MEASURING E-COMMERCE IN BRAZIL

ICT ENTERPRISES SURVEY AND WEBSCRAPPING

Working group on measuring e-commerce and the digital economy | 29 November 2022
The Regional Center for Studies on the Development of the Information Society (Cetic.br) has as its mission the production of statistical data and analyses on the impacts of digital technologies on society.

1. ICT statistical data production center & knowledge center
   - ICT public statistics
   - Disaggregated data
   - Statistics for SDGs
   - Qualitative studies

2. Capacity-building on survey methodology for the production and use of ICT statistics

3. Laboratory of ideas and methodological innovation

4. Analysis of socioeconomic impacts of ICT & contributions to policies
REGULAR PRODUCTION OF STATISTICAL DATA & SECTORAL STUDIES

CGI.BR ICT SURVEYS
ICT ENTERPRISES 2021

METHODOLOGY

TARGET POPULATION
All active Brazilian enterprises with 10 or more employed persons

COVERAGE
National

DATA COLLECTION
Computer-assisted telephone interview (CATI)

DATA COLLECTION PERIOD
August 2021 to April 2022

SAMPLE
4,064 enterprises

OBJECTIVE
Measure the access and use of information and communication technologies (ICT) in Brazilian enterprises

MÓDULOS
General information on ICT systems
Internet use
Security
E-commerce
Skills
Software
New technologies
Privacy and personal data protection

E-COMMERCE
Increase in ecommerce practices

Total number of Internet users (%)

<table>
<thead>
<tr>
<th>Year</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>34%</td>
<td>39%</td>
<td>46%</td>
<td></td>
</tr>
</tbody>
</table>

Source: ICT Households

Types of products most bought online in the pandemic (2021)

Internet users 16 Years or older who purchased goods and services online (%)

- Ordered meals on site or applications: 44% (▲ 29 pp)
- Paid for streaming services: 43% (▲ 9 pp)

Online purchase channel (2020)

Internet users 16 Years or older who purchased goods and services online (%)

- Instant messages: 46% (▲ 20 pp)
- Social media: 28% (▲ 20 pp)

Payment method (2021)

- Credit card: 76%
- Pix: 72%
ENTERPRISES THAT SOLD ON THE INTERNET IN THE LAST 12 MONTHS

Total number of enterprises with Internet access (%)

In 2019, 57% of enterprises sold on the Internet

2019:
- Small enterprises: 57%
- Accommodation and food services: 59%

ENTERPRISES THAT SOLD ON THE INTERNET IN THE PAST 12 MONTHS BY TYPE OF ONLINE MEDIA USED FOR TRANSACTIONS

Total number of enterprises that sold on the Internet (%)

MEASURING ECOMMERCE
WEBSCRAPING
PILOT PROJECT WITH UN BIG DATA REGIONAL HUB OF BRAZIL
USE OF TRADITIONAL SURVEY DATABASES IN CONJUNCTION WITH ADMINISTRATIVE RECORDS AND DATA COLLECTION ON THE WEB

Brazil has the 5° largest domain name registrations among OECD and G20 countries

Domains registered under individual taxpayer registry and company registration number

The main objective is to produce e-commerce indicators based on web scraping of websites (use of e-commerce related words to assess whether the company sell online).
USE OF TRADITIONAL SURVEY DATABASES IN CONJUNCTION WITH ADMINISTRATIVE RECORDS AND DATA COLLECTION ON THE WEB

As a pilot study we chose to web scrap only companies’ proprietary websites

Cetic.br identified DNS 1,930

1,299 key company identifier has at least one DNS

302 key company identifier have two DNS

122 key company identifier have three DNS

One key company identifier have 370 DNS

Sample of 5,252 companies from the brazilian official register

For 1,930 companies, 5,257 .com.br domain names were found

For 3,322 companies no websites in .com.br were found

MAIN FINDINGS AND FUTURE WORK

- It is not possible to locate the website used to sell online based only on the company identifier.

- The use of searching engines to look for websites, based on company name, address, phone number, does not reach good results (for the 3,322 enterprises not found in .com.br register, were found more than 95,000 sites with at least one of the information).

- It is possible to web scrap and estimate the e-commerce websites for the population of .COM.BR domains based on a probabilistic sample of the .br domain register.

- But how to connect webscrapping data and official NSO registers remains a challenge - Need to developed more matching methods.

- Websites may be just one way for selling online – how to deal with companies that uses several online channels for selling online?

THANK YOU!

leonardomelo@nic.br

ww.cetic.br