



# RISK ASSESSMENT SCALE

For privacy risks in the disclosure of statistical information.



NATIONAL INSTITUTE OF STATISTIC AND GEOGRAPHY

---

DECEMBER 2021

---

# STATISTICAL AGENCIES DILEMMA



## Providing data

- About the economic, demographic, social and environmental situation (more and more precision).
- Assuring the availability and quality of official statistics

## Protecting data (providers)

- Avoid the identification of data providers
- Use data from data providers for statistical purposes only

So...

# WHAT TO DO?

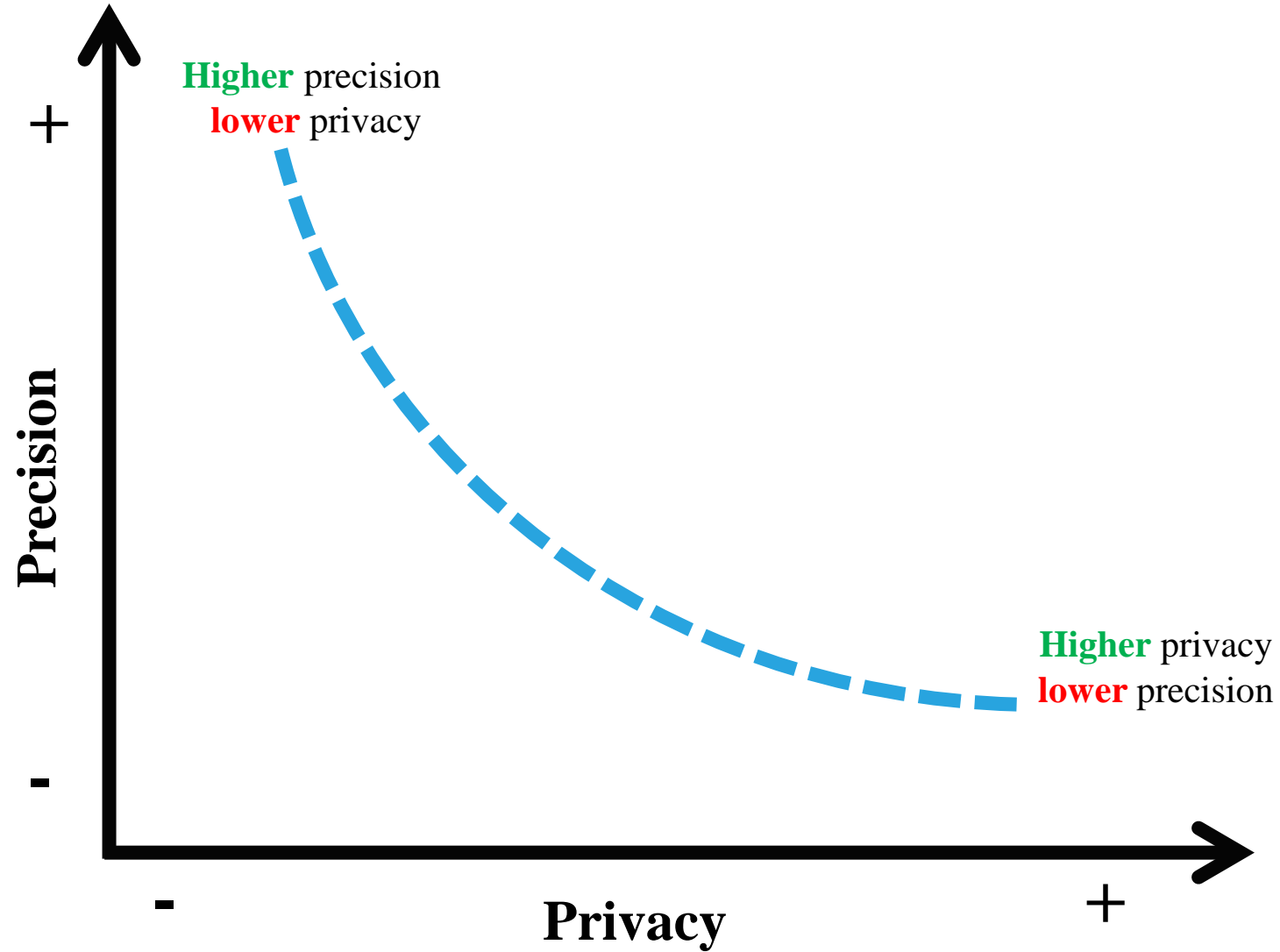


Reference	Requirements
UN Fundamental Principles of Official Statistics – Implementation guidelines	<p>Statistical data producers apply statistical disclosure control methods prior to the release of statistical information.</p> <p>Review by authorized staff of all data prepared for dissemination for possible indirect disclosure.</p>
United Nations National Quality Assurance Frameworks Manual for Official Statistics (UN NQAF Manual)	<p>There should be a balance between the acceptable level of risk of identification of individuals and usability of the data.</p> <p>Appropriate processes are in place to assess the risk of disclosure of sensitive information</p>




# HOW TO DO IT?

# WHERE IS THE BALANCE?



# THE PROPOSAL **MUST ALLOW AT LEAST**



- Characterize different levels of risk that a data provider is identified in the information that is made available to the public.
- That the different risk levels specify the characteristics of the information as well as the circumstances that correspond to it.
- That its application allows statistical offices to find a balance between their role of providing quality information and preserving statistical confidentiality.

# THE PROPOSAL

Identification risk level	Description
High	Identification is immediate. Just by accessing the information it is possible to recognize the person or company to which the data corresponds; or when the identification is deduced from the combination of different data represented in the same product through which the information is presented.
Middle	Identification is achieved by adding or subtracting some classes or groups from the same tabulation and combining the result with other statistical or geographic information products.
Low	Identification is achieved by combining the information with different public and private data repositories using analysis techniques, software and computer equipment.
Null	It is not possible to carry out the identification by any means or by the combination of any Information variable.

# OPERATION POLICIES



- a) The scale can be applied to statistical information as well as georeferenced statistical information.
- b) The scale should be used when the official statistical information is analyzed before it is disseminated
- c) When the risk of identification is high or medium, it is recommended not to disseminate the information with the current attributes and circumstances.
- d) The scale must be applied again each time the information prior to dissemination is modified.





# THANK YOU



Conociendo  
**México**

800 111 46 34  
[www.inegi.org.mx](http://www.inegi.org.mx)  
[atencion.usuarios@inegi.org.mx](mailto:atencion.usuarios@inegi.org.mx)

    **INEGI Informa**