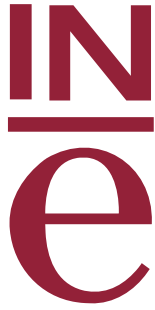


INSTITUTO NACIONAL DE ESTADISTICA



Hotel Occupancy Survey 2019

Methodology

Index

1. Presentation
2. Objectives
3. Statistical unit
4. Survey scope
5. Definitions
6. Framework of the survey and sample design
7. Estimators
8. Collection of the information
9. Dissemination of the information
10. Statistical secrecy and significance
11. Variation coefficients
12. Linking coefficients
13. Treatment of the partial non-response

1. Presentation

This publication presents the results corresponding to the Hotel Occupancy Survey (formerly the Survey on the Movement of Guests in Hotel Establishments).

The data provided offers information on the two aspects considered when analysing tourist trends: with regard to demand, there is information on guests, overnight stays and average stay, distributed by country of residence of the guests and category of the establishment they are staying in, or by Autonomous Community of origin in the case of Spanish guests. As regards supply, the information includes the estimated number of establishments open for the season, the estimated number of bedplaces, the occupancy rate, and the information on employment in the sector, in terms of the category of the establishment.

This information is offered monthly, on national, Autonomous Community, provincial and tourist area levels.

The National Statistics Institute owes a debt of gratitude to all the Managers and Owners of the Hotel Establishments that have taken part, for their help, without which this publication would not have been possible.

2. Objectives

The main objective of the Hotel Occupancy Survey is to ascertain the behaviour of a series of variables which allow us to describe the basic characteristics of the hotel sector.

3. Statistical unit

This statistical operation is monthly, and the data refers to the activity carried out by the hotel establishments located within the national territory.

The target population of the study consists of all the hotel establishments existing within the national territory, registered as such in the corresponding Tourist Offices of the Autonomous Communities. A hotel establishment is understood to be all units that render hotel accommodation services (hotel, apartment hotel, motel, hostel, B&B, boarding house, guest house), situated in the same geographic location and in which one or more persons work on behalf of the same company. Hotel establishments are classified according to their category, gold and silver, and within these, by the number of stars. The category of the establishment is assigned by the Tourist Offices of the Autonomous Communities, and this varies from one Autonomous Community to another.

4. Survey scope

All hotel establishments located in the national territory are investigated.

5. Definitions

5.1 Estimated number of open hotel establishments

A hotel establishment open for the season is understood to be that establishment in which the month of reference is included within its opening period.

5.2 Estimated bedrooms

The number of bedrooms estimated by the survey of the hotel establishments open for the season.

A bedroom is the room or group of rooms constituting an indivisible unit intended for being leased as accommodation. Bedrooms may be single rooms, double rooms or multiple rooms, depending on them being equipped to shelter one, two or more persons.

5.3 Estimated bedplaces

The number of bedplaces estimated by the survey of the hotel establishments open for the season.

The number of bedplaces is equal to the number of fixed beds in the establishment. Extra beds are not included and double beds are considered as two bedplaces.

5.4 Guests

All persons who stay for one or more consecutive nights in the same accommodation.

Guests are classified by their country of residence, and persons residing in Spain are classified by the Autonomous Community they regularly live in.

5.5 Overnight stays or occupied bedplaces

An overnight stay or occupied bedplace is understood to be each night that a guest stays in the establishment.

As occurs with guests, overnight stays are broken down according to place of residence.

5.6 Average stay

This variable is an approximation to the number of days which, on average, the guests stay at the establishments, and is calculated as the quotient between the overnight stays and the number of guests.

5.7 Occupancy rate by room

The ratio, as a percentage, between the average daily number of rooms occupied in the month and the total number of rooms available.

5.8 Occupancy rate by bedplaces

Ratio, as a percentage, between the total number of overnight stays and the product of the bedplaces, including extra beds, by the days the overnight stays refer to.

Extra beds are those that do not have a fixed nature and are not included in the bedplaces declared officially by the establishment but do appear in the directory.

5.9 Occupancy rate by bedplaces at weekends

Ratio, as a percentage, between overnight stays on Friday and Saturday that are in the reference week and the product of the bedplaces, including extra beds for those two days, by the days the overnight stays refer to, in this case, two.

5.10 Extra beds

Extra beds are understood to mean all those that are not fixed and that are not counted amongst the official number of bed places declared by the establishment and that appear in the directory.

The establishment's fixed furniture (sofa bed, fold-out bed etc.) only counts as an extra bed when it is used as such.

Equally, cots will also be considered to be extra beds.

5.11 Employed personnel

Defined as the group of paid and unpaid persons who contribute with their work to the production of goods and services in the establishment, during the reference period of the survey, even when they work outside the premises.

5.12 Tourist site

Municipality with a significant tourist offer.

5.13 Tourist area

The group of municipalities in which there is an abundance of tourists. Information on the main areas of tourist interest is offered.

In the document **List of municipalities included in each area** , the main tourism areas of interest are detailed, along with the municipalities that form these areas.

6. Framework of the survey and sample design

As the framework for the selection of the sample the directories of the Tourist Offices of the Autonomous Communities and other sources will be used, in which, amongst others, the following data appears, for each establishment: name, address, category, normal opening period, number of bedplaces and rooms.

Estos directorios se actualizan permanentemente.

The survey's sample design is defined by the pair (stratum) {province, category}.

The survey is exhaustive in all provinces, except in some categories for the provinces detailed in the document [Sampling fractions](#).

In addition, there are establishments belonging to a non-exhaustive stratum, due to the province in which they are located and the category associated with them, which qualify as exhaustive by having a high number of bed places in relation to the total number of bed places that exist in the stratum {province, category} to which they belong.

7. Estimates

The estimators' expressions are described, taking into account the information received from the XML files.

The XML files are generated directly in the establishments' management systems following an outline and validations published by the INE. This file provides detailed information on the number of travellers and overnight stays for all days of the reference month, broken down by traveller's residence (NUTS III in the case of travellers residing in Spain, and country, in the case of non-residents).

The variables used in the Hotel Occupancy Survey are as follows:

E = number of establishments open during the month, existing in the directory

e = number of establishments that respond to the survey by means of the usual method, (incidents 1 and 2)

e' = number of establishments that answer the monthly survey by means of the usual method, (with incidents 1 and 2 in the monthly questionnaire)

e'' = number of establishments sending the XML file (incidences 1 and 2)

c = number of sample establishments that are closed within their opening period and responding by means of the usual method (incident 3)

c' = number of establishments that are closed within their opening period and responding by means of the usual method (with incident 3 in the monthly questionnaire)

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

D^{fs} = number of Fridays and Saturdays in the reference month

dm = number of days the establishment has been open during the reference month

P = number of bedplaces according to the directory

P' = number of supplementary bedplaces used

H = number of rooms according to the directory

V = number of guests checked in

VM = number of guests checked in during the whole month

N = number of occupied bedplaces (overnight stays)

NM = number of occupied bedplaces (overnight stays) during the whole month

EM = average stay

T = employed personnel

B = number of occupied rooms

BD = number of occupied double rooms

BI = number of occupied single rooms

GP = occupancy rate by bedplaces

GH = occupancy rate by rooms

The **subindices** used are:

i = establishment (hotel, campsite, apartment)

j = province

k = category (5, 4, 3, 2, 1 gold: 3, 2 and 1 silver)

m = modality (professional situation, place of residence, month, etc.)

We distinguish between two groups of **estimators**:

- Weekly information plus XML (for all strata).
- Weekly information, monthly information plus XML (only 3, 4 and 5 gold-star establishments plus Galicia).

Group A estimators: Weekly information plus XML (for all strata)

A) Estimate of the number of establishments open in the month

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

B) Estimate of the number of bedplaces in the establishments open in the month

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

C) Estimate of number of guests checked in during the month

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

m = place of residence

D) Estimate of occupied bedplaces(Overnight stays)

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

m = place of residence

E) Estimate of the number of extra bed places occupied

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

where ρ_{jk} is the elevation factor by bed places for province j and category k.

F) Estimated average stay

a) Average stay per category

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

b) Average stay by country of residence

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

m = place of residence

G) Estimate of employed personnel

$$\widehat{T}_{jkm} = \left(\sum_{i=1}^{e_{jk}} T'_{ijkm} + \sum_{i=1}^{e''_{jk}} T'_{ijkm} \right) \cdot \rho_{jk}$$

where $T'_{ijkm} = T_{ijkm} \cdot AC_{ijk}$, where AC_{ijk} is the percentage of employed personnel dedicated to the activity of the accommodation

m = professional situation

H) Estimate of the number of occupied rooms

Previously, it was necessary to calculate the number of existing rooms in open establishments

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

a) Total occupied rooms

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

where β_{jk} is the elevation factor by rooms for province j and category k

b) Occupied rooms with double use

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

where β_{jk} is the elevation factor by rooms for province j and category k

c) Occupied rooms with single use

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

where β_{jk} is the elevation factor by rooms for province j and category k

I) Estimate of the occupancy rate (percentage)

a) Occupancy rate by capacity

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

For the set of categories

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

m = place of residence

b) Level of occupation by rooms

Beforehand, it was necessary to calculate the number of rooms in establishments open:

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

and the number of rooms occupied:

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

then,

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

and for the set of categories

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

c) Weekend occupancy rate

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

where:

$$\widehat{N}^{fs}_{jk} = \left[\left(\sum_{i=1}^{e_{jk}} N^{fs}_{ijk} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N^{fs}_{ijk} \right] \cdot \rho_{jk}, \text{ bedplaces occupied at weekends:}$$

$$\widehat{N}^{fs}_{jk} = \left[\left(\sum_{i=1}^{e_{jk}} N^{fs}_{ijk} \cdot \frac{D^{fs}}{2} \right) + \sum_{i=1}^{e''_{jk}} N^{fs}_{ijk} \right] \cdot \rho_{jk}, \text{ supplementary beds occupied at weekends}$$

m = place of residence

For the set of categories

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

Group B estimators: Weekly information, monthly information plus XML

The following variables are used:

e' = number of establishments that answer the monthly survey (with incidents 1 and 2 in the monthly questionnaire)

e'' = number of establishments that answer the monthly survey (with incidents 1 and 2 in the monthly questionnaire);

c' = number of establishments that are closed within their opening period (with incident 3 in the monthly questionnaire). Always less than or equal to c .

dm = number of days the establishment has been open during the reference month

VM = number of guests checked in during the whole month

NM = number of occupied bedplaces (overnight stays) during the whole month

BD = number of occupied double rooms

BI = number of occupied single rooms

A) Estimate of the number of establishments open in the month

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

B) Estimate of the number of bedplaces in the establishments open in the month

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

C) Estimate of number of guests checking in during the month

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

α_{jk} , the raising factor by vacancy for that of province j and category k

The monthly estimator broken down by place of residence

$$\widehat{VM}_{jkm} = \widehat{VM}_{jk} \cdot \frac{\widehat{V}_{jkm}}{\sum_m \widehat{V}_{jkm}}, \text{ where } m \text{ is the place of residence}$$

D) Estimate of occupied bedplaces (Overnight stays)

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

α_{jk} , the elevation factor by bed places for province j and category k

The monthly estimator broken down by place of residence

$$\widehat{NM}_{jkm} = \widehat{NM}_{jk} \cdot \frac{\widehat{N}_{jkm}}{\sum_m \widehat{N}_{jkm}}, \text{ where } m \text{ is the place of residence}$$

E) Estimate of occupied extra bedplaces or overnight stays in said extra bedplaces

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

F) Estimate of the average stay

a) Average stay by category

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

b) Average stay by country of residence

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

m = place of residence

G) Estimate of employed personnel

$$\widehat{T}_{jkm} = \left(\sum_{i=1}^{e_{jk}} T'_{ijkm} + \sum_{i=1}^{e''_{jk}} T'_{ijkm} \right) \cdot \varpi_{jk}$$

where: $T'_{ijkm} = T_{ijkm} \cdot AC_{ijk}$, with AC_{ijk} the percentage of employed personnel dedicated to the activity of accommodation¹ ω_{jk} , the elevation factor by bed places for province j and category k and professional situation m

H) Estimate of occupied rooms

Previously it is necessary to calculate the number of rooms existing in the open establishments:

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

a) Total of occupied rooms:

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

Note: B_{ijk} for establishments sending the questionnaire by XML, it is the result of adding double rooms occupied by two persons, double rooms occupied by one person and others.

b) Occupied double rooms:

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

δ_{jk} , the raising factor by vacancy for that of province j and category k ;

c) Occupied single rooms:

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

δ_{jk} , the raising factor by vacancy for that of province j and category k ;

I) Estimate of the occupancy rate

a) Net occupancy rate by bedplaces

¹ In the base of establishments sending the XML file, it is assumed that the percentage is equal to 100

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

For the set of categories

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

b) Net occupancy rate by rooms

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

and for the set of categories

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

8. Collection of the information

The enquiry pertaining to basic data refers to one full week of each month, randomly selected, in such a way that all the establishments cover the whole month between them.

A second enquiry is performed regarding the total number of guests checked in and overnight stays caused during the month the data refer to, in the strata of three, four and five gold stars.

Information is provided monthly by hotel establishments, via a questionnaire which is forwarded to the National Statistics Institute. Similarly, it is possible to submit information through a telematic transmission or by uploading an XML file, or through the Internet via the ARCE system, by completing the on-screen questionnaire directly.

9. Dissemination of the information

The information is presented on different levels of geographical breakdown: national, Autonomous Community, provincial, tourist areas and tourist sites.

All areas have been considered (group of municipalities) as well as the municipalities in which there is an abundance of tourists.

10. Statistical secrecy and significance

RULES FOR STATISTICAL SECRECY

Since 2019, Information may be provided about all those strata (or data geographical groups) where the number of establishments with an incidence of 1 (open with movement) is greater than or equal to 4.

In this way, if in a stratum (province-category), the condition of statistical secrecy is not fulfilled, establishments from that stratum will be aggregated to those in the category immediately below (except for 1 gold, which aggregates to 2 gold), until reaching the condition necessary in order to be able to offer data (gold categories will not be aggregated to silver categories)¹.

This stands with the exception of Ceuta and Melilla, about which information for two groups can be provided: silver stars and gold stars.

STATISTICAL SIGNIFICANCE RECOMMENDATIONS

Results of the main variables estimated by the surveys are statistically significant if, within the domain in which the information is provided, the following conditions are met, in general:

Since 2019, it has been established that there must be, 8 establishments at least with an incidence of 1 (with data) or 2 (without movement) must be in the monthly module for categories 5, 4 and 3 gold stars, and in the weekly module for the rest of the categories, in which at least 4 of them will have an incidence of 1 in the monthly module.

The sampling error must be lower than 20%, or where it is higher than 20%, the collection percentage must be higher than 80%.

For years prior to 2019, the applicable statistical significance recommendations stipulate that there must be at least 12 establishments in the sample with incidents 1 (with data) or 2 (no movement) in the monthly module for the categories of 5, 4 and 3 gold stars and in the weekly module for the rest of categories, of which at least 3 of them will have 1 incidence in the monthly module.

¹From January 2005, gold and silver categories can be joined.

11. Variation coefficients

Calculation of sampling errors of the monthly variables that include the number of travellers in a month and the number of occupied bed places, both in the case where the number of days that the establishment is open in the reference month, dm , is considered and in the case where this is not considered:

Number of travellers checked in in the month: \hat{VM}_{jk}

Number of bed places occupied (overnight) in the month: \hat{NM}_{jk}

Where \hat{Y}_{jkm} is the estimator of any of these variables, for each province, j , category, k , and modality (residents or non-residents), m .

The estimate of the relative sampling error (%) will be calculated for each month as follows:

\hat{Y}_{jkm} is the estimator of any of these variables for each province, j , category, k , and modality (residents or non-residents), m .

The estimation of the relative sampling error (%) will be computed for every month as follows:

$$CV\hat{\left(\hat{Y}\right)} = \frac{\sqrt{\hat{V}\left(\hat{Y}\right)}}{\hat{Y}} \cdot 100;$$

$$CV\hat{\left(\hat{Y}_j\right)} = \frac{\sqrt{\hat{V}\left(\hat{Y}_j\right)}}{\hat{Y}_j} \cdot 100$$

, for each province, j ;

$$CV\hat{\left(\hat{Y}_{jk}\right)} = \frac{\sqrt{\hat{V}\left(\hat{Y}_{jk}\right)}}{\hat{Y}_{jk}} \cdot 100$$

, for each province, j , and each category, k ;

$$CV\hat{\left(\hat{Y}_m\right)} = \frac{\sqrt{\hat{V}\left(\hat{Y}_m\right)}}{\hat{Y}_m} \cdot 100$$

, for each modality (residents or non-residents), m ;

where:

$$\hat{V}\left(\hat{Y}\right) = \sum_j \hat{V}\left(\hat{Y}_j\right) = \sum_j \sum_k \hat{V}\left(\hat{Y}_{jk}\right);$$

$$\hat{V}(\hat{Y}_j) = \sum_k \hat{V}(\hat{Y}_{jk});$$

$$\hat{V}(\hat{Y}_{jk}) = \sum_m \hat{V}(\hat{Y}_{jkm})$$

$$\hat{V}(\hat{Y}_m) = \sum_j \sum_k \hat{V}(\hat{Y}_{jkm})$$

where $\hat{V}(\hat{Y}_{jkm})$ will be calculated as follows depending on the type of estimator.

Group A estimators: Weekly information plus XML, without considering information on the number of days that the establishment is open in the reference month, dm

$$\hat{V}(\hat{Y}_{jkm}) = (1 - f_{jk}) \cdot \frac{e_{jk} + e''_{jk}}{(e_{jk} + e''_{jk}) - 1} \cdot \frac{\hat{P}_{jk}^2}{\left(\sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm}{D} \right)^2} \cdot \sum_s (Y_{ijkm} - \hat{R}_{jkm} P_{ijk})^2$$

where

$$f_{jk} = \frac{e_{jk} + e''_{jk} + c_{jk}}{E_{jk}} \quad \text{y} \quad s = e + e''$$

therefore:

$$Y_{ijkm} \quad \text{si } i \in e''$$

$$Y_{ijkm} = \left\langle Y_{ijkm} \cdot \frac{D}{7} \quad \text{si } i \in e \right.$$

$$\hat{R}_{jkm} = \frac{\sum_{i=1}^{e_{jk}} Y_{ijkm} \cdot \frac{D}{7} + \sum_{i=1}^{e''_{jk}} Y_{ijkm}}{\sum_{i=1}^{e_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk}}$$

Group B estimators: Monthly information plus XML (only 3, 4 and 5 star establishments), considering information on the number of days that the establishment is open in the reference month, dm

$$\hat{V} \left(\hat{Y}_{jkm} \right) = (1 - f'_{jk}) \cdot \frac{e'_{jk} + e''_{jk}}{(e'_{jk} + e''_{jk}) - 1} \cdot \frac{PM_{jk}^2}{\left(\sum_{i=1}^{e'_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} + \sum_{i=1}^{e''_{jk}} P_{ijk} \cdot \frac{dm_{ijk}}{D} \right)^2} \cdot \sum_s (Y_{ijkm} - \hat{R}_{jkm} P_{ijk})^2$$

where

$$f_{jk} = \frac{e'_{jk} + e''_{jk} + c'_{jk}}{E_{jk}} \quad \text{y} \quad s = e' + e'',$$

therefore:

$$Y_{ijkm} = \begin{cases} Y_{ijkm} & \text{si } i \in e'' \\ \langle YM_{ijkm} & \text{si } i \in e' \end{cases}$$

$$\hat{R}_{jkm} = \frac{\sum_{i=1}^{e'_{jk}} YM_{ijkm} + \sum_{i=1}^{e''_{jk}} Y_{ijkm}}{\sum_{i=1}^{e'_{jk}} P_{ijk} + \sum_{i=1}^{e''_{jk}} P_{ijk}}$$

12. Linking coefficients

Methodological changes or extraordinary updates of the directories constituting the survey framework imply that the data published introducing said changes is not directly comparable to the previously published data.

In order to avoid this break in the series, and allow for comparability, linking coefficients are calculated to be applied to the published data before introducing the improvements.

Linking coefficients

Given the strong seasonal nature of these series, we have decided to obtain different linking coefficients for each month, thus maintaining the interannual variation rates of the series, although the same does not apply to the inter-monthly rates.

For each month, the linking coefficient for a variable X is obtained as the quotient between the estimated value of said variable in said month of year T, considering all of the available information (methodological changes and/or improvements in the directory) and the estimated value in the same month without considering the innovations.

$$CX_{i,T} = \frac{X_{i,T}^{conmejoras}}{X_{i,T}^{sinmejoras}}$$

where:

$i = 1 \dots 12$ months

$X_{i,T}^{conmejoras}$ = Estimated value of variable X in month i of year T, using all of the information available.

$X_{i,T}^{sinmejoras}$ = Estimated value of variable X in month i of year T without considering the innovations.

Linked series

The series that are linked and comparable with the data published are calculated by multiplying the published series (prior to the date on which the improvements are introduced) by the linking coefficient of the corresponding month.

$$X_{i,T-j}^{enlazada} = X_{i,T-j}^{pub} \times CX_{i,T}$$

where:

$i = 1..12$ months

$X_{i,T-j}^{pub}$ = Published estimated value of variable X in month i of year $T - j$
(where $j > 0$)

$CX_{i,T}$ = Linking coefficient for variable X in month i

Given that the linking coefficients are calculated independently for each geographical breakdown, each category or each nationality, the linked series lose their additivity.

13. Treatment of partial non-response

TREATMENT OF THE EMPTY STRATA OF THE HOS

For provinces, tourist areas and tourist sites, by category:

- 1) If there are only refusals and closed establishments for a stratum, the estimation is ZERO, due to the fact that all refusals are considered as closed.
- 2) If there are only refusals, all refusals are considered as open (since we do not have any record of any closed) and said stratum has to be imputed as a figure in point 4.
- 3) If there are open establishments in an area or a site in a stratum, but any of them are in the sample, that stratum is imputed in line with the following point.
- 4) The imputation stratum is the first non-empty stratum that is found decreasing by category to 1 silver, and if not, it is found increasing to 5 gold. Once the estimators of the donor stratum have been computed, the imputation to the empty stratum is carried out, according to the bedplaces open in the directory of the empty stratum, as we have considered that there are not closed bedplaces.

For example:

Assuming stratum k of province j to be empty, and that it will be imputed by the data of stratum k':

- The estimation of the variables: number of establishment open in the month, number of bedplaces and number of rooms of the establishments open in the month, coincides with the item of data in the directory.
- For the rest of the variables:

$$\hat{V}_{jk} = \underbrace{\sum_{i=1}^{e_{jk'}} V_{ijk'} \cdot \frac{D}{7} \cdot \frac{\hat{P}_{jk}}{\sum_{i=1}^{e_{jk'}} P_{ijk'}}}_{\text{Stratum k' information.}} \cdot \underbrace{\left(\sum_{i=1}^{e_{jk'}} V_{ijk'} \cdot \frac{D}{7} \cdot \frac{1}{\sum_{i=1}^{e_{jk'}} P_{ijk'}} \right)}_{\text{Information.}} \cdot \hat{P}_{jk}$$

Although for an empty stratum the number of establishments is greater than or equal to 3, its estimation will be aggregated to the donor stratum.

(From December 2004 onward)

It is necessary that, due to the treatment of the empty strata in provinces, areas and sites, the totals, by category, are not distorted by the increase of the variable in the recipient category, and by the decrease of the donor category. Thus, it is necessary to compute the contribution of the empty stratum as a difference between the estimation of the recipient stratum and that of said stratum if it had not received the population of the empty stratum.

If there is more than one empty category from the same province or area, the difference will be distributed proportionally to the number of bedplaces in the directory of the empty strata (similar to the procedure carried out for Campsites, for national data by category).

In turn, if empty strata are found for provinces or the rest of the geographical breakdowns with the monthly questionnaire, the weekly estimations will be used in these strata.

(From July 2005 onward, for the province, area and site strata of the HOS)

inci = Incidence of the weekly questionnaire

inci1 = Incidence of the monthly module

In turn, the following cases will require a specific treatment, due to the fact that they have a monthly module:

1.- In a stratum where:

- (No. establishments with $inci1=1$) $\neq 0$ (there may be other closed establishments, 3, or refusals, 4)
- y (No. establishments with $inci=1$) $=0$ (there may be closed establishments, 3, or refusals, 4).

The monthly modules are used to compute the total guests and overnight stays in order to distribute them by nationality, and to compute the rest of the variables, it is necessary to look for a "donor stratum" of estimations obtained with the week. The donor stratum will be obtained within the same province, and it will be the first non-empty stratum, and with movement of guests (in the week information) that is found following the aforementioned order (in the HOS).

2.- In a stratum where:

- (No. establishments with $inci1=1$) $\neq 0$
- and (No. establishments with $inci=1$) $=0$
- and (No. establishments with $inci1=2$) $\neq 0$ (there are more establishments in the stratum but some of them have an incidence of 2)

The monthly modules are used to compute the total guests and overnight stays. In order to distribute them by nationality and to compute the rest of the variables, it is necessary to look for a "donor stratum" of estimations obtained with the week, but only to apply it to the establishment or establishments with a monthly incidence=1. The donor stratum will be obtained within the same province, and it will be the first non-empty stratum (in the week information) that is found following the aforementioned order (in the HOS).

3.- In a stratum where there are no establishments with $inc=1$ in the week and there is one or more with: $inci=2$ or 3 in the weekly questionnaire and $inci=1$ in the monthly questionnaire, the incidence in the weekly questionnaire will be reassigned as a *refusal*, 4. This will leave us in cases 1 or 2, then, and this case will be solved by either of those two ways. (This case will be checked first).