

<1>Data Collection Report

ICT Panel – 6th edition

<2>Introduction

The Brazilian Internet Steering Committee (CGI.br), through the Regional Center for Studies on the Development of the Information Society (Cetic.br), a department of the Brazilian Network Information Center (NIC.br), presents the evaluation of data collection and processes associated with the results of the ICT Panel survey – 6th Edition: Web survey on e-waste.

<2>Data collection instruments

<3>Information about the data collection instruments

Data was collected using a structured web questionnaire containing closed-ended questions with predefined answers (single or multiple answers) and one open-ended question. It was self-administered, without the mediation of an interviewer.

<3>Themes

The survey investigated activities carried out online and the devices used to access the Internet, using as a reference the indicators validated by the ICT Households survey, carried out by CGI.br, and usage indicators related to the following themes:

- Activities carried out online
- Ownership of electronic devices
- Disposal of electronic devices

<2>Field data collection

<3>Data collection method

The computer-assisted web interviewing [CAWI] collection method was employed, using a programmed and self-administered online questionnaire.

<3>Data collection period

Data was collected between February 1 and 14, 2022.

<3>Collected sample

A total of 2,515 respondents, 16 years old or older, answered the survey out of a total of approximately 16,000 respondents contacted.

<2>Data processing

<3> Weighting procedures

For the ICT Panel - 6th Edition, the ICT Households 2022 survey was used as the primary reference, without the need for additional updates for new totals drawn from the Continuous National Household Sample Survey (Continuous PNAD).

<4>Stage 1 - Estimation of the number of Internet users represented in the ICT Panel - 6th Edition

Step I. Adjustment of the score model of propensity to be an Internet user for ICT Households 2022

Table 1
Model adjustment statistics

Independent variables in the model	ICT Households 2022	
	R ²	Correct classification rate ⁽¹⁾
Sex, age, level of education, social class, computer user indicator	0.34	81%

(1) = Percentage of individuals correctly classified based on the adjusted model.

Step II. Estimation of propensity scores for respondents to the ICT Panel - 6th Edition

Based on the model adjusted with data from the ICT Households 2022 survey, the propensity scores were estimated for the set of respondents in the ICT Panel - 6th Edition. Next, the distributions of the propensity scores in the ICT Households 2022 sample were compared with the scores of the ICT Panel - 6th Edition sample for Internet users. The results are shown in Table 2. The distribution of scores for respondents to the ICT Panel - 6th Edition has a different profile from that observed for the population of Internet users 16 years old or older according to ICT Households 2022.

Table 2
Comparison of the distribution of Internet user propensity scores

Survey	Minimum	Q1	Median	Mean	Q3	Maximum
ICT Households 2022	0.0186	0.6582	0.9026	0.7933	0.9864	0.9976
ICT Panel – 6th Edition	0.2976	0.9812	0.9928	0.9615	0.9962	0.9994

Step III. Determination of the common support population for the ICT Households 2022 survey and the ICT Panel - 6th Edition

Given that the distributions of the scores obtained in the two surveys are different, cut-outs were tested on the survey's target population in order to establish the set of Internet users 16 years old and older that were represented by the respondents of the ICT Panel - 6th Edition. The results are shown in Table 3.

Table 3
Comparison of the distribution of weights of the respondents to the ICT Panel - 6th Edition, according to alternatives for cutting out the propensity scores

Calibrated weight statistics	Minimum	Q1	Median	Mean	Q3	Maximum
Without score cut-outs	0.2948	0.9541	0.9783	0.9498	0.9985	0.9979
Cut-out for scores greater than or equal to 2/3	0.6684	0.9562	0.9785	0.9556	0.9885	0.9979
Cut-out for scores greater than or equal to 3/4	0.7504	0.9584	0.9793	0.96	0.9887	0.9979
Cut-out for scores greater than or equal to 4/5	0.8006	0.9608	0.9801	0.9653	0.9889	0.9979

Considering the results of the comparison of weights of the respondents to the ICT Panel - 6th edition, it was decided to use the cut-out of propensity scores greater than or equal to 2/3 as the identification of the population of Internet users 16 years old or older.

With this cut-out, the represented population corresponds to just over 113 million (approximately 85%) of users 16 years old or older. It should be noted that the population represented - with the highest propensity scores to be Internet users - consists of a younger, more educated population in higher social classes.

<4>Stage 2 - Estimation of pseudo-probabilities of inclusion to determine the weights of the respondents to the ICT Panel - 6th Edition

The process of estimating pseudo-weights consists of estimating pseudo-probabilities of inclusion of respondents from the ICT Panel (non-probability sample) in the ICT Households survey (probability sample) based on a model; and using their reciprocals as weights, just as in a traditional probability sampling survey.¹

The most parsimonious model considering the independent variables (**X**) available and common to both databases contains the following variables: macro-region, social class, indicator of Internet use on computers, level of education, age group, sex, and number of residents in the household. Based on this model, the pseudo-probabilities of inclusion of respondents from the ICT Panel - 6th Edition into the ICT Households 2022 survey were estimated. The reciprocals of these pseudo-probabilities are the initial weights allocated to each respondent from the ICT Panel - 6th Edition.

These initial weights were calibrated to the estimated marginal totals of the variables macro-region, sex, age group, level of education, and Internet activity practices. The weights thus calibrated were considered for the estimation of all the outcome indicators of interest and the associated measures of precision.

<3>Estimation of errors

Replication was the method used to estimate errors.¹

¹ More details are available in the ICT Panel's "Methodological Report".