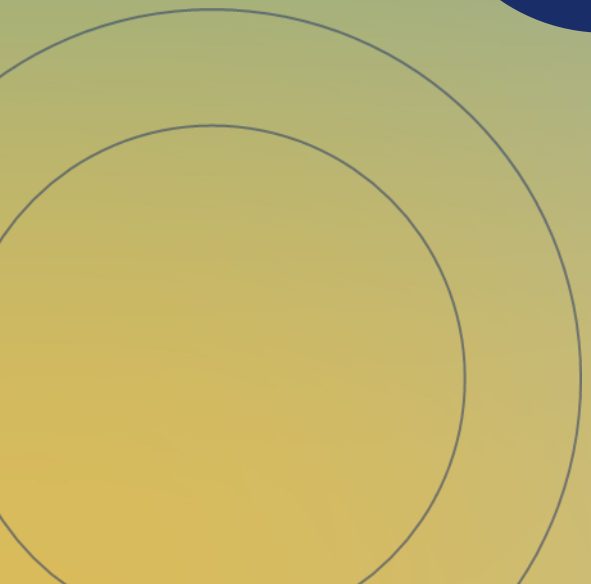
EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL



Instituto Nacional de Estatística
Statistics Portugal

eurostat

The conference is partly
financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

I can see clearly now outliers are gone

How to improve data quality of official statistics?

Sónia Mota, *scmota@bportugal.pt*, Banco de Portugal
Susana Santos, *smsantos@bportugal.pt*, Banco de Portugal
Beatriz Amorim, *bamorim@bportugal.pt*, Banco de Portugal

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Agenda

1. Focus on high quality standards
2. How can we maintain statistical quality within the ongoing flux of information challenges?
3. How can we see clearly now – a fresh approach to quality control?
4. Conclusion

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

The conference is partly financed by the European Union

"The global financial crisis highlighted the necessity of collecting relevant and more **granular data**. As a result, the scope of central bank statistics is moving beyond aggregates (...) to also cover micro (more detailed) data, in order to identify warning signs in the financial system at an early stage" (Source: ECB)

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

How can we maintain statistical quality within the ongoing flux of information challenges?

The conference is partly financed by the European Union

How can we maintain statistical quality within the ongoing flux of information challenges?

As statistical compilation relies further on highly **granular data**, and **tighter deadlines**, it is crucial to ensure that **speed** does not compromise **data quality**, and that **decision-making** rests on **high-quality data**

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

How can we maintain statistical quality within the ongoing flux of information challenges?

European Statistics Code of Practice

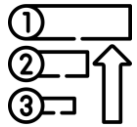
Principles of Accuracy and Reliability

Principle of Coherence and Comparability

Model of quality indicators

Enables comparisons across different statistical areas

Identify system weaknesses



Set priorities for future statistical activities

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Principle of Accuracy and Reliability

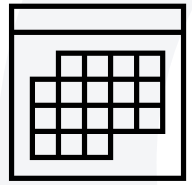
The impact of revisions

Reasons for carrying out revisions:

- Inclusion of new or additional information
- Incorporation of time-lagged data sources
- Adjustments arising from methodological or classification changes
- Correction of errors and omissions in the source data

IMPROVE QUALITY

Two main strategies:



Revision policy publicly available, disclosing any relevant changes

Monitoring revisions - different indicators

The conference is partly financed by the European Union

Principle of Accuracy and Reliability

The impact of revisions

Root Mean Square Relative Error (RMSRE)

$$\text{RMSRE} = \sqrt{\frac{\sum_{t=1}^{N}(X_t - Y_t)^2}{\sum_{t=1}^{N}(\overline{X_t} - X_t)^2}}$$

N: number of observation

X_t: last observation

Y_t: first observation

$\overline{X_t}$: average of the last data evaluation

Directional Reliability Indicator (Q)

$$Q = \frac{N_{11} + N_{22}}{N}$$

N: number of observation

N_{11}: number of times where, at the same time, $\Delta X_t > 0$ and $\Delta Y_t > 0$

N_{22}: number of times where, at the same time, $\Delta X_t < 0$ and $\Delta Y_t < 0$

Bias

$$Bias = \frac{\sum_t bias(t)}{N}$$

$$bias(t) = \begin{cases} 1, if\ last\ version\ higher\ than\ first\ version \\ 0, if\ last\ version\ equal\ to\ first\ version \\ -1, if\ last\ version\ lower\ than\ first\ version \end{cases}$$

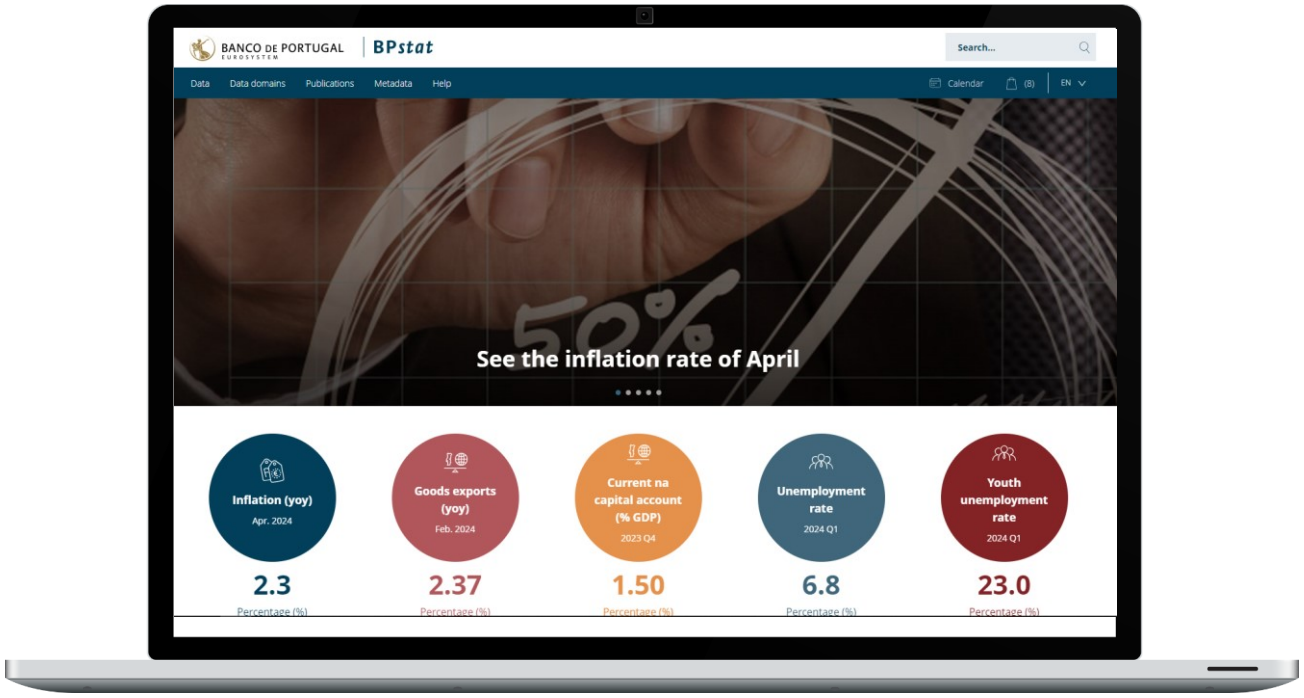
N: number of observation

t: reference date

The conference is partly financed by the European Union

Principle of Accuracy and Reliability

The impact of revisions



- Data warehouse system
- Stores the historical values of all series published on the Banco de Portugal's statistics portal – BPstat
- Data since February 2020
- Analyse either all series or selectively focus on a specific group

Monthly revision (including off-calendar and reverted)

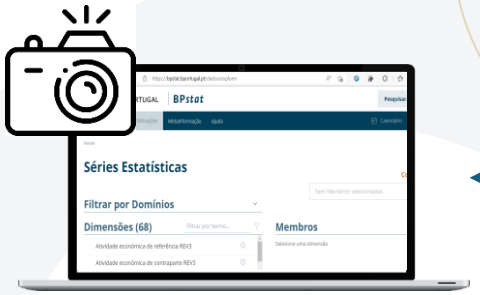
DifferenceP = last - penultimate

1st PUBLICATION of each "series_datereference“

(e.g., series 12504756 and dateref 2021-12 first published in **2022-02**)

PENULTIMATE PUBLICATION of each "series_datereference“

(e.g., series 12504756 and dateref 2021-12 published in **2022-11**)



LAST PUBLICATION of each "series_datereference"

(e.g., series 12504756 and dateref 2021-12 published in **2022-12**)

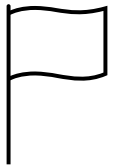
Revision indicators (RMSRE, Q, Bias)

Difference1 = last - first version

The conference is partly financed by the European Union

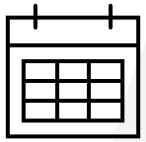
Monthly revision (including off-calendar and reverted)

DifferenceP = last - penultimate

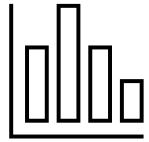


Flag Score:
number of series with a score >=3 (more relevant revisions)

Flag Reverted Revisions:
number of series where the last = first <> penultimate



Flag Out-of-Calendar Revisions:
number of series with revisions outside the calendar



Flag Relevant Series:
number of series classified as more relevant

The conference is partly financed by the European Union

Principle of Coherence and Comparability

What is data consistency?

"Statistics must be consistent: (1) over time, (2) within the dataset that is published in a single release, (3) across datasets and (4) across different frequencies for the same dataset, and that, where appropriate, they must be (5) comparable with statistics of other regions and countries."

The ECB Statistics Quality Framework and quality assurance procedures

The conference is partly financed by the European Union

Principle of Coherence and Comparability

What is data consistency?

Internal consistency

Ensure coherence within the same statistical domain or compliance with established reporting rules.

External consistency

Evaluates whether the data is comparable with other domains or sources of information.

Temporal consistency

Month-on-month and year-on-year rates of change.

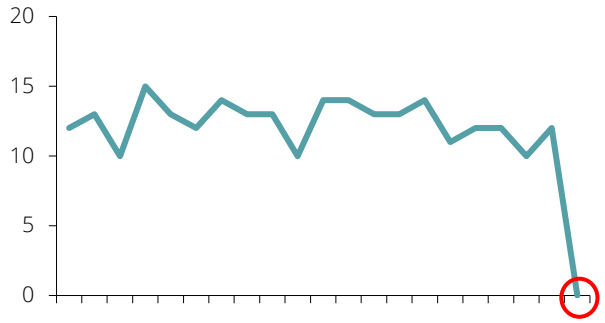
The conference is partly financed by the European Union

Principle of Coherence and Comparability

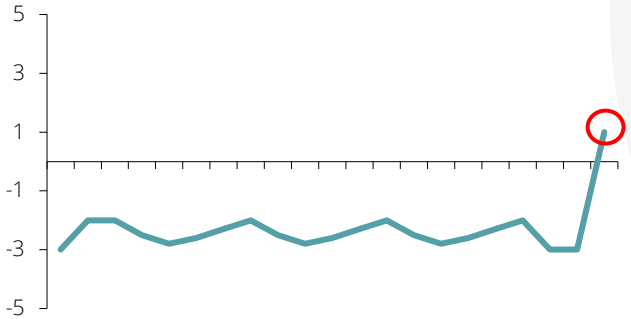
What is data consistency?

Outliers

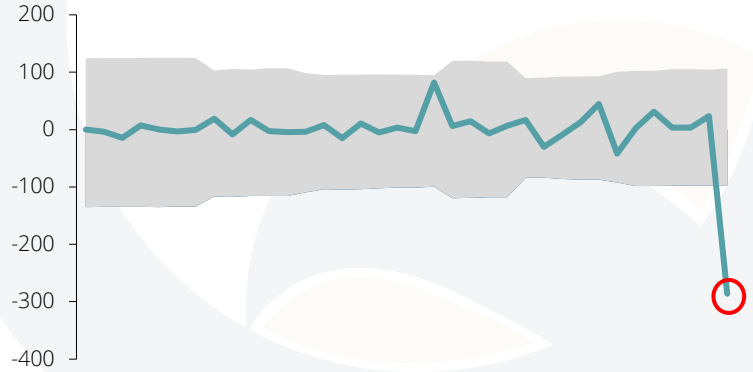
Implausible zeroes



Unexpected signal changes



Unexpected changes in magnitude



INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

How can we see clearly now – a fresh approach to quality control?

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL
eurostat
The conference is partly financed by the European Union

Integrated quality control system

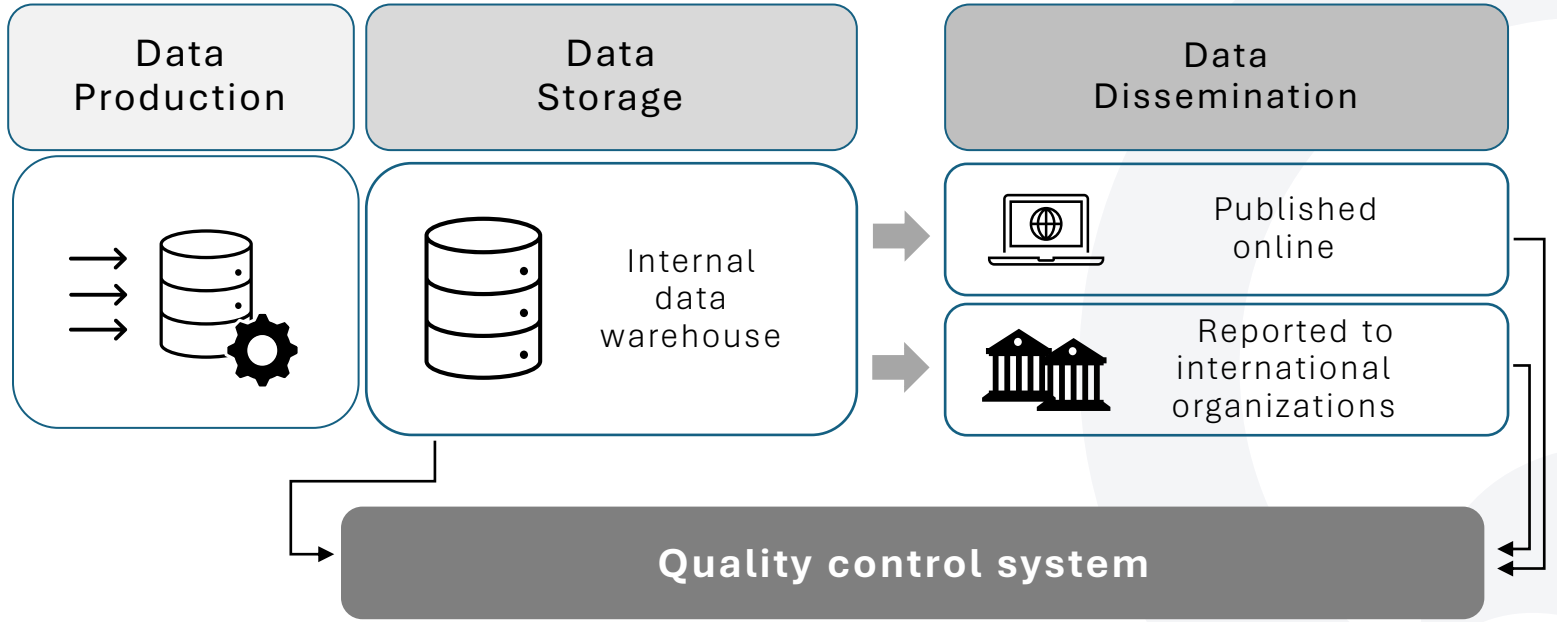
Collection of Statistical Tests

Operating Rules

Web Application

The conference is partly financed by the European Union

How is this integrated process of quality control applied in practice?



This integrated process of quality control, adheres to standardized protocols to ensure data quality across the statistical production cycle

The conference is partly financed by the European Union

Concluding remarks

- Speed should never compromise quality
- Robust quality indicators
- Prioritize both timely acquisition and data integrity
- Trust in central bank's statistics

In Banco de Portugal, there is an ongoing effort to develop and implement more flexible exploring tools enabling a more efficient, transparent, and reliable data quality control.

I CAN SEE CLEARLY NOW OUTLIERS ARE GONE

HOW TO IMPROVE DATA QUALITY OF OFFICIAL STATISTICS?

Thank you for your attention!!

BANCO DE PORTUGAL
EUROSYSTEM

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

Susana Santos, *smsantos@bportugal.pt*

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union