

Advance Estimator of the Resident Tourism Survey



Madrid, May 2026

Contents

1. Identification of the Operation	3
2. Origin of the Demand and Identifiable Need	3
3. Project Objectives.....	4
4. Type of Operation	4
5. Content.....	5
6. Project Characteristics	5
Dissemination Plan and Frequency	7
8. Implementation Schedule	8
9. Cost Estimate	8

1. Identification of the operation.

The following information will be indicated:

- PEN code and name of the statistics operation submitted for review.
9622: Resident Tourism Survey (ETR/FAMILITUR)
- Unit responsible for the operation:
General Subdirectorate of Statistics of Economic Sectors of the INE [Instituto Nacional de Estadística (National Statistics Institute)]
- Collaborating units/bodies, if applicable.
- Background of the operation:
Resident Tourism Survey, conducted at the INE since February 2015.
- Date of the last review, if applicable. 17-
03-2015
- Purpose of the review request, outlining the reasons leading the unit to request a new review

Development of a procedure to calculate an advance estimate of the total number of tourist trips made by residents in Spain, providing more timely information and with less time lag compared to the publication of the ETR, integrating partial information collected from the ETR and mobile phone positioning data. Applying this methodology will provide an indicator 35 days after the reference month.

2. Origin of the Demand and Identifiable Need.

The importance of tourism in Spain is undeniable. Spain ranks second in the world in both international tourist arrivals and international tourism revenue, according to UN Tourism data¹. On the other hand, tourism represents 12.6% of GDP².

Currently, the only official statistics that provide data on trips made by residents in Spain do so with a delay of up to five months compared to the reference month, in the case of the first month of each quarter. However, the need for information is constantly growing, and users increasingly require more timeliness in the publication of the data.

On the other hand, since 2020 the INE has been working on the project to measure domestic and outbound tourism based on mobile phone positioning. The collaboration between the INE (National Institute of Statistics) and mobile phone carriers has been ongoing, with the aim of developing a series of algorithms that allow carriers to transform the positioning data recorded by phone antennas into data on tourists, trips, overnight stays, etc., following the statistical definitions in force in this field.

¹ [UN Tourism Barometer](#)

² [Tourism Satellite Account 2024](#)

This work has made it possible to publish experimental statistics³ that provide monthly results on the number of tourists residing in Spain who travel within the national territory, broken down by municipality of origin and destination, and trips abroad, broken down by municipality of origin and country of destination.

These years of working with this information source of information and the experience acquired has determined that the data provided by mobile telephony can be very useful to calculate an advance indicator of the number of trips in the Resident Tourism Survey, thus gaining promptness with respect to the official survey.

The main rationale for this project lies in improving the timeliness of information, significantly reducing the existing time lag in the publication of the ETR. This improvement allows for a monthly leading indicator to be available approximately thirty-five days after the reference month, compared to the delay of up to five months in the publication of the ETR results for the first month of each quarter.

3. Project Objectives

The main objective of the project is to provide an advance estimate of the total number of tourist trips made by residents in Spain, apportioning this variable in a more timely manner and with less of a time lag compared to the publication of the RTS. The aim is to significantly reduce the time it takes for information to be available, going from several months to publish the final results—with a delay that can reach up to five months in the case of the first month of each quarter—to a period of thirty-five days after the reference month.

Furthermore, the advance estimate would be published with the same time lag as other tourism surveys, such as FRONTUR-EGATUR or the Surveys of Occupancy in Tourist Accommodation.

For illustrative purposes, the following comparison of information availability is provided:

Month of travel	Publication of ETR	Diff. ref. month	ETR preview	Diff. ref. month
January	June	5	Early March	35 days
February	June	4	Early April	35 days
March	June	3	Early May	35 days
April	September	5	Early June	35 days
May	September	4	Early July	35 days
June	September	3	Early August	35 days
July	December	5	Early September	35 days
August	December	4	Early October	35 days
September	December	3	Early November	35 days
October	March	5	Early December	35 days
November	March	4	Early January	35 days
December	March	3	Early February	35 days

4. Type of operation

This is a current operation that will measure levels.

³ [Experimental statistics](#)

5. Content

5.1 POPULATION UNDER STUDY

The target population for this study is the population aged 15 years or older residing in main family homes throughout the Spanish territory. However, the population under study includes children under 15 years of age, for whom indirect information will be collected regarding the trips made by this subpopulation.

5.2 GEOGRAPHICAL OR TERRITORIAL SCOPE

The entire national territory.

5.3 DATA REFERENCE PERIOD(S)

Current statistical operation on a monthly basis.

5.4 STUDY AND CLASSIFICATION VARIABLES

STUDY VARIABLES

Number of trips made by residents in Spain.

5.5 BASE STATISTICS

- Resident Tourism Statistics (ETR)
- Experimental statistics: measuring domestic and outbound tourism using mobile phones.

6. Characteristics of the Project

6.1 OBSERVATION UNITS TO WHICH THE PRIMARY DATA REFER

The units of analysis are the trips made by residents in Spain.

6.2 DATA COLLECTION METHODOLOGY

The ETR collects information on overnight trips made during the two calendar months prior to the interview.

This design implies a time lag with respect to the reference month in the publication of the final results, derived both from the collection period itself (the trips of a given month are not fully collected until two months have elapsed) and from the quarterly periodicity of the publication of the survey. Specifically, the trips made in month m are collected in month $m+1$ and month $m+2$. Furthermore, if month m coincides with the first month of the quarter, the final results are not published until $m+5$, that is, two months after the end of the quarter.

To produce this advance estimate, a partial sub-sample is used, consisting of interviews conducted during the first 25 days of the collection month, from which only the trips corresponding to the immediately preceding month are selected. In this way, it is not necessary to wait to complete the information collection for the two months contemplated by the survey design, but information on the trips made in the reference month is available before the end of the following month.

The mobile phone data captures the trips made during the month of study.

6.3 SAMPLE DESIGN

Sampling framework: Georeferenced Address Framework

Sample design: stratified three-stage sampling, where the first-stage units are census sections, the second-stage units are dwellings, and a person aged 15 or over has been selected within each dwelling.

Sample size: subsample with data from the first 25 days of ETR collection (around 45% of the final sample for the month).

6.4 MEANS OF DATA COLLECTION AND PROCESSING

The data processing aims to obtain an advance estimate of the total number of trips made by residents in Spain during the reference period. To do this, data collected in the 25-day subsample of the ETR are combined with auxiliary information from mobile phone data and various calendar variables.

The information from the subsample is raised to the total population by applying raising factors, calculated specifically for that subsample from the survey's sample design. The subsample constitutes a fundamental explanatory variable in prediction models, as it provides an early signal of travel behaviour in the reference month.

However, since it is a partial subsample limited to the first 25 days of collection, the direct estimate shows greater variability and may not correctly reflect the pattern of trips in the reference month. For this reason, the subsample is not used in isolation, but is complemented with the aforementioned auxiliary information to correct for potential biases and improve the accuracy of the advance estimate.

The methodology is based on the combination of different prediction models, including time series models with exogenous variables, such as ARIMAX, neural network models applied to time series, as well as other machine learning models that can capture non-linear relationships between variables. These models use the total number of trips in the ETR as the dependent variable and the information from the subsample, mobile phone data and calendar variables as explanatory variables, allowing for the capture of both the temporal dynamics of the series and the relationships between the different information sources.

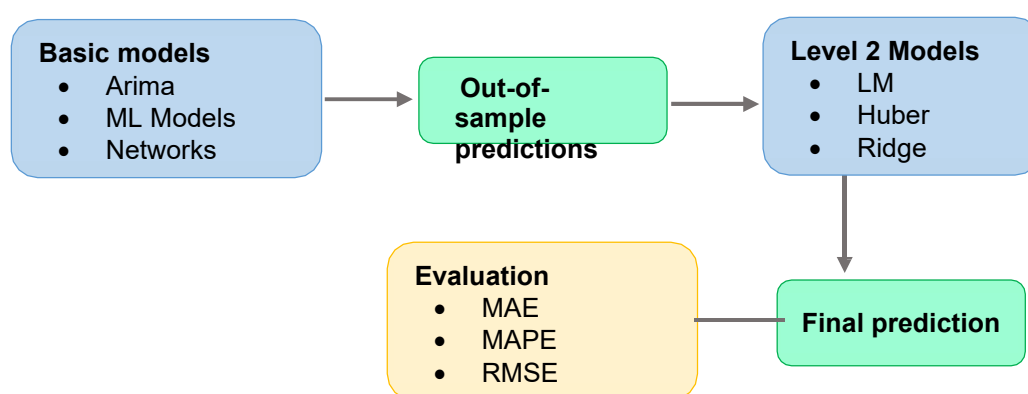
Each model generates its own estimate of the total number of trips. Subsequently, these estimates are combined using a second-level model (stacking), in which different specifications are considered, such as linear models, Huber regression or Ridge regression, with the aim of obtaining a final prediction that improves upon the individual estimates. This model learns a weighted combination of individual predictions, assigning greater weight to those models that perform better depending on the context.

The combination model takes into account both the overall performance of the models and their ability to capture different components of the series, such as

trends, seasonality, or possible changes in dynamics. This way, the stacking integrates these contributions and takes advantage of the complementarity between models.

To ensure the consistency of the process, both the base models and the second-level model are trained and evaluated using temporal validation procedures based on moving windows, which can simulate a real prediction context and evaluate the performance of the models in terms of indicators such as mean absolute error (MAE), root mean squared error (RMSE), or mean absolute percentage error (MAPE).

In this moving window approach, models are run iteratively using subsets of historical data that are shifted in time, so that in each iteration the model is trained with the information available up to a certain point and its predictive capacity is evaluated on later periods. This procedure allows a context for real-world prediction to be reproduced, in which only the information available up to that moment is available at any given time, and allows for the evaluation of the behaviour of the models in different periods, thus providing a more robust measure of their predictive capacity.



6.5 DATA COLLECTION FREQUENCY

Monthly.

7. Dissemination plan and frequency

7.1 DISSEMINATION PLAN

The advance estimate of the number of trips made by the population residing in Spain and the annual rate of change will be published with a delay of 33-35 days after the end of the reference month.

7.2 DISSEMINATION FREQUENCY

Monthly

8. Implementation schedule

The methodology regarding the implementation schedule is currently being developed, and the design and testing of different mathematical models are being finalised.

The plan is to make the first publication as experimental statistics in September 2026, with data from July 2026.

9. Cost estimate

The Advance Estimator of the Resident Tourism Survey presented in this document is calculated using information already available, both an ETR subsample and mobile phone positioning information, and therefore does not involve any additional cost.

For informational purposes, the cost of both operations is included (in thousands of euros):

- ETR budget allocation (according to the Annual Program 2026): 3,538.67
- Economic compensation to mobile phone carriers: 154.10

The financial compensation set in the agreements between the INE and Telefónica and MasOrange amounts to 231,150 euros annually. These agreements include information on domestic, outbound and inbound tourism, so it is considered that only two-thirds of the amount would be attributable to this statistical operation.