

# Tipologia i cicle de vida de les dades

## Pràctica 1

### Estudiants:

Victor Garcia Domingo

Ivan Jalencas Lobera

### 1. Context. Explicar en quin context s'ha recol·lectat la informació. Explicar per què el lloc web triat proporciona aquesta informació.

La informació s'ha recol·lectat a partir de la pàgina web oficial del Centre Sismològic Euro-mediterrani (EMSC). Tal i com s'explica a la pàgina, aquest centre va ser fundat l'any 1975 a partir de les recomanacions de la Comissió Sismològica Europea (ESC), que forma part de l'Associació Internacional de Sismologia i Física de l'Interior de la Terra (IASPEI). Alhora, aquesta és una branca de la Unió Internacional de Geodèsia i Geofísica (IUGG).

El 1987, el Consell d'Europa va encarregar l'EMSC de proveir-li amb avisos sísmics amb l'objectiu de prevenir moviments sísmics, i protegir-s'hi i organitzar-se en cas d'ocurrència. A través d'aquesta web, informen de tots aquells moviments que ocorren tant a Europa com al món.

### 2. Definir un títol pel dataset. Triar un títol que sigui descriptiu.

Moviments sísmics a Espanya (agost-octubre 2020).

### 3. Descripció del dataset. Desenvolupar una descripció breu del conjunt de dades que s'ha extret (és necessari que aquesta descripció tingui sentit amb el títol triat).

El data set conté dades temporals, espacials i tècniques d'una sèrie de moviments sísmics ocorreguts a Espanya.

### 4. Representació gràfica. Presentar una imatge o esquema que identifiqui el dataset visualment



**5. Contingut. Explicar els camps que inclou el dataset, el període de temps de les dades i com s'ha recollit.**

- Les dades del dataset corresponen al mes de setembre de 2020.
- S'han recollit a partir de la URL de cerca de terratrèmols, la qual té la següent aparença:  
[https://www.emsc-csem.org/Earthquake/?filter=yes&start\\_date=&end\\_date=&min\\_depth=&max\\_depth=&min\\_mag=&max\\_mag=&region=&min\\_intens=&max\\_intens=](https://www.emsc-csem.org/Earthquake/?filter=yes&start_date=&end_date=&min_depth=&max_depth=&min_mag=&max_mag=&region=&min_intens=&max_intens=)

Es compona dels següents paràmetres:

- filter: si es vol filtrar els resultats
- start\_date: des de quina data
- end\_date: fins a quina data
- min\_depth: profunditat mínima
- max\_depth: profunditat màxima
- min\_mag: magnitud mínima
- max\_mag: magnitud màxima
- region: regió
- min\_intens: intensitat mínima
- max\_intens: intensitat màxima

- Els camps són els següents:
  - Date:
    - Tipus: data
    - Format: DD/MM/AAAA
    - Descripció: data del moviment sísmic
  - Time UTC:
    - Tipus: data
    - Format: HH/MM/SS
    - Descripció: hora del moviment sísmic
  - Latitude:
    - Tipus: real amb signe
    - Format: quatre posicions, dos enters i dos decimals
    - Descripció: latitud del moviment sísmic
  - Latitude Cardinal:
    - Tipus: caràcter
    - Format: una posició
    - Descripció: direcció cardinal de la latitud del moviment sísmic
  - Longitude:
    - Tipus: real amb signe
    - Format: quatre posicions, dos enters amb dos decimals
    - Descripció: longitud del moviment sísmic
  - Longitude Cardinal:
    - Tipus: caràcter
    - Format: una posició
    - Descripció: direcció cardinal de la longitud del moviment sísmic
  - Depth:
    - Tipus: enter
    - Format: no té limitació de grandària
    - Descripció: profunditat a la que ha succeït el moviment sísmic
  - Magnitude Type:
    - Tipus: caràcter
    - Format: dos caràcters
    - Descripció: tipus de magnitud del moviment sísmic
  - Magnitude:
    - Tipus: real
    - Format: dues posicions, un enter i un decimal
    - Descripció: magnitud del moviment sísmic
  - Region:
    - Tipus: caràcter
    - Format: cinc posicions
    - Descripció: regió a la que ha succeït el moviment sísmic
  - Last Update:
    - Tipus: data
    - Format: DD/MM/AAAA HH/MM/SS
    - Descripció: última actualització de la notificació del moviment sísmic
  - Earthquake ID:
    - Tipus: enter
    - Format: sis posicions

- Descripció identificador del moviment sísmic

**6. Agraïments. Presentar el propietari del conjunt de dades. És necessari incloure cites de recerca o anàlisis anteriors (si n'hi ha).**

Aquest dataset s'ha elaborat a partir de la informació recol·lectada mitjançant tècniques de web scraping de la web <https://www.emsc-csem.org/>. El Centre Sismològic Euromediterrani es dedica bàsicament a la recol·lecció dels paràmetres principals dels sismes ocorreguts a la regió euromediterrània per el seu estudi.

**7. Inspiració. Explicar per què és interessant aquest conjunt de dades i quines preguntes es pretenen respondre.**

El conjunt de dades pot fer-se servir per a construir models de predicció de terratrèmols de gran intensitat a partir de patrons de sismicitat. Un altre ús pot ser la capacitat de poder comprovar com evoluciona, al llarg del temps, l'activitat sísmica de zones en concret. Això, a la llarga, també es podria fer servir en models predictius.

**8. Llicència. Seleccionar una d'aquestes llicències pel dataset resultant i explicar el motiu de la seva selecció:**

Les dades original que extraïem formen part d'una web amb llicència Copyright. Com a tal, està prohibit qualsevol ús, reproducció o transmissió sense consentiment per escrit. Una excepció a aquesta prohibició és l'ús acadèmic de les dades. Per tant, podem emprar les dades sense la necessitat d'un consentiment per escrit.

Tal com dèiem en la secció d'agraïments, el dataset s'ha elaborat a partir de les dades obtingudes a partir de la web <https://www.emsc-csem.org/> a fi de poder elaborar aquesta pràctica i, en qualsevol cas, les dades continuen sent del seu propietari.

Aquesta pràctica té un ús acadèmic i no es comercialitza de cap manera. A més, permetem que es faci treball derivatiu sempre i quan es distribueixin amb les mateixes restriccions.

Per aquests motius, hem decidit publicar aquest treball sota la llicència CC BY-NC-SA 4.0

Contribucions	Signa
Recerca prèvia	VG, IJL
Redacció de les respostes	VG, IJL
Desenvolupament codi	VG, IJL