

# PAUTES GENERALS PER GARANTIR LA QUALITAT DE LES DADES OBERTES

## ELUDIR FORMATS DE DADES NO PROCESSABLES

Compartir les dades en **formats oberts i reutilitzables**.



## FER SERVIR UNA CODIFICACIÓ DE CARÀCTERS ESTANDARDITZADA

És recomanable emprar una **codificació de caràcters** internacionalment usada, com per exemple, **UTF-8**.



## ANOMENAR LES COLUMNES ADEQUADAMENT

Fer servir només caràcters en minúscula. Els camps i les seves especificacions han d'estar recollits al diccionari de dades. Tampoc no s'han de fer servir caràcters especials, títols o signes de puntuació. Els espais han de ser substituïts per guions.

## EVITAR VALORS ABSENTS

Per evitar confusions, cal **marcar clarament els valors absents com a valors nuls (N/A)**.



## PREVINDRE LA DUPLICITAT DE REGISTRES

Estandarditzar la recollida de dades i el seu emmagatzematge, centralitzant el procés en un únic sistema d'informació, de manera que les duplicitats siguin fàcilment detectables i puguin ser eliminades automàticament.

## ESTANDARDITZAR VALORS DE DADES

Per normalitzar l'estructura i els valors dels camps, és recomanable fer servir **vocabularis de referència**. L'estructura ha de ser documentada al **diccionari de dades**.



## ATRIBUTS DE LA QUALITAT DE LES DADES



Consistència



Accessibilitat



Precisió



Comprensibilitat



Traçabilitat

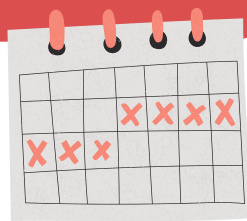
## PROPORCIONAR UNA QUANTITAT ADEQUADA DE DADES PER FACILITAR-NE L'ANÀLISI

Assegurar-se que es publica una **quantitat raonable de dades** perquè hi hagi prou context.



## FORMAT DE DATA I HORA

Les dates s'han de codificar utilitzant la norma ISO, és a dir, **aaaa-mm-dd** per a la data i **hh:mm:ss** per a l'hora.



## FORMAT DE DADES NUMÈRIQUES

Fer servir com a separador decimal el punt (internacionalització). Evitar separadors de miler. Valors negatius amb signe (-). En columnes amb valors sencers, no empreu separadors decimals ni barregeu text amb valors numèrics.

## EVITAR LA BARREJA D'ESCALES NUMÈRIQUES

Intentar que l'escala no variï al llarg del temps. En cas que sigui necessari, proporcioneu les dades en ambdues escales i documenteu el canvi.



## DEFUGIR LA BARREJA DE RANGS

Publicar les dades amb el **nivell més alt de desagregació**. Si no és possible, mantenir la consistència en tots els valors de la variable.



## INCORPORAR VARIABLES AMB INFORMACIÓ GEOGRÀFICA

Publicar les dades amb **coordenades geogràfiques en dues columnes independents**: "latitud" i "longitud".



## EVITAR LA INCORPORACIÓ DE SUBTOTALS, TOTALS O AGRUPAMENTS

Presentar el **major nivell de desagregació possible** de les dades.



## ELUDIR LA FRAGMENTACIÓ DE DADES

Millorar l'**organització i l'etiquetatge** dels continguts, sent necessari establir connexions entre els diferents *datasets*.



Font: "Guía práctica para la mejora de la calidad de datos abiertos: Secretaría de Estado de Digitalización e Inteligencia Artificial del Ministerio de Asuntos Económicos y Transformación Digital" (Setembre, 2022).

<https://datos.gob.es/ca/documentacion/guia-practica-para-la-mejora-de-la-calidad-de-datos-abiertos>

