreto estadístico (art. 17.1 de la LFEP).

##### Obligación de facilitar los datos

Las Leyes 4/1990 y 13/1996 establecen la **obligación de facilitar los datos** que se soliciten para la elaboración de esta Estadística. Los servicios estadísticos podrán solicitar datos de todas las personas físicas y jurídicas, nacionales y extranjeras, residentes en España (artículo 10.1 de la LFEP).

Todas las personas físicas y jurídicas que suministren datos, tanto si su colaboración es obligatoria como voluntaria, **deben contestar de forma veraz, exacta, completa y dentro del plazo** a las preguntas ordenadas en la debida forma por parte de los servicios estadísticos (art. 10.2 de la LFEP).

**El incumplimiento** de las obligaciones establecidas en esta Ley, en relación con las estadísticas para fines estatales, **será sancionado** de acuerdo con lo dispuesto en las normas contenidas en el presente Título (art. 48.1 de la LFEP).

Las infracciones muy graves serán sancionadas con multas de **3.005,07 a 30.050,61 €**. Las infracciones graves serán sancionadas con multas de **300,52 a 3.005,06 €**. Las infracciones leves se sancionarán con multas de **60,10 a 300,51 €** (art. 51.1, 51.2 y 51.3 de la LFEP).

Nota: este cuestionario está disponible en las distintas lenguas cooficiales de las comunidades autónomas.

Mod.: EOH1-08

## 2. Entrada de viajeros

Se entiende por viajero toda persona que llega y se aloja en el establecimiento. Se anotarán sólo las personas que entran cada día, aunque pernocten más de un día.

Lugar de residencia	Día 1°	Día 2°	Día 3°	Día 4°	Día 5°	Día 6°	Día 7°
<b>Unión Europea</b>							
<b>ESPAÑA</b>							
1.1 Andalucía							
1.2 Aragón							
1.3 Asturias (Principado de)							
1.4 Balears (Illes)							
1.5 Canarias							
1.6 Cantabria							
1.7 Castilla y León							
1.8 Castilla - La Mancha							
1.9 Cataluña							
1.10 Comunitat Valenciana							
1.11 Extremadura							
1.12 Galicia							
1.13 Madrid (Comunidad de)							
1.14 Murcia (Región de)							
1.15 Navarra (Comunidad Foral de)							
1.16 País Vasco							
1.17 Rioja (La)							
1.18 Ceuta y Melilla							
<b>RESTO DE LA UNIÓN EUROPEA</b>							
2. Alemania							
3. Austria							
4. Bélgica							
5. Bulgaria							
6. Chipre							
7. Dinamarca							
8. Eslovaquia							
9. Eslovenia							
10. Estonia							
11. Finlandia							
12. Francia							
13. Grecia							
14. Hungría							
15. Irlanda							
16. Italia							
17. Letonia							
18. Lituania							
19. Luxemburgo							
20. Malta							
21. Países Bajos							
22. Polonia							
23. Portugal							
24. Reino Unido							
25. República Checa							
26. Rumania							
27. Suecia							
<b>RESTO DE PAÍSES</b>							
28. Islandia							
29. Liechtenstein							
30. Noruega							
31. Rusia							
32. Suiza							
33. Turquía							
34. Ucrania							
35. Resto de Europa							
36. Argentina							
37. Brasil							
38. Canadá							
39. Estados Unidos							
40. Méjico							
41. Resto de América							
42. Sudáfrica							
43. Resto países africanos							
44. Japón							
45. República China							
46. República de Corea del Sur							
47. Resto de Asia							
48. Australia							
49. Resto del mundo							

### 3. Plazas ocupadas (número de pernoctaciones)

Se entiende por plaza ocupada o pernoctación cada noche que un viajero se aloja en el establecimiento. Se anotará el número de personas que duermen cada noche, independientemente del día en que entraron. *Ej.: Un matrimonio con un hijo, dará lugar a tres pernoctaciones o plazas ocupadas.*

Lugar de residencia	Día 1º	Día 2º	Día 3º	Día 4º	Día 5º	Día 6º	Día 7º
<b>Unión Europea</b>							
<b>ESPAÑA</b>							
1.1 Andalucía							
1.2 Aragón							
1.3 Asturias (Principado de)							
1.4 Balears (Illes)							
1.5 Canarias							
1.6 Cantabria							
1.7 Castilla y León							
1.8 Castilla - La Mancha							
1.9 Cataluña							
1.10 Comunitat Valenciana							
1.11 Extremadura							
1.12 Galicia							
1.13 Madrid (Comunidad de)							
1.14 Murcia (Región de)							
1.15 Navarra (Comunidad Foral de)							
1.16 País Vasco							
1.17 Rioja (La)							
1.18 Ceuta y Melilla							
<b>RESTO DE LA UNIÓN EUROPEA</b>							
2. Alemania							
3. Austria							
4. Bélgica							
5. Bulgaria							
6. Chipre							
7. Dinamarca							
8. Eslovaquia							
9. Eslovenia							
10. Estonia							
11. Finlandia							
12. Francia							
13. Grecia							
14. Hungría							
15. Irlanda							
16. Italia							
17. Letonia							
18. Lituania							
19. Luxemburgo							
20. Malta							
21. Países Bajos							
22. Polonia							
23. Portugal							
24. Reino Unido							
25. República Checa							
26. Rumania							
27. Suecia							
<b>RESTO DE PAÍSES</b>							
28. Islandia							
29. Liechtenstein							
30. Noruega							
31. Rusia							
32. Suiza							
33. Turquía							
34. Ucrania							
35. Resto de Europa							
36. Argentina							
37. Brasil							
38. Canadá							
39. Estados Unidos							
40. Méjico							
41. Resto de América							
42. Sudáfrica							
43. Resto países africanos							
44. Japón							
45. República China							
46. República de Corea del Sur							
47. Resto de Asia							
48. Australia							
49. Resto del mundo							

#### 4. Habitaciones ocupadas y plazas supletorias

	Día 1°	Día 2°	Día 3°	Día 4°	Día 5°	Día 6°	Día 7°
4.1 Número de habitaciones ocupadas							
4.2 Plazas supletorias y otras plazas adicionales utilizadas (no incluidas en la capacidad del hotel)							
	<b>Total</b>	<b>Dobles (uso doble)</b>	<b>Dobles (uso individual)</b>	<b>Otras</b>			
4.3 Número de habitaciones-día ocupadas <sup>1</sup>							

<sup>1</sup> Ej.: Si una habitación doble con uso doble se ocupa los 7 días, en el apartado correspondiente se anotará 7.

#### 5. Personal ocupado (no debe incluir personal de vacaciones, de baja laboral, ...)

	Personal remunerado		
	Personal no remunerado	Fijo	Eventual
Personal ocupado durante la semana de referencia			

#### 6. Tarifa Promedio Diaria (ADR) por tipo de cliente (en euros, no incluye IVA)

Indique el ADR (Average Daily Rate o Tarifa Promedio Diaria) para cada tipo de cliente por día incluyendo sólo alojamiento –sin desayuno, media pensión o pensión completa- de una habitación doble con baño.  
 Señale también el porcentaje aproximado de habitaciones ocupadas por cada tipo de cliente sobre el total – el mismo del apartado 4.3- en la semana de referencia (este porcentaje puede ser cero en algún caso).

	ADR en euros	% <sup>2</sup>
Tour operador tradicional		
Agencia de viajes tradicional <sup>1</sup> (incluye bonos y talones)		
Empresas		
Particulares (incluye tarifa normal y tarifa fin de semana)		
Grupos		
Contratación directa en la web del hotel o en la cadena hotelera		
Tour operador on-line		
Agencia de viajes on-line		
Otros		

<sup>1</sup> Deben considerar en este tipo de cliente los huéspedes que han entregado bonos y/o talones de hoteles como contraprestación por los servicios prestados.

<sup>2</sup> Este porcentaje será cero si no ha habido habitaciones ocupadas durante la semana de referencia.

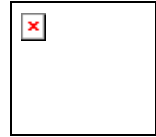
100 %<sup>2</sup>

Observaciones: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

**Gracias por su**

# Encuesta de Ocupación en Alojamientos Turísticos. Establecimientos Hoteleros

## Datos mensuales



### Identificación del establecimiento

Este cuestionario se enviará, una vez cumplimentado, en los **cinco días naturales siguientes** a los que se refieren los datos, a la Delegación Provincial del Instituto Nacional de Estadística, a la que podrá dirigirse para cualquier aclaración.

### Número de viajeros y pernoctaciones en el mes de referencia

Indique el número total de viajeros y pernoctaciones que se registraron en su establecimiento en el mes de referencia.

	Viajeros	Pernoctaciones
Total	_____	_____

### Periodo de apertura en el mes de referencia

Indique el número de días que ha estado abierto su establecimiento en el mes de referencia.

Número de días \_\_\_\_\_

(Firma o sello del establecimiento)

**Gracias por su colaboración**

#### Naturaleza, características y finalidad

Esta estadística se enmarca dentro de la Encuesta de Ocupación en Alojamientos Turísticos. Los objetivos son ampliar y actualizar la información de los establecimientos del sector según directrices de la Unión Europea.

#### Legislación obligatoria

#### Estadística de cumplimentación

##### Secreto Estadístico

Serán objeto de protección y quedarán amparados por el **secreto estadístico**, los datos personales que obtengan los servicios estadísticos tanto directamente de los informantes como a través de fuentes administrativas (art. 13.1 de la Ley de la Función Estadística Pública de 9 de mayo de 1989, (LFEP)). Todo el personal estadístico tendrá la obligación de preservar el secreto estadístico (art. 17.1 de la LFEP).

##### Obligación de facilitar los datos

Las Leyes 4/1990 y 13/1996 establecen la **obligación de facilitar los datos** que se soliciten para la elaboración de esta Estadística.

Los servicios estadísticos podrán solicitar datos de todas las personas físicas y jurídicas, nacionales y extranjeras, residentes en España (artículo 10.1 de la LFEP).

Todas las personas físicas y jurídicas que suministren datos, tanto si su colaboración es obligatoria como voluntaria, **deben contestar de forma veraz, exacta, completa y dentro del plazo** a las preguntas ordenadas en la debida forma por parte de los servicios estadísticos (art. 10.2 de la LFEP).

El **incumplimiento** de las obligaciones establecidas en esta Ley, en relación con las estadísticas para fines estatales, **será sancionado** de acuerdo con lo dispuesto en las normas contenidas en el presente Título (art. 48.1 de la LFEP).

Las infracciones muy graves serán sancionadas con multas de **3.005,07 a 30.050,61 euros**. Las infracciones graves serán sancionadas con multas de **300,52 a 3.005,06 euros**. Las infracciones leves se sancionarán con multas de **60,10 a 300,51 euros** (art. 51.1, 51.2 y 51.3 de la LFEP).

Nota: este cuestionario está disponible en las distintas lenguas cooficiales de las comunidades autónomas.

Mod. EOHM-08