



# Описание данных и задание SQL

## Таблица `books`

Содержит данные о книгах:

- `book_id` — идентификатор книги;
- `author_id` — идентификатор автора;
- `title` — название книги;
- `num_pages` — количество страниц;
- `publication_date` — дата публикации книги;
- `publisher_id` — идентификатор издателя.

## Таблица `authors`

Содержит данные об авторах:

- `author_id` — идентификатор автора;
- `author` — имя автора.

## Таблица `publishers`

Содержит данные об издательствах:

- `publisher_id` — идентификатор издательства;
- `publisher` — название издательства;

## Таблица `ratings`

Содержит данные о пользовательских оценках книг:

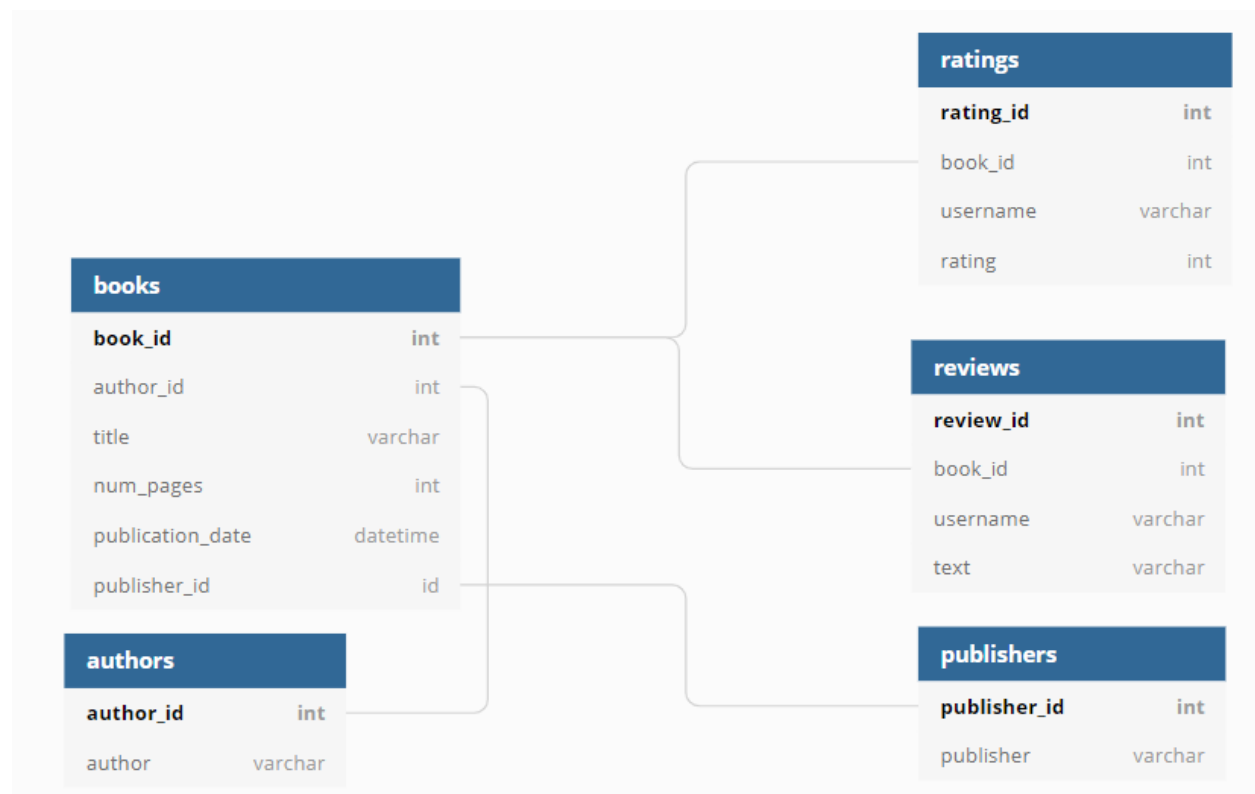
- `rating_id` — идентификатор оценки;

- `book_id` — идентификатор книги;
- `username` — имя пользователя, оставившего оценку;
- `rating` — оценка книги.

### Таблица `reviews`

Содержит данные о пользовательских обзорах на книги:

- `review_id` — идентификатор обзора;
- `book_id` — идентификатор книги;
- `username` — имя пользователя, написавшего обзор;
- `text` — текст обзора.



## Задания

- Посчитайте, сколько книг вышло после 1 января 2000 года;
- Для каждой книги посчитайте количество обзоров и среднюю оценку;

- Определите издательство, которое выпустило наибольшее число книг толще 50 страниц — так вы исключите из анализа брошюры;
- Определите автора с самой высокой средней оценкой книг — учитывайте только книги с 50 и более оценками;
- Посчитайте среднее количество обзоров от пользователей, которые поставили больше 48 оценок.

### **Как выполнить задание?**

- Опишите цели исследования;
- Исследуйте таблицы — выведите первые строки;
- Сделайте по одному SQL-запросу для решения каждого задания;
- Выведите результат каждого запроса в тетрадке;
- Опишите выводы по каждой из решённых задач.