On viuen les dades

Tres formes, sis fonts i un conjunt de dades vostre

Noms _____

Necessiteu un navegador i un editor de text. Cap compte, cap contrasenya. La Part A s'omple amb tota la classe; la Part B, per parelles. Escriviu aquí.

Referència

Forma	Què vol dir	Exemple
Estructurades	Files i columnes fixes, totes les files amb la mateixa forma.	Una taula d'una base de dades, un CSV net.
Semiestructurades	Etiquetades, però la forma pot variar.	JSON, XML.
No estructurades	Sense cap model de dades.	Fotos, àudio, el PDF d'una carta, un xat.

Part A · Sis fonts (amb la classe)

Per a cada font: la seva forma i què necessiteu per entrar-hi.

Font	Exemple quotidià	Forma	Què necessiteu per arribar-hi
BD relacional	Els expedients de l'institut; les comandes d'una botiga.		
CSV	Una exportació de l'app del banc; el quartet de la setmana passada.		
Excel	El full de càlcul que t'envia el cap per correu.		
JSON	El que es descarrega una app del temps.		
API	Spotify, l'horari d'un bus.		
Portal de dades obertes	Portal del GOIB, TidyTuesday.		

A1 Quines fonts demanen credencials?

Sempre _____ Normalment _____ Mai _____

A2 "Un CSV és un Excel." Obriu-ne un de cada en un editor de text. Què hi veieu?

CSV _____ `.xlsx` _____

Part B · La caça (14 min)

Trieu un portal i trobeu-hi **un conjunt de dades que us interessi**: turisme, habitatge, platges, busos, música, jocs...

- **Portal de dades obertes del GOIB**: `dadesobertes.caib.es`
- **TidyTuesday**: `github.com/rfordatascience/tidytuesday`, un conjunt de dades públic nou cada setmana des del 2018. Mireu a `data/<year>/<date>/`, obriu el fitxer `.csv` i premeu **Download raw file**. Algunes propostes:

Tema	Carpeta	Fitxer
Cançons de Spotify	`2020/2020-01-21`	`spotify_songs.csv`: popularitat, ballabilitat, energia, tempo...
Taylor Swift	`2023/2023-10-17`	`taylor_album_songs.csv`: totes les cançons dels àlbums amb les seves característiques d'àudio.
Pokémon	`2025/2025-04-01`	`pokemon_df.csv`: tipus, estadístiques, alçada, pes.
Jocs de Steam	`2021/2021-03-16`	`games.csv`: mitjana i pic mensual de jugadors per joc.

Descarregueu-lo i obriu-lo en un **editor de text**, no en un full de càlcul: heu de veure què separa els valors.

Camp	La vostra resposta
Títol i adreça web	
Formats oferts (CSV, JSON, XLSX...)	
Forma	
Separador i marca decimal	
Files × columnes, aproximadament	
Llicència	
Una pregunta que podria respondre	

B1 Compareu amb una parella que hagi fet servir l'altre portal. Els vostres fitxers separen els valors i escriuen els decimals de la mateixa manera?

B2 Què necessitaríeu per posar el vostre conjunt de dades dins d'una eina com Superset? Demana credencials?

Heu acabat quan La taula de la Part A és plena · la fitxa de la Part B està completa · la vostra pregunta és una que un gràfic podria respondre. **Parella ràpida?** Obriu `pokemon_df.csv` a l'editor de text: quantes columnes té la capçalera, quantes línies té el fitxer i què hi ha escrit al `type_2` de `charmander`?

A casa Un gimnàs vol un quadre de comandament de quanta gent entrena cada hora. Les dades són a la base de dades MySQL de la seva app de reserves i en un fitxer Excel setmanal de recepció. Un paràgraf: per a cada font, la seva forma i què necessiteu per treure'n les dades.