

# Proyecto BI - Análisis de *Amazon Kindle Books*



**Ainoa Iglesias Dasilva**

[alu0101164403@ull.edu.es](mailto:alu0101164403@ull.edu.es)

2023, Gestión de Comunicación y Conocimiento empresarial

## CONTEXTO

Amazon Kindle se destaca actualmente como una de las plataformas líderes para la distribución y consumo de libros electrónicos. La plataforma cuenta con una amplia variedad de libros y cuenta cada vez con más popularidad.

Kindle cuenta con una suscripción mensual llamada Kindle Unlimited, la cual da acceso a más de 4 millones de ejemplares.

A través de la web [Kaggle](#) se ha obtenido una extensa fuente de datos que proporciona información detallada de una parte de los libros de Kindle incluyendo título, autor, precio, número de reseñas, calificación, categorías y fecha de publicación, entre otros.

Se quiere mostrar de manera sencilla a autores y vendedores datos que sirvan para mejorar las ofertas y alcance en la plataforma.

## OBJETIVOS

El objetivo principal de este análisis es dar a los autores y vendedores datos clave que les permita mejorar sus ofertas y alcance en la plataforma. Los objetivos específicos son:

- Comprender tendencias en ventas y popularidad de autores.
- Análisis de géneros y su popularidad.
- Analizar la influencia de 'Best Seller' y 'Elegido por el editor' en la popularidad.
- Tendencias de publicación en el tiempo.
- Obtener predicción sobre géneros populares en el futuro.

## ACTORES-ROLES

- **Editoriales:** Tomar decisiones estratégicas basadas en el análisis de tendencias de géneros y en las relaciones entre calificación, precios, géneros, número de páginas.
- **Autores:** Utilizar el análisis para mejorar la calidad y contenido de sus futuras publicaciones, orientándose hacia lo que es popular según los datos de análisis de géneros.
- **Equipos de Marketing:** Mejorar campañas y promociones de libros mediante el análisis de géneros, teniendo en cuenta la influencia de las etiquetas BestSeller y

EditorsPick.

- **Lectores:** Contribuir con reseñas y calificaciones, siendo indispensables para los análisis de tendencias y la toma de decisiones por parte de los anteriores.

## KPIs

- **Comprender tendencias en ventas y popularidad de autores.**
  - N° libros publicados por autor
  - Calificación media de cada autor
  - N° reseñas por autor
  - N° ventas por autor
  - Popularidad de cada autor (suma de ventas y descargar en unlimited, promedia de valoraciones y número de comentarios)
- **Análisis de géneros y su popularidad.**
  - N° libros totales por género
  - N° reviews totales por género
  - N° descargas totales por género
  - Calificación promedio por género
- **Analizar la influencia de un 'Best Seller' y 'Elegido por el editor' en la popularidad**
  - % libros con esas etiquetas
  - Media de reseñas para libros con estas etiquetas
  - N° ventas libros con estas etiquetas y sin etiquetas
- **Tendencias de publicación en el tiempo**
  - Número de libros publicados por año.
  - N° libros por género publicado cada año
  - % de géneros publicados por año

## ACCIONES/DECISIONES QUE PUEDEN TOMAR LOS ACTORES

- **Editoriales**
  - Adquisición de libros basado en géneros/tendencias populares
  - Ajustar precios para maximizar la popularidad.
  - Desarrollo de estrategias de marketing.
- **Autores:**

- Orientar el contenido de sus futuros libros según las tendencias.
- Orientarse a géneros y temas populares.
- **Equipos de Marketing:**
  - Desarrollo de estrategias promocionales basadas en géneros y libros populares.
  - Considerar las etiquetas 'best seller' y 'editor pick' en las campañas.
- **Lectores:**
  - Proporcionar calificaciones y comentarios a sus compras en kindle para mejorar los datos estadísticos.
  - Participar activamente en la comunidad Kindle.

## FUENTES DE INFORMACIÓN

Los datos analizados provienen de la web [Kaggle](#), con información de 130 mil libros electrónicos de Amazon Kindle. Estos datos fueron recopilados mediante [web scraping](#) en octubre de 2023.

Además, para enriquecer el conjunto de datos, se agregaron varias columnas con [datos sintéticos generados con Python](#), número de ventas, popularidad, número de páginas, número de descargas en unlimited y el número de listas de deseos en los que se encuentra el libro. Para generarlos, el número de páginas se toma como un valor completamente aleatorio mientras que en el resto de columnas se intentó seguir algo más realista pero de manera simple, así por ejemplo la popularidad depende del género del libro, las estrellas y número de valoraciones, y las ventas y descargas de unlimited dependen de la popularidad.

## ANÁLISIS DE FUENTES DE INFORMACIÓN

Se dispone de dos tablas de datos en formato CSV que constituyen la base fundamental para alcanzar los objetivos previamente establecidos. Estas tablas, disponibles en el siguiente [enlace](#), brindan una visión detallada de 130 mil libros electrónicos de Amazon Kindle, recopilados mediante scraping en octubre de 2023.

En este análisis se llevó a cabo tareas de comprensión, limpieza y preparación de los datos. Para ello se comprueba los formatos de las columnas, que datos son útiles para cumplir los objetivos y entender la estructura y una lectura de la web donde se obtuvo los datos para saber que es cada columna.

Como se dijo anteriormente, se decidió añadir columnas adicionales a los datos originales, generados con datos sintéticos. Esto es para dar una visión más completa y realista del proyecto.

Durante la revisión del fichero csv que contiene los datos, se comprobó que algunas columnas que se esperaban como booleanos (true, false) en realidad contenían texto. De manera similar, la columna fecha estaba en formato texto.

113399	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	4.6	0.299	VERDADERO	12	FALSO	FALSO	FALSO	FALSO	Humor & Entertainment
113400	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	4.6	0.799	FALSO	25	FALSO	FALSO	FALSO	FALSO	2019-08-27 Children's eBooks
113401	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	4.7	0.3298	FALSO	11	FALSO	FALSO	FALSO	FALSO	Engineering & Transportation
113402	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	4.5	743.299	VERDADERO	10	FALSO	FALSO	FALSO	FALSO	2017-04-30 Computers & Technology
113403	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	4.6	89.699	FALSO	10	FALSO	FALSO	FALSO	FALSO	2020-03-04 Computers & Technology
113404	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	0.0	0.0	FALSO	10	FALSO	FALSO	FALSO	FALSO	2023-09-27 Computers & Technology
113405	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	5.0	1.2473	FALSO	1	FALSO	FALSO	FALSO	FALSO	Arts & Photography
113406	<a href="https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg">https://m.media-amazon.com/images/I/51jZ8qZ8qZ8q.jpg</a>	0.0	0.999	FALSO	17	FALSO	FALSO	FALSO	FALSO	2023-09-26 Business & Money

En al menos las columnas autor y fecha se pueden encontrar valores vacíos, en el caso de los autores no es muy relevante para el análisis pero sí las fechas, por lo que se omitirán esas filas al cargarse las tablas en HOP.

De igual manera, puesto que había información en distintos alfabetos, se eliminaron todas las filas con un alfabeto distinto al latino.

132803	B0BX6GWLW	Хроника Ego (of	Marc Фрай	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	0.0	0	14.99	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132804	B0BL1WY8V9	Цифровая бомба	Александр Чер	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.6	5	0.0	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132805	B07V9V84JN	40 اربعون	أحمد العفري	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.6	516	11.2	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132806	B0BTCTCY5S	لنكات يهودي مصري	أثير لزي	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.1	10	7.99	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132807	B07B2LYKSB	ألف قصص سائفة	حاتم حسيني	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.5	87	9.99	FALSO	19	FALSO	FALSO	FALSO	Foreign Language

132869	B0CCMSTBJN	தேவநாபிபுலம்	Meera Krish	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	3.9	66	5.48	VERDADERO	19	FALSO	FALSO	FALSO	Foreign Language
132870	B0CD87FBNR	உயிர் உருவம்	சுமதி சரஸ்வதி	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.4	248	6.49	VERDADERO	6	FALSO	FALSO	FALSO	Parenting & Relationships
132871	B0BJJ853H8	タスクベースで	国際基督教大学	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.1	0	13.48	FALSO	22	FALSO	FALSO	FALSO	Reference
132872	B07JY8P2WP	ハリー・ポッター	J.K.ローリング	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.7	0	0.0	VERDADERO	25	FALSO	FALSO	FALSO	Children's eBooks
132873	B0192CTNQL	ハリー・ポッター	J.K. Rowling	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.7	2335	8.99	VERDADERO	16	VERDADERO	FALSO	FALSO	Teen & Young Adult
132874	B0BJV1LKX4	メイドの物語	CLIFTONL YARI	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.6	48	3.35	VERDADERO	1	FALSO	FALSO	FALSO	2022-10-19 Arts & Photo graphy
132875	B07J5VG4YD	九ノ月 (Trad. 南条 政成)	曹公 實典	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.2	206	06.08	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132876	B07N2QFBR0	包拯傳士的內	羅古魯 黃典	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.2	23	8.49	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132877	B0872RV2F8	一本你希望父母	菲莉絲 張瑞	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.4	0	7.56	FALSO	6	FALSO	FALSO	FALSO	Parenting & Relationships
132878	B09KZ25X1F	一行禪師講 心	一行禪師	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.7	17	8.53	FALSO	19	FALSO	FALSO	FALSO	Foreign Language
132879	B07X8G5PM8	一行禪師講 主	一行禪師	Amazon.com Se	<a href="https://m.media-amazon.com/images/I/51jK000000.jpg">https://m.media-amazon.com/images/I/51jK000000.jpg</a>	4.7	14	6.75	FALSO	19	FALSO	FALSO	FALSO	Foreign Language

Hay columnas que son innecesarias como la url de la imagen del libro que se omitirán durante la creación de nuevas tablas en el modelado de datos con Apache Hop.

## MODELADO DE DATOS

Para almacenar la información se ha optado por utilizar MySQL junto con Workbench. Tras crear la base de datos `kindle_db`, se cargó los datos del fichero CSV en una tabla.

### ● Dimensiones y medidas

#### ○ Dimensiones

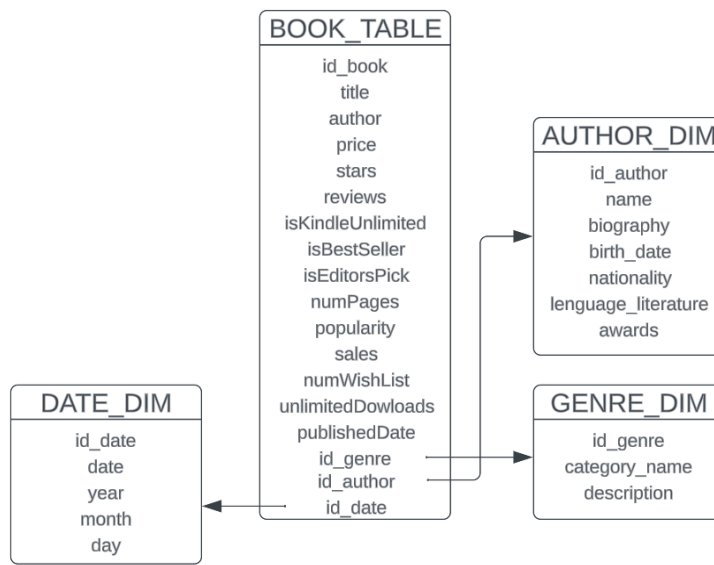
- Autores: id(identificador único), nombre, biografía, fecha nacimiento, nacionalidad, premios, idioma de literatura.
- Géneros: id(identificador único), nombre, descripción.
- Etiquetas (best seller y elegido por autor): id(identificador único), nombre, descripción.
- Fecha: id (identificador único), año, mes, día, trimestre.
- Libros: id(identificador único), título, precio, autor (clave foránea), fecha publicación (clave foránea), número de páginas, genero (clave foránea), idioma, isKindleUnlimited, isBestSeller, isEditorsPick, numero listas deseados, número ventas, número descargas en unlimited, popularidad.

- Medidas
  - N° libros publicados por autor
  - Calificación media de cada autor
  - N° reseñas por autor
  - N° libros por género
  - Calificación promedio por género
  - % libros con esas etiquetas
  - Media de reseñas para libros con estas etiquetas
  - % libros disponibles en unlimited
  - Correlación entre precio y calificación media
  - Número de libros publicados por año.
  - Número de libros publicados por mes.
  - Variación en la cantidad de libros publicados a lo largo del tiempo
  - N° libros por género publicado cada año
  - % de géneros publicados por año
  - Variación de publicaciones por género

En cuanto a estas medidas, puesto que se pueden calcular tanto en Apache Hop como en las herramientas de visualización, se ha optado por no añadirlas a Hop ya que podrían significar un mayor tiempo en la carga y actualización de los datos al añadir estos cálculos.

## ● Relaciones entre datos

La tabla principal es la de libros (Book\_table) la cual se relaciona con el resto a través de claves foráneas como el id\_author o id\_genre. Se hace uso de un esquema estrella como el siguiente:



Para poder hacer predicciones en el tiempo más precisa se creó una tabla Date\_dim la cual almacena más detalles sobre las fechas de publicación de los libros la cuál también se liga con una clave foránea a la tabla principal.

## ● Generación de Datos sintéticos

Para la generación de los datos sintéticos, se implementó un [script en python](#) que genera un número específico de filas y las incorpora al fichero CSV. Este proceso incluye los datos de todas las columnas del conjunto de datos.

En el caso de columnas como el título, nombre autor, estrellas, número de valoraciones, precio, los booleanos de las etiquetas, la fecha de publicación y el número de páginas se generan de manera completamente aleatoria, limitándose a rangos especificados.

Para lograr una aproximación más realista y minimizar la alteración de los análisis reales, se empleó una ponderación para obtener el género. Asimismo, se aplicaron los mismos cálculos para determinar la popularidad, número de ventas, número de descargas y número de listas de deseos.

Este script tiene programada su ejecución para todos los días a las 10am, generando correctamente 5 filas en cada ejecución. Esta imagen muestra las 5 primeras filas generadas:



```

133099 133098,B0C35JDLK8,Anna Karenina,Leo Tolstoy,De Marque,https://m.media-amazon.com/images/I/61RXEYmCtkL.AC_UY218_.jpg,https://www.amazon.com/dp/B0C35JDLK8,4,7,0,0,0,False,29,True,False,False,20
133100 133099,B090M75Y1,Scary Smart: Scary Smart: The Future of Artificial Intelligence and How You Can Save Our World,Mo Gaudat,Amazon.com Services LLC,https://m.media-amazon.com/images/I/71tjf6R-k
133101 133100,B0CFMFPJTN,The Count of Monte Cristo: The Original Unabridged and Complete Edition (Alexandre Dumas Classics),Alexandre Dumas,De Marque,https://m.media-amazon.com/images/I/71zcCB5PvuL-k
133102 133101,B0C1KMFJJC,War And Peace,Leo Tolstoy,De Marque,https://m.media-amazon.com/images/I/715QpZlzmil.AC_UY218_.jpg,https://www.amazon.com/dp/B0C1KMFJJC,4,7,0,0,0,False,29,True,False,False,20
133103 133102,B0BS1KYZB6,Pride and Prejudice,Jane Austen,De Marque,https://m.media-amazon.com/images/I/81Scutrtj4L.AC_UY218_.jpg,https://www.amazon.com/dp/B0BS1KYZB6,4,7,0,0,0,False,29,False,False,f
133104 133104,e5be896a-d3ae-4640-b00a-5a7c96437638,Poor son develop today.,Mark Barber,, , ,4,2,328,45.25,True,29,False,True,False,2020-07-01,Nonfiction,739,15,145335,22,113880
133105 133105,a6488e8a-86af-4343-9f70-8295240e7cd5,Religious affect above consumer of free.,Ryan Dunn,, , ,4,0,191,6.3,False,4,False,True,False,2020-12-14,History,872,9,71442,29,0
133106 133106,8e84c2f4-a761-4c2c-9d1d-30dd90baa6c,Candidate media provide.,Karen Fields,, , ,2,9,783,31.32,True,3,True,False,False,2022-11-28,Education & Teaching,62,24,32424,3,70464
133107 133107,9d4851c5-9a66-4591-8745-e26306014091,Whom manage wrong Fire mean stand.,Joseph Jones,, , ,4,6,270,45.29,False,31,False,False,True,2020-11-23,Self-Help,392,13,7644,1,0
133108 133108,56cc814d-0907-45ad-99e0-dc9d175e4984,Tree future director firm apply.,Charles Drake,, , ,2,7,67,42.35,True,21,True,True,False,2020-08-13,"Mystery, Thriller & Suspense",670,3,14535,15,2

```

Tras incorporar los nuevos datos a CSV, el script también ejecuta una secuencia para que se actualice automáticamente la tabla en MySQL.

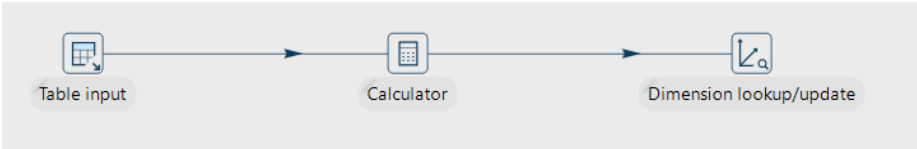
- **Apache Hop (ETL)**

En este paso, se crean las tablas de las dimensiones tal como se especificaron anteriormente. Lo primero que se hizo fue cargar el fichero CSV en una tabla de MySQL ya que al intentar cargarlo directamente en Apache Hop no funcionaba correctamente.

Para cada tabla habrá una clave primaria y una technical key. Las tecncial key que son is\_book, author\_id, genre\_id y date\_id son usada en la tabla principal, books\_tables, como claves foráneas que conectan con las tablas genre\_dim, author\_dim y date\_dim ya que no tendrían que estar duplicados.

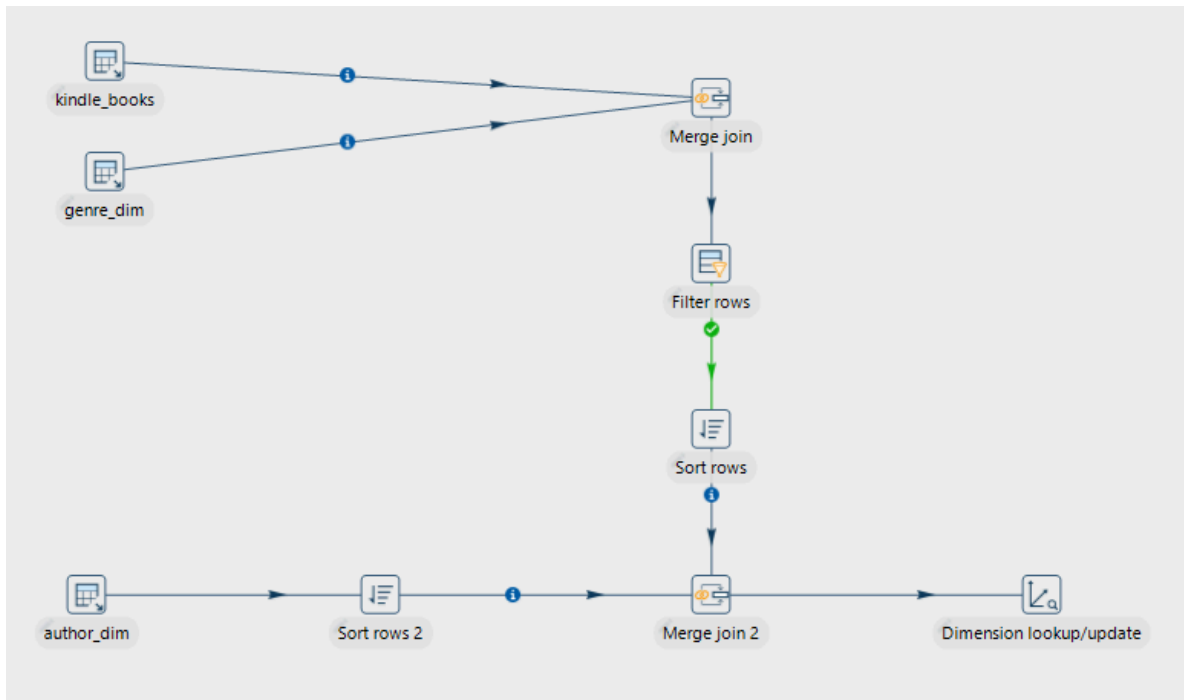
A continuación se muestran los pipelines y workflows creados en Apache Hop.

- **Date\_dim:** a partir de la tabla que carga el CSV, se obtienen las fechas existentes en los datos y se hace una separación en columna del año, día y mes. Si fuera necesario en un futuro sería sencillo añadir otras columnas como el trimestre.

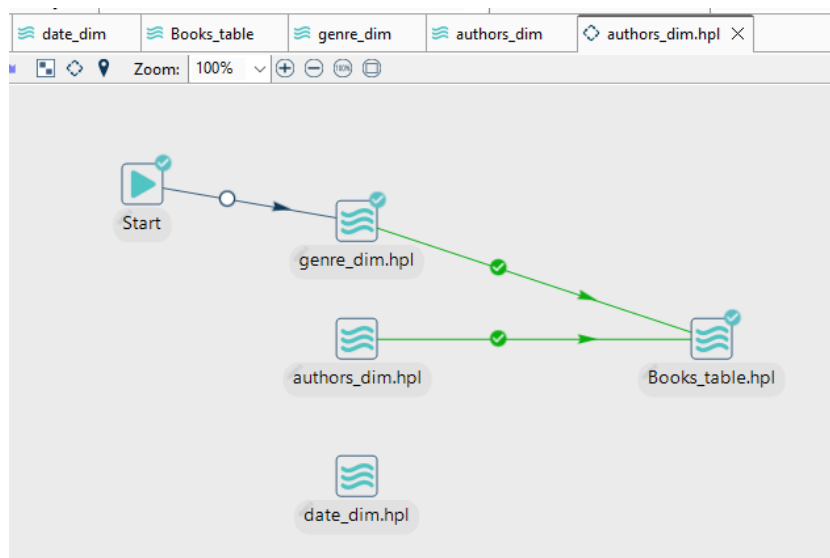




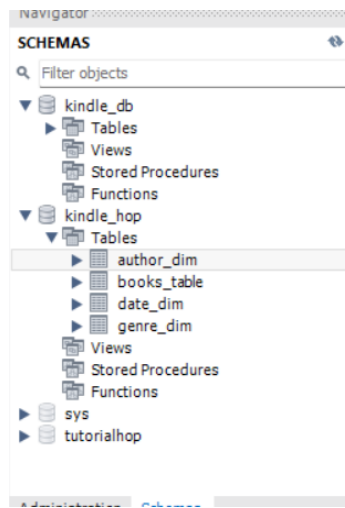
- Books\_table: a los datos de la tabla inicial se añaden las columnas con los id de género y autores. Además la salida contendrá solo las columnas útiles eliminando por ejemplo las columnas url. Para este paso fue necesario usar SORT para ordenar las columnas por las que se hará la igualdad en los merge join.



- Workflow: en el workflow se añaden los pipelines anteriores para poder ejecutar y actualizar las tablas desde un solo sitio.

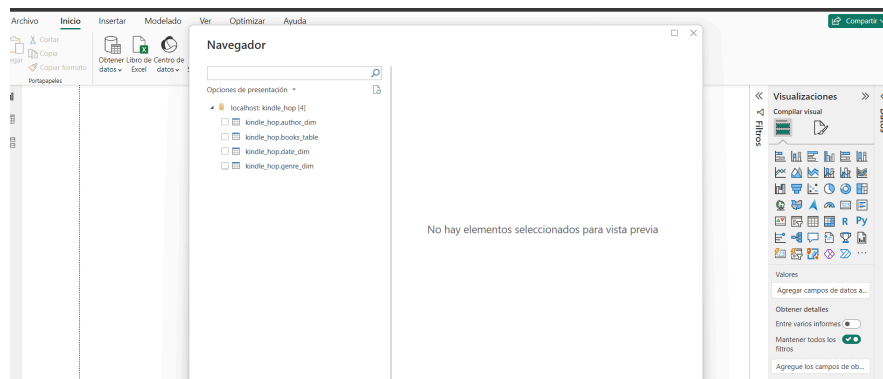


El esquema en la base de datos queda del siguiente modo:

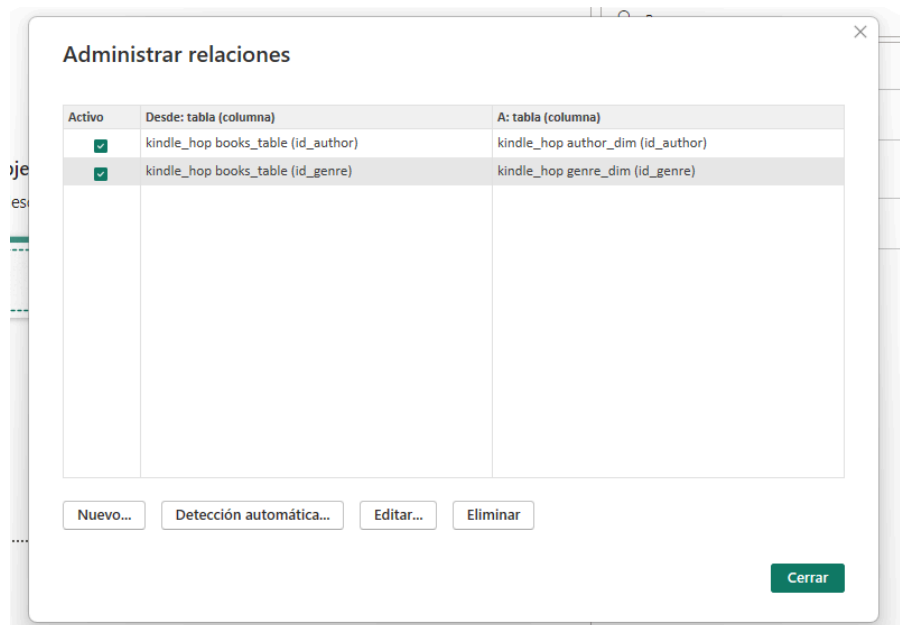


## ANÁLISIS DE LOS DATOS

Haciendo uso de Power BI, se ha procedido al análisis de los datos para obtener los objetivos especificados anteriormente.

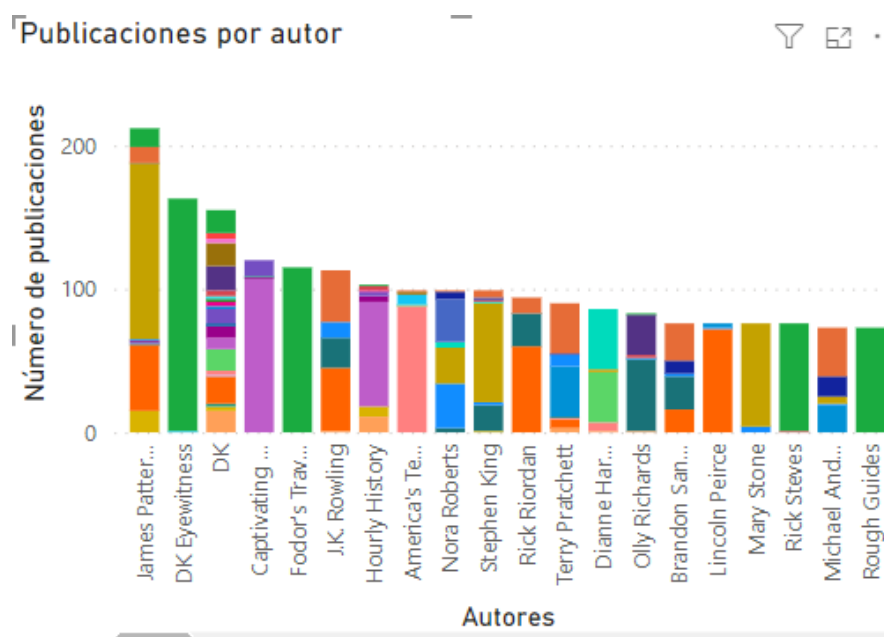


Las tabla principal está relacionada por la clave foránea a las demás.

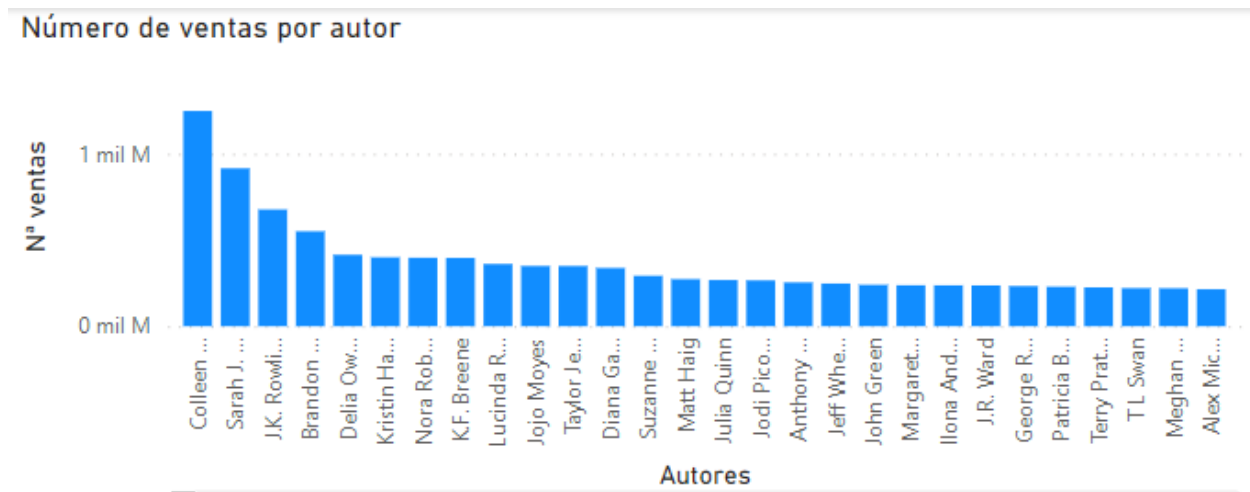


## ● Dashboard

**Tendencias en ventas y popularidad de los autores.** Para este punto se ha decidido usar una gráfica de barras apiladas que muestre el número de libros publicados por autor, de manera que podamos observar si hay relación entre el número de libros y otros datos y los géneros de sus publicaciones.



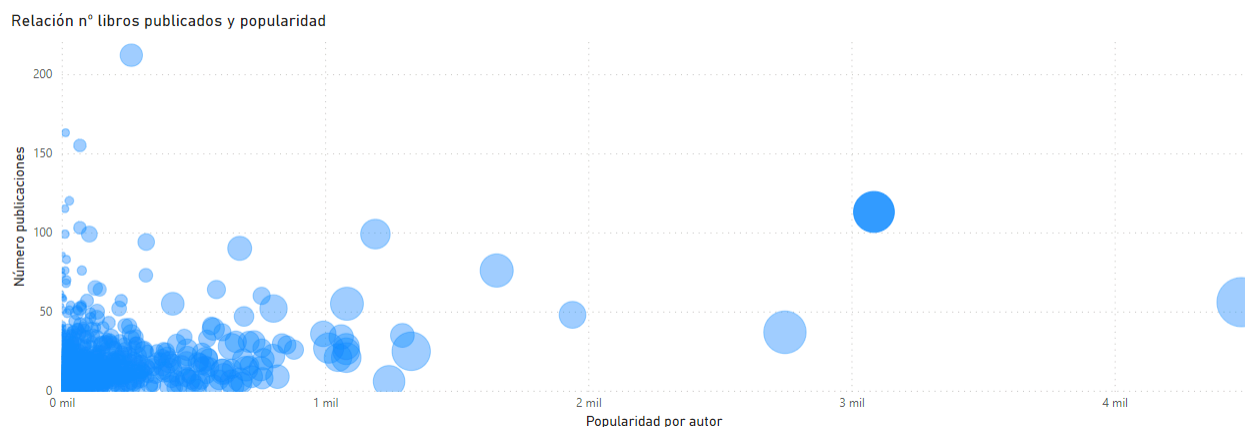
La siguiente gráfica de barras muestra la contribución de cada autor por el número de ventas totales.



Power Bi permite filtrar datos muy fácil pulsando sobre las columnas de las tablas lo que muestra su posición en el resto de gráficos del dashboard.

La siguiente gráfica es de dispersión, se busca identificar patrones, como autores con muchas publicaciones y baja popularidad o viceversa. También se visualiza la cantidad de revisiones para evaluar el nivel de participación de los lectores.

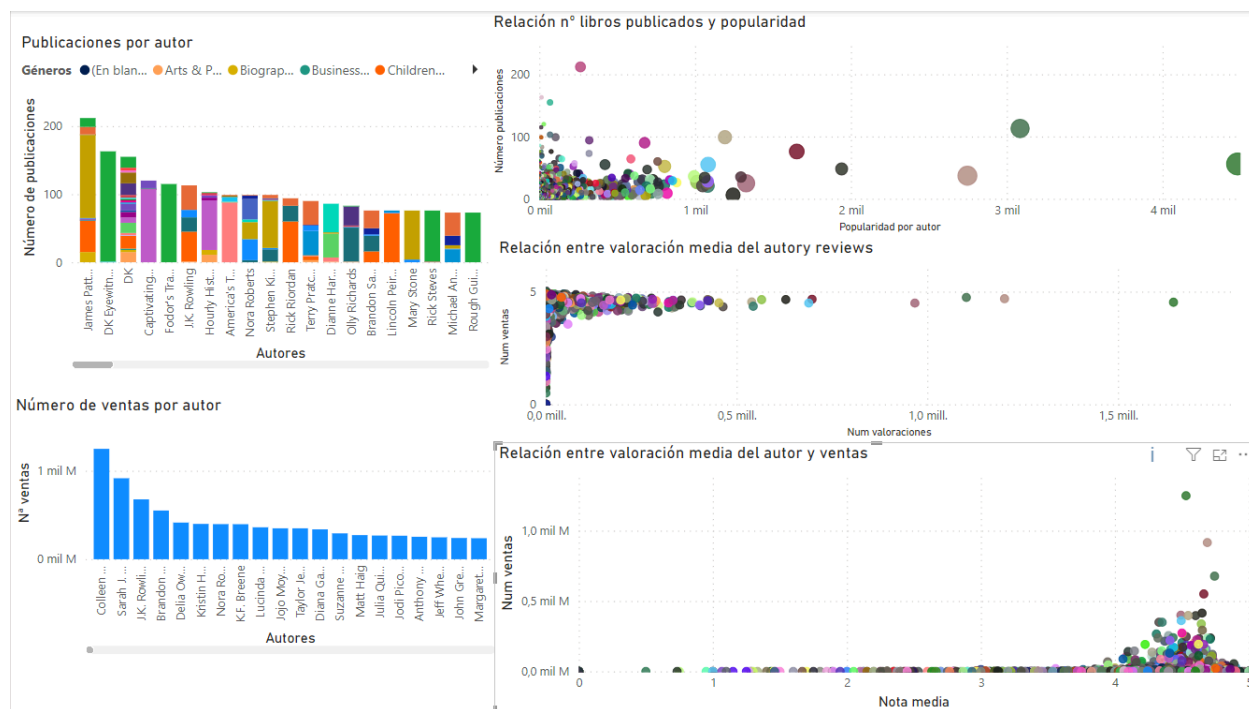
La popularidad de un autor se ha calculado como una nueva medida en PowerBi sumando las ventas totales de sus libros, descargas en kindle unlimited.



Por último se añadieron dos gráficas más de dispersión que tienen la finalidad de explorar si la nota de los usuarios se relaciona con el número de ventas. ¿Los libros mejor valorados tienden a vender más? Y la segunda analiza si la nota de los usuarios

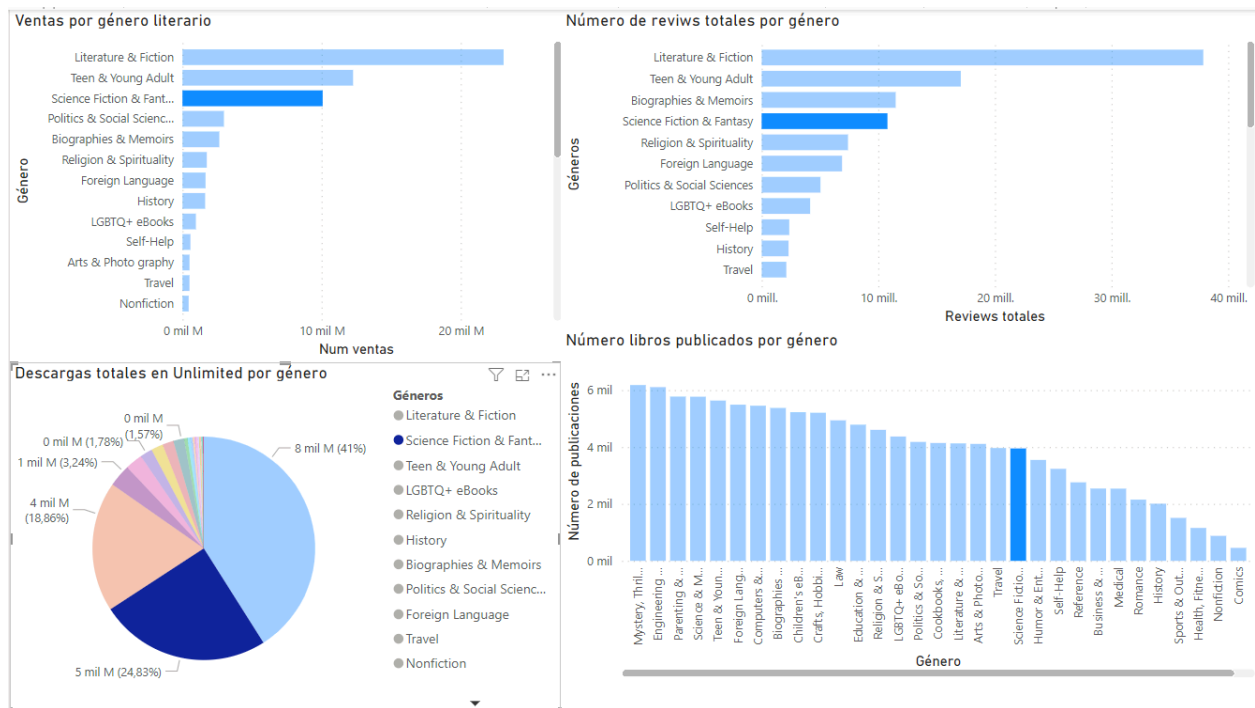
está vinculada con la cantidad de revisiones. ¿Los libros mejor valorados reciben más comentarios?

Esta sería una vista general de este dashboard.



**Análisis de géneros y su popularidad.** Estas gráficas cuentan con datos para saber si hay géneros más populares como las ventas por género, número de reviews por género, número de publicaciones y número de descargas en unlimited.

Este sería el dashboard relacionado con este objetivo.

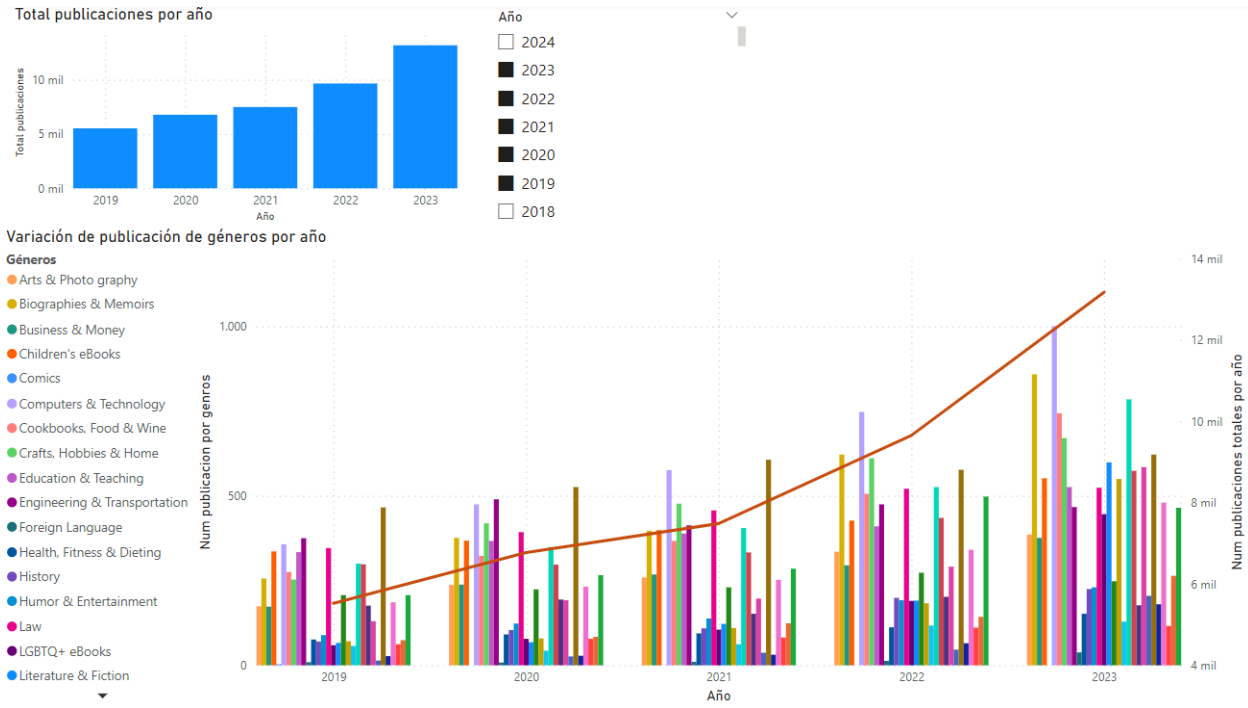


De estas vistas se puede observar que no hay mucha relación entre ser un género popular y la cantidad de libros que se publican de ese mismo género.

**Tendencias de publicación en el tiempo.** Se busca obtener información que pueda indicar si hay alguna tendencia en la publicación de libros, qué es lo que influye en estas tendencias.

Se muestra una tabla en la que se pueden seleccionar los años a mostrar en las gráficas para que la visualización sea más sencilla. Una gráfica de barras para ver el número de publicaciones por año. Y otra para mostrar el número de publicaciones por géneros y año.





## CONCLUSIÓN

El análisis realizado sobre la plataforma de libros electrónicos Kindle da insights valiosos, a pesar de algunos desafíos encontrados en el proceso. La granularidad limitada de los datos de ventas y la imposibilidad de usar Apache Superset afectaron la profundidad del análisis.

El uso de datos sintéticos para añadir la columna de número de ventas totales de cada libro, afectan en la realidad de los datos pero sirvió para mostrar mejores análisis.

En cuanto al objetivo, creo que en parte fue alcanzado ya que se adapta a los datos existentes pero no fue tan sencillo como lo esperado. Esto era dar a los autores y a la plataforma que escoge los libros a publicar, información útil que sirva para mejorar su alcance.

En cuanto a los datos analizados, la conclusión final es la siguiente:

- No existe patrón entre la calificación media obtenida por los autores y libros y el número de ventas y reviews.
- Hay una pequeña tendencia en algunos autores más populares a tener mayor número de publicaciones.

- No hay una relación clara que indique que cuantas más publicaciones tenga un autor, más ventas tenga.
- La mayor parte de autores con mayor número de publicaciones tienen uno o dos géneros preferidos.
- En cuanto a las gráficas de tendencias en el tiempo, hay una clara tendencia al aumento de publicaciones que también puede deberse a que cada año la plataforma añade más libros a su biblioteca.
- En general los géneros mantienen su popularidad en el tiempo.
- Parece haber una relación en cuanto al número de ventas y el número de descargas y el número de reviews recibidos por los libros y sus géneros.
- No hay una relación clara entre la popularidad de un género y el número de libros que se publican de ese mismo género.

## BIBLIOGRAFÍA

1. Carpeta Drive:  
[https://drive.google.com/drive/folders/1uwJNVEZ--LP6GWGniB4o3lCVnkrfhE40?usp=drive\\_link](https://drive.google.com/drive/folders/1uwJNVEZ--LP6GWGniB4o3lCVnkrfhE40?usp=drive_link)
2. Código python datos sintéticos.
  - a. Nuevas columnas:  
[https://drive.google.com/file/d/1otxhsSeQK-wAvFRfQKwxf1fW7OmJmavw/view?usp=drive\\_link](https://drive.google.com/file/d/1otxhsSeQK-wAvFRfQKwxf1fW7OmJmavw/view?usp=drive_link)
  - b. Actualización diaria:  
[https://drive.google.com/file/d/1JpIdm7L1lh5JlDDrOozmWNcmYACMVx8/view?usp=drive\\_link](https://drive.google.com/file/d/1JpIdm7L1lh5JlDDrOozmWNcmYACMVx8/view?usp=drive_link)
3. Repositorio Github:  
<https://github.com/alu0101164403/GCCE-ProyectoKindle.git>