

## Context.

Per tal de generar el conjunt de dades, la informació s'ha recol·lectat a partir de la pàgina web <https://www.normacomics.com/>

La web és una tenda digital especialitzada en còmics, motiu pel qual podem veure una gran varietat de productes a la web.

Norma Comics és considerada una de les llibreries més important d'Europa en l'àmbit dels còmics. Es diferencien tres àrees de còmics: europeu, americà i manga. També disposen de més àrees com per exemple, seccions de còmic infantils, figures, regals... que per al nostre objectiu resulten menys interessants.

La secció que més ens interessa és la secció de novetats, ja que per al nostre conjunt de dades volem obtenir les novetats de l'última setmana. La tenda és tot un referent, la llibreria ha guanyat el premi a la millor tenda de còmics de 2018. Per tant les dades obtingudes venen d'una font de dades de confiança.

## Títol del dataset.

Novedades semanales - còmics

## Descripció del dataset.

El conjunt de dades ens mostra els còmics nous apareguts durant la setmana al web de Norma Còmics.

Representació gràfica.



La imatge l'hem dissenyada nosaltres mateix. Hem seleccionat 9 còmics de la secció de novetats de la web Norma Comics, i a més els còmics triats formen part de les àrees més representatives en les que ells treballen (europeu, americà i manga). Hem inclòs tres paraules, Novetats Setmanals Comics.

## Contingut.

Per cada còmic, que correspon a un registre del conjunt de dades, tenim:

ISBN	Número Estándar Internacional de Libres (International Standard Book Number en anglès) - Clau primària - Cadena de caràcters
titulo	Títol del còmic - Cadena de caràcters
coleccion	Col·lecció a la que pertany el còmic. El domini és: Americà, Europeu, Manga, Altres - Cadena de caràcters
precio	Preu del còmic en divisa euros. - Numèric Nota: Amb dos decimals
autor	Autor del còmic - Cadena de caràcters Nota: Un còmic pot tenir més d'un autor
editorial	Editorial del còmic - Cadena de caràcters
serie	Sèrie a la que pertany el còmic. - Cadena de caràcters
fecha_venta	Dada de venda del còmic (DD/MM/YYYY). - Date
formato	Format del còmic. - Cadena de caràcters
paginas	Nombre de pàgines del còmic. - Numèric.

El període de temps de les dades és setmanal, perquè el que volem mostrar són els còmics novetat. Per tant, el conjunt de dades s'actualitzarà per exemple els divendres de matinada. Les dades es recullen de la web Norma Comics.

## Agraïments.

Per a poder generar el conjunt de dades, s'ha accedit a la pàgina web de Norma Comics (<https://www.normacomics.com/>)

La web és una tenda digital especialitzada en còmics on va ser completament renovada l'any 2010. Norma Comics no sols té tenda digital, té tenda física en Barcelona i en altres parts d'Espanya. La tenda de Barcelona és considerada una de les llibreries més importants d'Europa en l'àmbit dels còmics, on podem destacar que l'any 2018 va guanyar el premi a la millor tenda de còmics.

## Inspiració.

El conjunt de dades presentat respon a la intenció de mantenir actualitzada una base de dades de còmics. Setmanalment, amb el csv obtingut a partir de les dades de la web de Norma Comics, serà possible afegir totes les novetats.

En l'àmbit de mineria de dades, es podria crear un model capaç de respondre preguntes com quines editorials tenen més probabilitat de fer nous llançaments, per d'aquesta manera apostar per una editorial o una altra. També es poden construir models per fer un seguiment dels nostres autors o sèries preferits, o inclús predir el preu d'un còmic en el suposat cas que ens falti aquesta informació.

## Llicència.

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

- Released Under CC0: Public Domain License
- Released Under CC BY-NC-SA 4.0 License
- Released Under CC BY-SA 4.0 License
- Database released under Open Database License, individual contents under Database Contents License
- Other (specified above)
- Unknown License



La llicència de Creative Commons BY-NC-SA permet ajustar i construir a partir de l'obra original amb finalitats no comercials, sempre i quan sigui reconeguda l'autoria i qualsevol nova creació es mantingui sota una llicència amb els mateixos termes.

Sobretot és important el fet de mantenir la no-comercialitat de l'eina donada la procedència comercial de les dades.

## Codi.

Es pot consultar el codi font a través del següent enllaç o bé accedint a la carpeta codi.

## Dataset

Es pot consultar el dataset a través del següent enllaç.

## Taula de contribucions

Contribucions	Signa
Recerca prèvia	Ivan Borrego García, Cristina Sanchis Puerto
Redacció de les respostes	Ivan Borrego García, Cristina Sanchis Puerto
Desenvolupament codi	En curs