

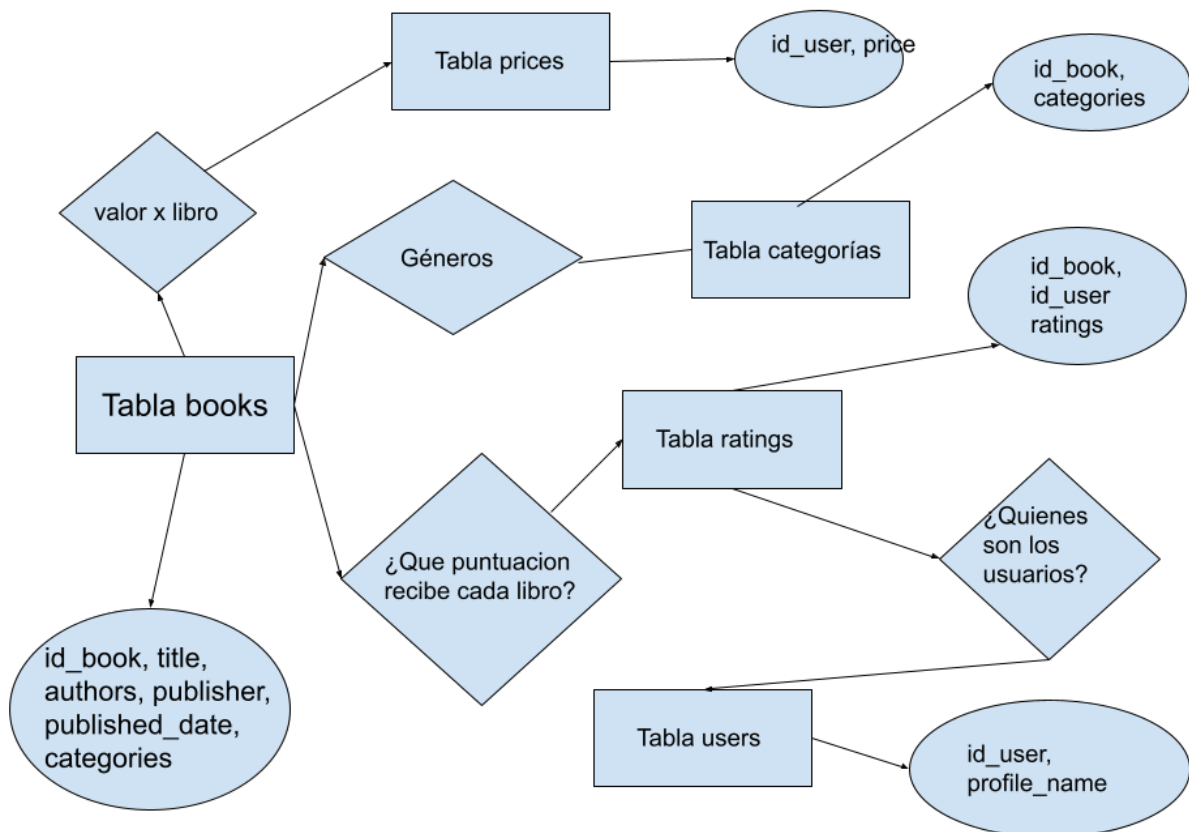
**Coderhouse****1° Entrega de proyecto final**

Base de datos elegida, y temática de la misma:

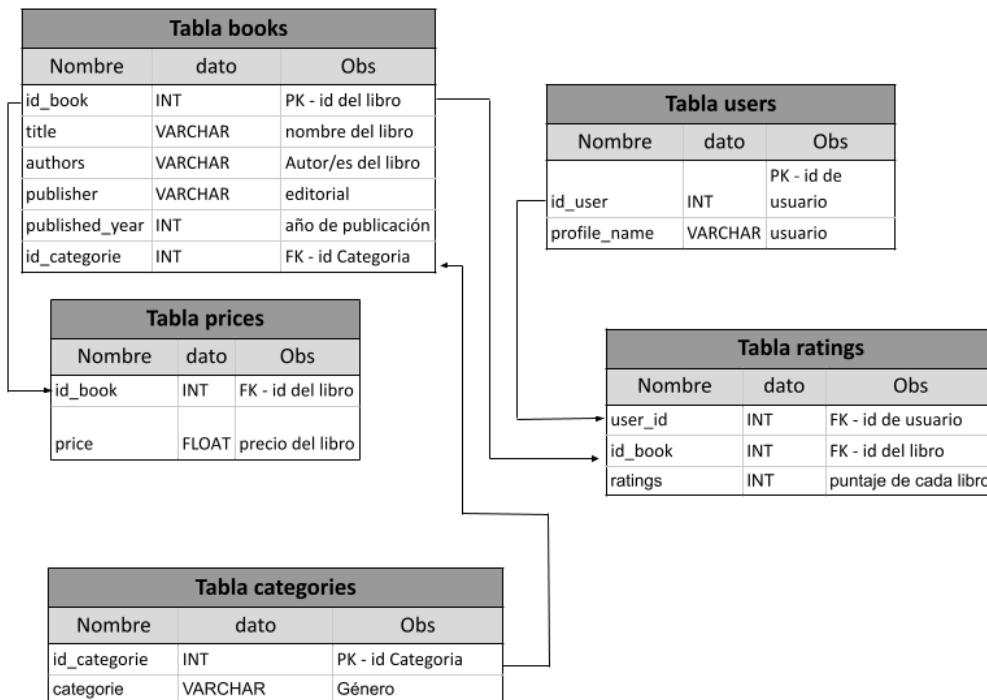
La BD elegida originalmente contiene una lista de libros, con diversos datos tales como el título, autor, año de publicación, entre otros.

El Dataset incluía una segunda tabla con 3 millones de puntuaciones de usuarios de la plataforma Amazon.

Debido al gran tamaño de la base de datos, opte por una reducción de la información, cuidando de no dejar datos huérfanos. También de pasar de 2 tablas, a 5, las cuales detallaré más adelante. Todo este proceso fue llevado a cabo a través de Python, usando la librería Pandas. El código está disponible en Github, así como el link al dataset que se encuentra alojado en la plataforma Kaggle.

**Diagrama entidad-relacion de la BD ya normalizada:**

### Descripción de las tablas



### Script de creación de Schema y Tablas

```
CREATE SCHEMA IF NOT EXISTS books;
```

```
USE books;
```

```
CREATE TABLE IF NOT EXISTS categories (
  id_categories INT NOT NULL PRIMARY KEY,
  categories VARCHAR (255)
);
```

```
CREATE TABLE IF NOT EXISTS books (
  id_book INT NOT NULL PRIMARY KEY,
  title VARCHAR (500),
  authors VARCHAR (500),
  publisher VARCHAR (500),
  published_year INT NOT NULL,
  id_categories INT NOT NULL, FOREIGN KEY (id_categories) REFERENCES categories
  (id_categories)
);
```

```
CREATE TABLE IF NOT EXISTS prices (  
id_book INT NOT NULL,  
FOREIGN KEY (id_book) REFERENCES books (id_book),  
price FLOAT NOT NULL  
);
```

```
CREATE TABLE IF NOT EXISTS users (  
id_user INT NOT NULL PRIMARY KEY,  
profile_name VARCHAR(255)  
);
```

```
CREATE TABLE IF NOT EXISTS ratings (  
id_user INT NOT NULL,  
FOREIGN KEY (id_user) REFERENCES users (id_user),  
id_book INT NOT NULL,  
FOREIGN KEY (id_book) REFERENCES books (id_book),  
score INT NOT NULL  
)
```

### **Script de inserción de filas:**

Aclaración: Teniendo en cuenta que las tablas tienen varias miles de filas, voy a dejar en este espacio una muestra de inserción de cada una de las tablas, que consta de 5 filas por tabla. En el repositorio de Github se puede encontrar el script completo.

Insert table categories:

```
use books;  
INSERT INTO categories (id_categories, categories) VALUES (1, 'Fiction');  
INSERT INTO categories (id_categories, categories) VALUES (2, 'Religion');  
INSERT INTO categories (id_categories, categories) VALUES (3, 'Biography &  
Autobiography');  
INSERT INTO categories (id_categories, categories) VALUES (4, 'Technology &  
Engineering');  
INSERT INTO categories (id_categories, categories) VALUES (5, 'Law');
```

Insert table books:

```
INSERT INTO books (id_book, title, authors, publisher, published_year, id_categories)  
VALUES (1, 'Whispers of the Wicked Saints', 'Veronica Haddon', 'iUniverse', 2005, 1);  
INSERT INTO books (id_book, title, authors, publisher, published_year, id_categories)  
VALUES (2, 'The Church of Christ: A Biblical Ecclesiology for Today', 'Everett Ferguson',  
'Wm. B. Eerdmans Publishing', 1996, 2);  
INSERT INTO books (id_book, title, authors, publisher, published_year, id_categories)  
VALUES (3, 'Saint Hyacinth of Poland', 'Mary Fabyan Windeatt', 'Tan Books & Pub', 2009, 3);  
INSERT INTO books (id_book, title, authors, publisher, published_year, id_categories)  
VALUES (4, 'Vector Quantization and Signal Compression (The Springer International Series
```

in Engineering and Computer Science)', 'Allen Gersho, Robert M. Gray', 'Springer Science & Business Media', 2012, 4);

Insert table prices:

```
INSERT INTO prices (id_book, price) VALUES (1, 10.95);
INSERT INTO prices (id_book, price) VALUES (2, 25.97);
INSERT INTO prices (id_book, price) VALUES (3, 13.95);
INSERT INTO prices (id_book, price) VALUES (4, 76.94);
INSERT INTO prices (id_book, price) VALUES (5, 14.95);
```

Insert Table Users:

```
INSERT INTO users (id_user, profile_name) VALUES (1, 'Book Reader');
INSERT INTO users (id_user, profile_name) VALUES (2, 'V. Powell');
INSERT INTO users (id_user, profile_name) VALUES (3, 'LoveToRead Actually Read Books');
INSERT INTO users (id_user, profile_name) VALUES (4, 'Clara');
INSERT INTO users (id_user, profile_name) VALUES (5, 'Tonya');
```

Insert table Ratings:

```
INSERT INTO ratings (id_user, id_book,score) VALUES (1, 1, 1);
INSERT INTO ratings (id_user, id_book,score) VALUES (2, 1, 4);
INSERT INTO ratings (id_user, id_book,score) VALUES (3, 1, 1);
INSERT INTO ratings (id_user, id_book,score) VALUES (4, 1, 5);
INSERT INTO ratings (id_user, id_book,score) VALUES (5, 1, 5);
```