



OTTAWA 2023

64TH WORLD STATISTICS CONGRESS



WELCOME.



OTTAWA 2023

64TH WORLD STATISTICS CONGRESS



CPS 838
Strengthening Modernization and Innovation
– a holistic approach

Mrs. Sofia Rodrigues
Mr. Francisco Lima
Mr. Jorge Magalhães
Mr. Paulo Saraiva
Mr. Almiro Moreira
Statistics Portugal

Tuesday 18 July, 8:30 AM - 9:40 AM

Introduction

Rapidly changing ecosystem

Imperative to embrace new technologies and methodologies to stay relevant and provide accurate and timely information

Streamline processes - do more with less

Statistics Portugal's initiatives in the new Ecosystem

National Data Infrastructure

Single point access to several types of data; make it available to multiple purposes and users

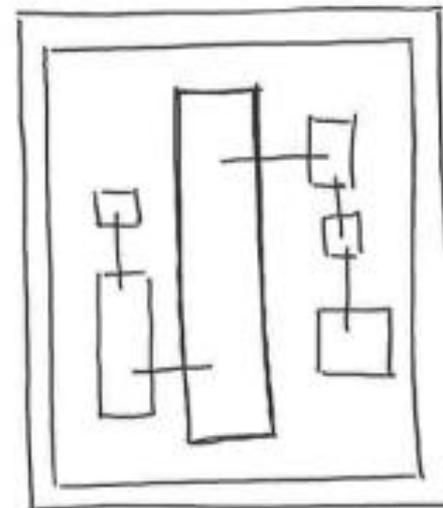
Main goals

Use of multiple data sources

Increase the analytical capacity and flexibility in data management

Exploitation of new data sources for statistical purposes

Innovation



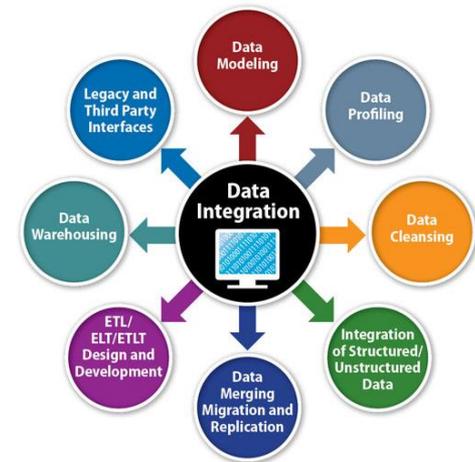
Organizational Initiatives

In 2019, adjustments in the internal organization for **strengthen** the capacity for data management and analysis Methodology and Information System Management and Data Collection Departments

These changes in competences were formalized in Statistics Portugal Statutory Laws

The Statistical Production Process Handbook was updated in 2020 to introduce National Data Infrastructure and put a greater emphasis on appropriation, validation, treatment, analysis and integration of sources, other than surveys

In 2020, creation of a dedicated unit: Administrative Data Unit, under the Management and Data Collection Department



Technological and Methodological Initiatives

**More intensive
use of Artificial
Intelligence and
Machine Learning**

**Administrative
and other Data
are currently
evaluated and
processed
centrally**

**Data protection
methods**

**Access control and
database auditing**

Dissemination and Communication

For internal use:

Consolidated and enriched databases (one database for all users)
Specific tailored tables

Methodological note

Flash report /
Dashboard
Service Level Agreement
Execution Report

For data providers:

Own files enriched with characterization variables

Methodological note

Dashboards
Feedback on the use of admin data

For scientific community:

Fully anonymized micro data files
Ready-made tables
Safe center

Methodological note

Characterization reports

For society:

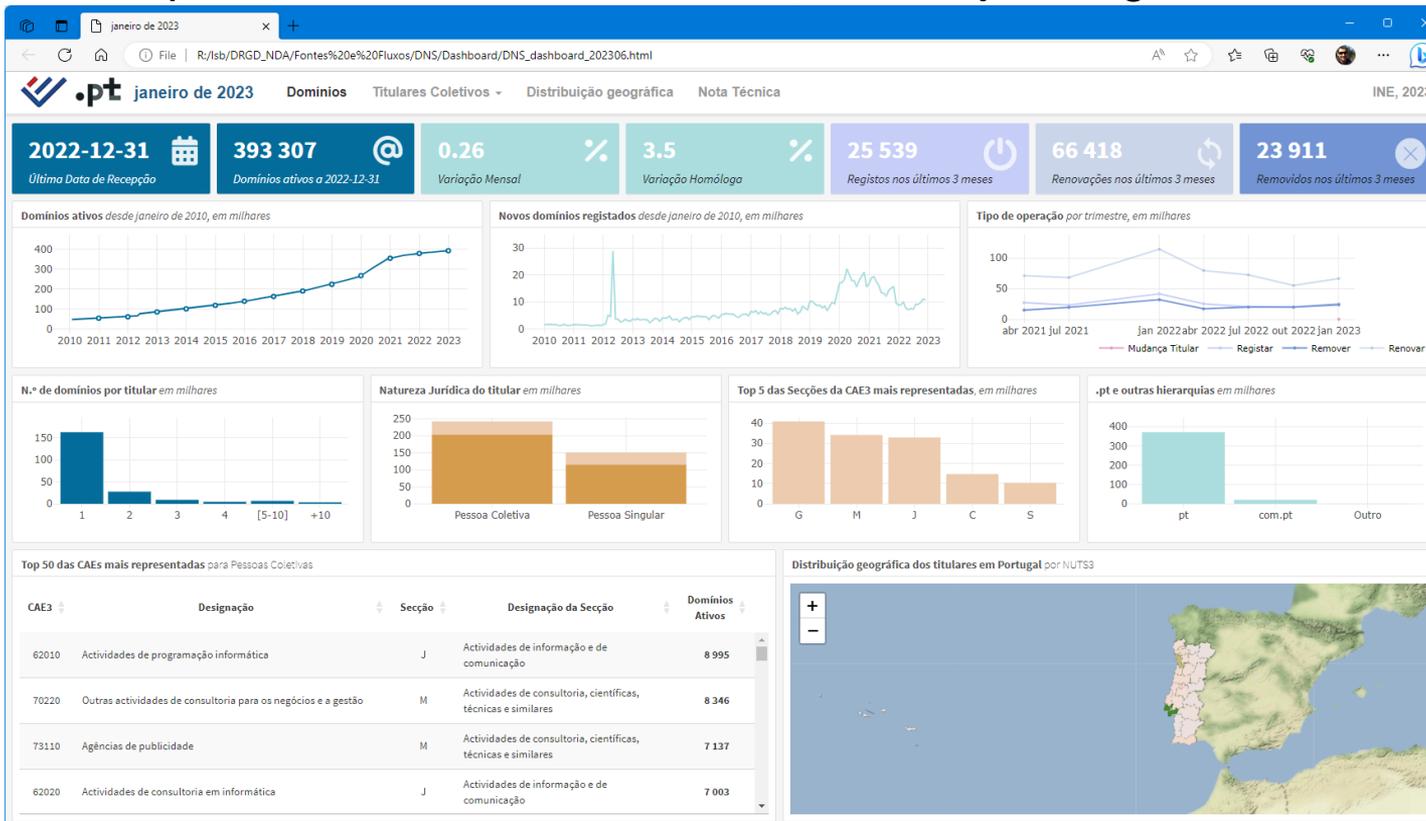
Statistical information
Statistics in development

Methodological note

Press releases
Media communications

Dissemination and Communication

Example of a dashboard: interactive, with major insights



Specific recruitment has been carried out for this unit, where scientific skills in terms of data appropriation, data analysis, programming skills and machine learning techniques are highlighted.

Up to now, this unit has **5 dedicated FTE**, and 1 research holder.

In addition:

- Resident researchers (hiring service providers) specialized in the processing and analysis of large datasets;
- Specialized consulting contracts with academia;
- Specific training in programming languages for data analysis
- “half-an-hourR” weekly meetings: stimulating the maintenance of a "community of R statisticians", at SP

E-Invoice: the proof of concept

Protocol with Tax Authority, Statistics Portugal receives
≈ 80 million records a month

A set of ETL procedures are performed, introducing quite significant improvements. Additionally, the database is also treated and enriched with:

- Inclusion of new attributes from other sources;
- Temporal referencing of characterization variables;
- Identification and imputation of anomalous values (systematic approach);
- Creation of new aggregated tables;
- Implementation of alternative plans (ongoing).

One of the main goals is to replace STS turnover variables with this dataset. In 2024, e-Invoice data will replace turnover for the Services sector.

Holistic approach: Each component of the system contributes for this achievement.



Consolidated database and Flash report delivered internally within 30 hours

StatsLab – latest press releases

One of the latest releases, in May 2023:

[Birth of the first child: what changes in the pattern of family expenses?](#)

Other examples:

[Gross monthly earnings per worker \(first published in oct 2019, became official statistics since nov 2021\)](#)

[Census with administrative data – progress report](#)

[Monthly Business Demography](#)

[Regional Economic Activity e-Invoice 2019-2020](#)

[Income statistics at the local level – Reported income on Personal Income Tax Indicators](#)

[Context indicators for the COVID-19 pandemic in Portugal – Data for Good](#)

Figure 1. Effects on Household Monthly Expenditure: Total Expenditure



Source: Statistics Portugal, based on the information from Tax Authority.

Conclusions

- Several initiatives implemented in a holistic approach, driven by the need to keep up with the demands of the evolving ecosystem
- Results achieved so far are very positive
- Innovation and efficiency (and user engagement) are key
- Address the challenges: sustainability; maintain quality and integrity
- Ongoing investment in technology and staff training and development



OTTAWA 2023

64TH WORLD STATISTICS CONGRESS



THANK YOU.