

Measurement of the hotel Average Daily Rate using Internet Distribution Systems

Ibai Roman^a,
Igor Ibarguren^a,
Jon Kepa Gerrikagoitia^a, and
Emilio Torres-Manzanera^b

^a CICtourGUNE, Cooperative Research Centre in Tourism, Spain
[ibairoman@tourgune.org; igoribarguren@tourgune.org;
jonkepagerrikagoitia@tourgune.org]

^b University of Oviedo, Spain
[torres@uniovi.es]

Abstract

Revenue management pricing practices demand just-in-time information. Traditional methodologies in the field of Tourism Statistics provide monthly data with a delay of months. In order to get a daily indicator of the hotel Average Daily Rate, a methodology based on Internet Distribution Systems is used. The experiment has collected daily rates of 8,000 Spanish hotels for over a year. Results show that this information is highly correlated with figures published by National Statistical Institute, with a coefficient of determination of 75%. This method provides a cheap and easy system to obtain daily hotel prices within a few hours.

Keywords: Hotel prices, Average Daily Rate, official statistics, Internet Distribution Systems, statistical methods.

1 Introduction

In the context of hotel performance measurement, there are four essential indicators to measure profitability: occupancy rate, revenue per available room, return on assets and return on equity (Chen, 2011, 2010). In hospitality and tourism organisations, where revenue management is widely practised, it is necessary to develop measurement instruments that are precise, accurate, and respond fast to changes.

Several national statistical institutes have designed statistical surveys to analyse the profitability in the hotel sector (Instituto Nacional de Estadística, 2009; Statistics New Zealand, 2011; Statistics Norway, 2012). In general, figures obtained by these Statistical Offices are based on monthly averages. There is usually a delay of several months after data was collected before it is published.

In order to develop a methodology to provide a just-in-time measurement of the current hotel prices, this research work presents a robust and efficient data collection process. It is based on the price information the main Internet Distribution Systems provide to consumers. The contribution of this paper focuses on testing the data obtained this way with figures published by the National Statistical Institute.

2 Methodology. Data collection and relationship

In this section a brief description of the methodologies carried out by Instituto Nacional de Estadística (2010) and by the authors is presented. The aim is to compare the daily data collected by the authors during one year concerning more than 8,000

Spanish hotels with the figures published by the Spanish National Statistical Institute (NSI) in the last year.

2.1 Spanish survey on monthly Average Daily Rate

The Instituto Nacional de Estadística (2010) has built indicators on the profitability of the hotel sector. The indicator named ADR (Average Daily Rate), along with the RevPAR (Revenue per Available Room), constitute an important source of information for hotel establishments, which enables them to evaluate their pricing policy or revenue management. The NSI collects, on a monthly basis, information relating to prices, which hotel businesspeople apply to their clients for the service in a double room with a bathroom, excluding taxes and any other service. This information is requested in the Hotel Occupancy Survey questionnaire (hereafter, HOS) (Instituto Nacional de Estadística, 2011).

Due to the changes that occurred in the sector, due, among other aspects, to the introduction of new technologies in sales processing, as of 2008 the question has been amended relating to prices, explicitly requesting the ADR for a double room with a bathroom, excluding taxes and any other service, and distinguishing between different types of customer and sales channels: traditional tour operators, traditional travel agencies, enterprises, groups, direct hiring on the hotel website or the hotel chain website, online tour operators and online travel agencies.

The survey comprises three different scopes: Time, population and geography. The HOS are monthly and the estimated data refers to the activity carried out by each hotel establishment during the month corresponding to the reference period. In the HOS, the data requested refers to seven consecutive days, such that establishments in the sample are distributed throughout the month, so that each day of the reference month is sufficiently represented. Concerning population under study for the Hotel Occupancy Survey, it comprises all establishments in their modalities of hotel, with the categories of five, four, three, two and one stars and guesthouses -whether or not these are residential inns and boarding houses of a single category and of three, two and one stars in accordance with the regulations of each Spanish autonomous community. Finally, the geographical scope of the survey encompasses the entire country.

2.2 Survey on daily hotel prices

Regarding to the methodology developed by the authors to collect daily hotel prices, it comprises a web crawler which uses screen scraping techniques to acquire prices and availability for twin bedded rooms on every available hotel for a given IDS (Internet Distribution System), geographical and time scope. For this experiment, the selected IDS was Booking.com, the geographical region is the same that the NSI's survey, Spain, and data was collected every day from July, 2011, until June, 2012.

The data extraction process is launched automatically every 24 hours. The web crawler creates specific URLs (Uniform Resource Locators) for a given region at a given date and processes the response to extract the prices for a twin bedded room on every hotel on the list obtained as response. If a hotel found on an earlier date is not found it is considered to be unavailable for that date. If there are different rates for the same hotel, the cheapest price is selected. Figure 1 details the technical process of collecting data.

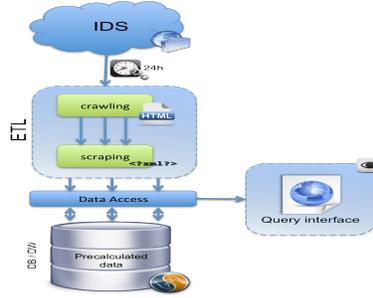


Fig. 1. Main phases for collecting and processing the data

Prices offered by the IDS include the Value Added Taxes, 8% during the period of observation.

2.3 Comparison of results

In order to check the accuracy of the method proposed by the authors, a relative measure of divergence was used to compare the published data by NSI and the data obtained by the authors:

$$\frac{IDS_m - NSI_m}{NSI_m} \cdot 100 \quad (1)$$

where m is the month, NSI_m is the ADR published by the NSI incremented by the value added tax, and IDS_m is the monthly average price obtained by the authors. The average price IDS_m follows a geometrical mean, in coherence with the index prices methodology of Instituto Nacional de Estadística (2009):

$$IDS_m = \left(\prod_{j,d,k} P_{m,jdkt} \right)^{1/n} \quad (2)$$

where $p_{m,jdkt}$ is the price of the j hotel in the k hotel category in the j -region for the d day of the month m . Additionally, a linear regression model has been calculated.

3 Results and discussion

The survey carried by the NSI covers a great number of establishments: in the winter season, it collects data from 13,000 hotels, whereas in the summer period, it goes up to 17,000 establishments. On the other hand, the number of hotels that offer their services through IDS is around 7,000 hotels (Fig. 2). In both surveys, the significance error level is less than 3% when disaggregated by regions ($p = q = 0.5$).

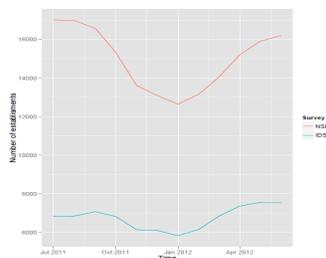


Fig. 2. Sample size of both surveys.

Figure 3 summarises the main results obtained by both methodologies. As it can be seen, prices obtained by IDS channels are cheaper than those published by NSI. The maximum divergence is observed in the off-peak season and during the peak tourist season prices tend to converge. The absolute difference in euros varies from €5 in July to €15 in January.

Date	Price (€)		Δ%
	IDS	NSI	
2011/07	76.67	81.32	-5.72
2011/08	79.98	88.24	-9.36
2011/09	69.18	75.82	-8.75
2011/10	64.66	72.25	-10.51
2011/11	60.83	74.52	-18.37
2011/12	60.87	74.84	-18.67
2012/01	58.75	73.76	-20.35
2012/02	59.43	74.09	-19.78

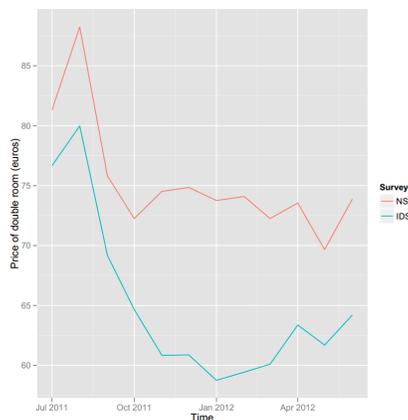


Fig. 3. Comparative of price evolution of both surveys.

In order to analyse this differences, we note in the first place that the price variables do not measure the same concept. On the one hand, the Average Daily Room Rate of NIS is the rate budgeted as the total rooms revenue divided by the number of rooms sold during the period:

$$\text{Average Room Rate} = \text{Rooms Revenue} / \text{Number of Rooms Sold}$$

The NIS average rate is calculated globally including different market segments: business, groups, tourists, airline crews, and other categories of guests served (Instituto Nacional de Estadística, 2010). On the other hand, the price calculated by IDS is the cheapest price on just one channel. And it is an offered product, not a sold product.

Therefore, the surveys have different purposes. The first one estimates the real ADR of the establishment, whereas the second one calculates the potential sale price.

However, there is a strong linear correlation between both series, the linear correlation is $\rho_{NIS,IDS} = 0.88$ and the linear regression models provides a R^2 of 75%:

$$NSI = 34.6317_{sd=6.9827} + IDS \cdot 0.6266_{sd=0.1069}, R^2_{adj} = 0.75 \quad (3)$$

This means that variation of the published data by National Statistical Institute is explained in a 75% by the data provided by the Internet Distribution Systems.

This fact has important consequences. The first one is that collecting data from IDS channels may lead to an estimation of the real ADR in the hotel sector, using the previous linear relationship. The second one is that it is necessary to balance among precision, publication time and economic cost. And the third consequence is that this method provides daily prices, improving the monthly data published by official institutes. It is clear that figures provided by National Statistical Institutes are precise and accurate. The proposed methodology, that collects data from Internet Distribution Systems, will never be as precise or accurate. But it is cheaper and easy to implement.

And these characteristics undoubtedly beat traditional methodologies in the area of revenue management. The use of IDS provides daily information on hotel prices. And hotel managers demand daily data instead of aggregated monthly data.

4 Conclusions

Revenue management needs just-in-time information on daily hotel prices. National statistical institutes publish monthly time series with a delay of months. Therefore, methodologies for precise experimental measurements, comparative analysis and collecting data about hotel rates is a topic of interest in the field of Tourism Statistics.

Web crawling of Internet Distribution Systems provides an excellent opportunity to test new statistical methodologies. This paper shows that there is a strong relationship between the results obtained by traditional methodologies and the ones collected by IDS methods. The linear model obtained a coefficient of determination of 75%. And this method provides daily data within a few hours.

References

- Chen, M.-H. (2010). The economy, tourism growth and corporate performance in the taiwanese hotel industry. *Tourism Management*, 31(5):665 – 675.
- Chen, M.-H. (2011). The response of hotel performance to international tourism development and crisis events. *International Journal of Hospitality Management*, 30(1):200 – 212.
- Instituto Nacional de Estadística (2009). *Hotel Price Index (HPI). Base 2008 (as of January 2009). Methodological note*. Instituto Nacional de Estadística, Madrid.
- Instituto Nacional de Estadística (2010). *Indicadores de Rentabilidad Hotelero. Metodología*. Instituto Nacional de Estadística. Madrid.
- Instituto Nacional de Estadística (2011). *Encuesta de ocupación en alojamientos turísticos*. Madrid
- Statistics New Zealand (2011). *Accommodation services and package holidays in the CPI*. Statistics New Zealand. .
- Statistics Norway (2012). *08100: Svalbard. Hotels and similar establishments. Utilisation and price per room*. Statistics Norway, Oslo. http://www.ssb.no/overnatting_en/