

Tutorial:How to comment on statistics

Statistics Explained

Principles for commentary

Text comments on statistics in all Eurostat’s publications should follow the four content principles below. The comments should be:

- Independent and objective: Comments should be independent from external interference as well as objective, i.e. not be biased by partisan views, questionable values etc.
- Relevant and attractive for the users: Comments should be oriented towards the information needs of the users. Both, visual and textual presentation should be attractive without compromising the other principles.
- Based on facts: All statements should be based on statistical evidence or other facts.
- Policy-relevant but not policy-prescriptive: If statistical data are relevant for an EU policy it is important to make this relationship explicit. However, all such statements have to refrain from evaluating or prescribing policies.

Guidance for the content of comments

How to improve the content of comments Appropriate comments should add value to a publication. Comments may highlight the most interesting results of the data, describe facts not immediately visible when looking at data, explain the meaning or the context of the data, or give information on data quality and data problems. Rather than repeating figures which a reader could extract himself from the tables or graphs, an effort should be made to include interesting comments helping the reader to better understand the data or to make better use of the data, while remaining objective.

All comments should be based on good statistical evidence or other well-justified facts. This includes a thorough check of the statistical significance of the results. In addition, only the known situation should be described. No conclusions should be drawn about the future or other insecure domains, unless scientifically undisputed extrapolation methods are available. Exceptions of the latter exist, e.g., in population statistics (demographic forecasts). It is recommended to always check that results and related comments are not significantly influenced by factors that haven’t been considered or controlled for in the statistical analysis.

Value judgments, in particular about policies and actors, as well as emotionally loaded language have to be avoided.

Journalists use the “inverted pyramid” structure for their articles, and it is recommended to do this as well for articles. The inverted pyramid simply means that the publication starts with the main messages followed by more specific information in decreasing order of importance and increasing level of detail. The reason for this is that the non-expert reader does not want to read the whole publication until he/she finds the conclusions. This and related guidelines on editorial style are described in the UN-ECE document ‘Making data meaningful’.

Explaining the wider context of the data It is often useful to explain why the statistical data is important from a reader’s perspective. This can be done by accompanying statistics with comments on the “real world” context such as issues related to:

- Actual political debates and news covered by the media;
- The everyday life and personal experiences of citizens (energy prices, health, transport, education, etc.);
- The current policy agenda and EU policies;
- Particular societal groups (teens and children, women, the elderly, foreigners etc.);
- Current events and calendar events (important fairs and conferences, world championships, Christmas, Europe Day etc.);
- Interesting scientific findings;

For example, it is recommended to refer to important Community programs and policies, legal acts and other suitable authoritative sources (such as Community web pages).

Comments on policies Commenting policies is an area where particular caution has to be applied. First, all comments have to refrain from evaluating and judging about policies. In addition, comments may not include any policy prescription. Nevertheless, it is recommended to cite important policies when reporting data.

In addition to the mere mentioning of policies, it can be considered to relate statistical results to aims, targets or other elements of policies as long as these elements can be measured and are clearly related to the data (by using agreed methods). In all such cases there should exist strong statistical evidence concerning all important aspects of the results. It is a good idea to cross-check the results with other sources and possibly by using alternative methods (if available). It also needs to be checked whether different subsets of the data (such as different countries, regions, sectors, time periods etc.) show different characteristics. If the results are negative with respect to the policy aims, it is particularly important to choose neutral language. It should also be avoided to focus negative comments too strongly on a small number of countries or other actors. A good idea might be to balance negative messages with some justifiable positive comments (in case of good statistical evidence). In all cases of policy comments it is important to strictly follow the rules of the approval procedure for the publication in question.

Comments on causal relationships Comments on causal relationships between indicators are an important area but also an area in which particular caution is necessary. In many cases, a certain statistical indicator (or rather the relevant entity in the world) has a significant statistical correlation with other indicators. However, statistical methods, such as correlation or regression methods, cannot prove causal relationships between indicators. Unless statistical correlation is backed by clear and undisputed scientific models, comments should refrain from stating causal relationships.

However, statistical analysis may reveal possible candidates for causal relationship. In cases in which the statistical evidence is strong and good scientific arguments for real world relationships between the indicators are available, it is recommended to comment in terms of "correlation", "(possible) influence" or "(possible) contribution" etc. Even when using such cautious expressions, it is mandatory to make sure that all essential influence factors are mentioned and - according to scientific understanding of the area in question - have in principle a causal influence on the phenomenon, in order to not invite critics that something important has been forgotten or has been wrongly taken into account in the analysis.

In specific cases, it might be useful to involve Member States in the discussion on causal relationships, in order to get a more widely agreed justification.

Comments on important (recent) changes or unexpected results Comments should reflect a fresh look on the new data to be published and tell the most interesting messages that can be extracted from the data including unexpected results. It should be avoided to always follow the standard analysis and to report on the same characteristics of the data.

The majority of readers are more interested in getting to know what has changed (recently) than what the

situation is or has been some while ago. Exemptions may be publications with certain legal obligations or a specific purpose which need to report on certain well-defined situations. The media are almost always interested in the most recent changes only.

Comments on important characteristics of detailed data. In some cases a certain subset of data may exhibit an important characteristic that differs from the aggregated total, such as an opposite trend. In such cases, it is often interesting to contrast the characteristic of the total with the characteristic of the subset.

Using information from other sources In some cases it can add value to a publication if information from other sources than Eurostat or the Commission is included, e.g. if it is important to compare the EU with other world countries and regions and relevant data is not available at Eurostat.

In addition, it may be the case that supportive comments can be found in other high-level studies. These arguments can be used in cases where Eurostat could not take the responsibility for the relevant comments.

Including important methodological information Sound methodological comments are an important element of statistical publications. Examples are important definitions, information about the data collection methods, breaks in series or other problems of comparability in space, time etc.

Methodological comments which are essential for the understanding of the data should be placed close to the data it refers to (especially if data could be misinterpreted if the comments were absent). More general methodological information should be included in annexes at the end of the publication.

Comments on the limitations of study results; questions or implications that require further analysis A statistical publication should include comments on the limitations of the data and of the conclusions. For example, it should be checked whether the data and methods used cover well-enough all entities (countries, sectors, time periods etc.) referred to in the results.

Depending on the research situation of the statistical area in question, it can also be beneficial to raise questions or implications that point to future analysis or research. This can add credibility to the study by demonstrating that the author is aware of the limitations of his or her conclusions and/or the relevant research in the area in question.