INSTITUTO NACIONAL DE ESTADISTICA

Hostel Occupancy Survey (YHOS)

Methodology

March 2015

Index

1	Presentation	3
2	Objectives	4
3	Statistical unit	5
4	Survey scope	6
5	Definition of variables	7
6	Survey framework and sample design	9
7	Estimators	10
8	Information collection	13
9	Coefficients of variation	14

National Statistics Institute

1 Presentation

This publication presents the results corresponding to the Hostel Occupancy Survey.

The data provided reflects the two aspects considered in the tourism study: from the demand perspective, information is offered on travellers, overnight stays and average stay, distributed by place of residence of the travellers, or by Autonomous City or Community of origin in the case of travellers resident in Spain; from the supply perspective, it provides the estimated number of open establishments, bedplaces, bedrooms, occupancy rates and information regarding employment in the sector.

Each month, this information is obtained on a national scope, and by Autonomous City or Community.

The National Statistics Institute would like to thank all of the professionals, businesspersons and institutions related to the tourist sector for the collaboration they have provided, which has been essential in conducting this survey.

National Statistics Institute

2 Objectives

The main objective of the Hostel Occupancy Survey is to ascertain the behaviour of a series of variables that allow us to describe the basic characteristics of this type of accommodation within the tourist sector, from both the supply and the demand points of view, thus answering the need of national organisations for knowledge regarding the sector, and meeting the requirements of international organisations.

National Statistics Institute

3 Statistical unit

These are all of the Hostels registered as such in the corresponding register of each Autonomous Community, according to the definitions appearing in the different Autonomous Community legal regulations regarding said accommodation.

Hostels are defined as establishments that offer accommodation services to the public, mainly in shared bedrooms, with or without complementary services, and usually with the possibility of participating in activities relating to the surroundings.

National Statistics Institute

4 Survey scope

This studies all of the Hostels that are part of the Spanish Youth Hostel Association (Red Española de Albergues Juveniles, REAJ, in Spanish) throughout the country.

This is with the exception of Ceuta, where there are no establishments of this type.

National Statistics Institute

5 Definition of variables

5.1 Estimated open establishments

The number of hostels open for the season as estimated by the survey.

A hostel open for the season is that for which the reference month is included within its opening period:

5.2 Estimated bedplaces

The number of bedplaces estimated by the survey in the hostels open for the season.

The bedplaces in a hostel are understood to be the number of fixed beds that the hostel has available; therefore, this does not include extra beds or double beds leading to two bedplaces.

5.3 Travellers

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

Travellers are classified by their place of residence. In the case of travellers resident in Spain, information is requested on the Autonomous City or Community of origin.

5.4 Overnight stay

An overnight stay is understood to be every night that a traveller stays in the hostel.

For example, a group of six persons staying at the establishment for two days gives rise to 12 overnight stays or occupied bedplaces.

5.5 Average stay

This variable is an approximation of the number of days that, on average, the travellers stay at the hostels, and it is calculated as the quotient between the overnight stays and the number of travellers.

5.6 Occupancy rate by bedplaces

Percentage-based relation between the total overnight stays and the product of the bedplaces and the days to which the overnight stays refer.

National Statistics Institute

It can be deduced from this definition that a hostel may have an occupancy rate lower than 100 percent, and yet, not have any available bedplaces (vacancies), since a bedroom or double bed may be occupied by one traveller, yielding a single overnight stay, though the bedroom has a greater capacity.

5.7 Weekend occupancy rate by bedplaces

Percentage-base relation between the total overnight stays and the product of the bedplaces and the two days to which the overnight stays refer (Friday and Saturday).

It can be deduced from this definition that a hostel may have an occupancy rate lower than 100 percent, and yet, not have any available bedplaces (vacancies), since a bedroom or double bed may be occupied by one traveller, yielding a single overnight stay, though the bedroom has a greater capacity.

5.8 Occupied personnel

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

National Statistics Institute

6 Survey framework and sample design

The survey has used the directories of the Autonomous Communities, and other auxiliary sources, such as the Spanish Youth Hostel Association (Red Española de Albergues Juveniles, REAJ, in Spanish), as the framework for selecting the informant units, and in which, the following data, among others, appears for each establishment: name, address, normal opening period, number of bedplaces, NIF.

These directories are updated on a permanent basis.

The survey is comprehensive in all provinces, except in Ceuta, where the survey is not conducted, due to the non-existence of this type of establishment there.

National Statistics Institute

7 Estimators

VARIABLES USED
E = number of establishments open during the month, existing in the directory
e = number of establishments that respond to the survey questionnaire using the customary method (incidence = 1 or 2)
c = number of establishments in the sample that state that they are closed in the questionnaire (incidence = 3)
dm = number of days that the establishment has been open during the reference month
D = number of days in the reference month (28, 29, 30 or 31)
P = number of bedplaces in open establishments, according to the directory
V = number of travellers checked in
N = number of occupied bedplaces or overnight stays
Dfs = total number of Fridays and Saturdays in the reference month
ES = average stay
T = employed personnel
H = number of bedrooms in open establishments, according to the directory
B = number of occupied bedrooms
GP = Occupancy rate by bedplaces.

SUBINDICES USED:
i = establishment
j = province
m = place of residence

PROCESSING OF INCIDENCES

In the event that during the data collection period it is not possible to collect information about an establishment, that information will continue to be requested during the two following months.

In the event that information about an establishment is not obtained during the collection period, it will be imputed to said establishment according to the data of the hostels of the stratum which have responded according to the formula shown below.

National Statistics Institute

Monthly information

Number of establishments open during the reference month:

$$\widehat{E}_j = E_j \frac{\left(\sum_{i=1}^{e_j} \frac{dm_{ij}}{D} \right)}{\left(e_j + c_j \right)}$$

Number of bedplaces in the establishments open during the reference month

$$\widehat{P}_j = \sum_{i=1}^{E_j} P_{ij} \cdot \frac{\left(\sum_{i=1}^{e_j} P_{ij} \cdot \frac{dm_{ij}}{D} \right)}{\left(\sum_{i=1}^{e_j} P_{ij} + \sum_{i=1}^{c_j} P_{ij} \right)}$$

Number of travellers checked in during the month

$$\widehat{V}_{jm} = \left[\sum_{i=1}^{e_j} V_{ijm} \right] \cdot \frac{\widehat{P}_j}{\sum_{i=1}^{e_j} P_{ij} \cdot \frac{dm_{ij}}{D}}$$

Number of occupied bedplaces (overnight stays)

$$\widehat{N}_{jm} = \left[\sum_{i=1}^{e_k} N_{ijm} \right] \cdot \frac{\widehat{P}_j}{\sum_{i=1}^{e_j} P_{ij} \cdot \frac{dm_{ij}}{D}}$$

Number of weekend overnight stays

$$\widehat{N}_j^{fs} = \left[\sum_{i=1}^{e_j} N_{ij}^{fs} \right] \cdot \frac{\widehat{P}_j}{\sum_{i=1}^{e_j} P_{ij} \cdot \frac{dm_{ij}}{D}}$$

Average stay

$$\widehat{ES}_j = \frac{\widehat{N}_j}{\widehat{V}_j}$$

Average stay by country of residence

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

National Statistics Institute

Employed personnel

$$\widehat{T}_j = \left(\sum_{i=1}^{e_j} T_{ij} \right) \cdot \frac{\widehat{P}_j}{\sum_{i=1}^{e_j} P_{ij} \cdot \frac{dm_{ij}}{D}}$$

Estimated bedrooms in open establishments

$$\widehat{H}_j = \sum_{i=1}^{E_j} H_{ij} \cdot \frac{\left(\sum_{i=1}^{e_j} H_{ij} \cdot \frac{dm_{ij}}{D} \right)}{\left(\sum_{i=1}^{e_j} H_{ij} + \sum_{i=1}^{c'_j} H_{ij} \right)}$$

Occupancy rate by bedplaces

$$\hat{GP}_j = \frac{\widehat{N}_j}{D \cdot \widehat{P}_j} \cdot 100$$

Weekend occupancy rate by bedplaces

$$\hat{GP}^{fs}{}_j = \frac{\widehat{N}_j^{fs}}{D^{fs} \cdot \widehat{P}_j} \cdot 100$$

National Statistics Institute

8 Information collection

The basic data on the hostels refers to one month.

The information is provided monthly by the hostels, via a questionnaire that is forwarded to the National Statistics Institute. Likewise, it is possible to submit the information online, through the IRIA system, directly completing the questionnaire on the screen.

National Statistics Institute

9. Coefficients of variation

Processing of incidences caused by the imputation

Since it is an exhaustive survey, sampling errors are not calculated.

But taking into account that, although it is exhaustive, a total response rate is not always possible, as there is information about certain hostels that cannot be obtained, the resulting estimates may contain an error due to the imputation.

The information about the hostels that did not responded is estimated or imputed by the average of the ones in the same stratum which did respond. In order to take into account this imputation, an approximation of the variability added by the imputation is calculated.

By using the average of the ones that did respond as imputation, the approximation of the calculation of this variability coincides with the usual formulation of the sampling error.

The justification for this is detailed below:

Estimation of the variance in the case of exhaustive stratum:

Being n=N the sample size of the exhaustive stratum and the size of the ones that responded=m.

From a model based approach, the variance of the imputed data can be calculated. In the event of imputation by the average of the ones that respond, it is obtained:

$$\hat{V}(\hat{Y}) = N^2 \hat{\sigma}^2 \frac{1-\lambda}{m}; \quad \hat{\sigma}^2 = \frac{1}{m-1}\sum_{k\in r}(y_k - \bar{y})^2; \lambda = tasa\ de\ respuesta = \frac{m}{N}$$

The result coincides with the usual formulation of the variance, from a design based approach, it is obtained:

$$\hat{V}(\hat{Y}) = N^2(1-f)\frac{\hat{\sigma}^2}{m}; \quad f = \frac{m}{N}$$

Thus, for this survey the errors derived from the imputation of the main monthly variables are calculated: travellers $\hat{V}_j$ and overnight stays $\hat{N}_j$, where:

a) travellers checked in (V in the questionnaire)

b) Occupied bedplaces or Overnight stays (N in the questionnaire)

Being $\hat{Y}_{ij}$ the estimator of any of these variables (see section 5.3). Shown below it is the calculation of the estimation of the relative sampling error or coefficient of variation (in percentage) for each month:

National Statistics Institute

-Coefficient of variation estimated for the total estimated of Y, at a national level and by province j:

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

-Coefficient of variation estimated for the total estimated of Y, by residents and non-residents:

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

where m = modality (residents or non-residents)

where:

$$\hat{V}\left(\hat{Y}\right)=\sum_j \hat{V}\left(\hat{Y}_j\right) \quad ; \quad \hat{V}\left(\hat{Y}_j\right)=\sum_m \hat{V}\left(\hat{Y}_{jm}\right) \quad ;$$

$$\hat{V}\left(\hat{Y}_m\right)=\sum_j \hat{V}\left(\hat{Y}_{jm}\right)$$

and $\hat{V}\left(\hat{Y}_{jm}\right)$ will be calculated as follows:

Estimators with monthly information. For Y = V (travellers) or N (overnight stays)

$$\hat{V}\left(\hat{Y}_{jm}\right)=(1-f_j)\cdot\frac{e_j}{e_j-1}\cdot\frac{\hat{P}_j^2}{\left(\sum_{i=1}^{e_j} P_{ij}\cdot\frac{dm}{D}\right)^2}\cdot\sum_s\left(Y_{ijm}-\hat{R}_{jm}\ P_{ij}\right)^2$$

where

- $f_j=\frac{e_j+c_j}{E_j}$

- s would correspond to the units that responded, thus s=e.

- $\hat{R}_{jm}=\frac{\sum_{i=1}^{e_j} Y_{ijm}}{\sum_{i=1}^{e_j} P_{ij}}$

National Statistics Institute