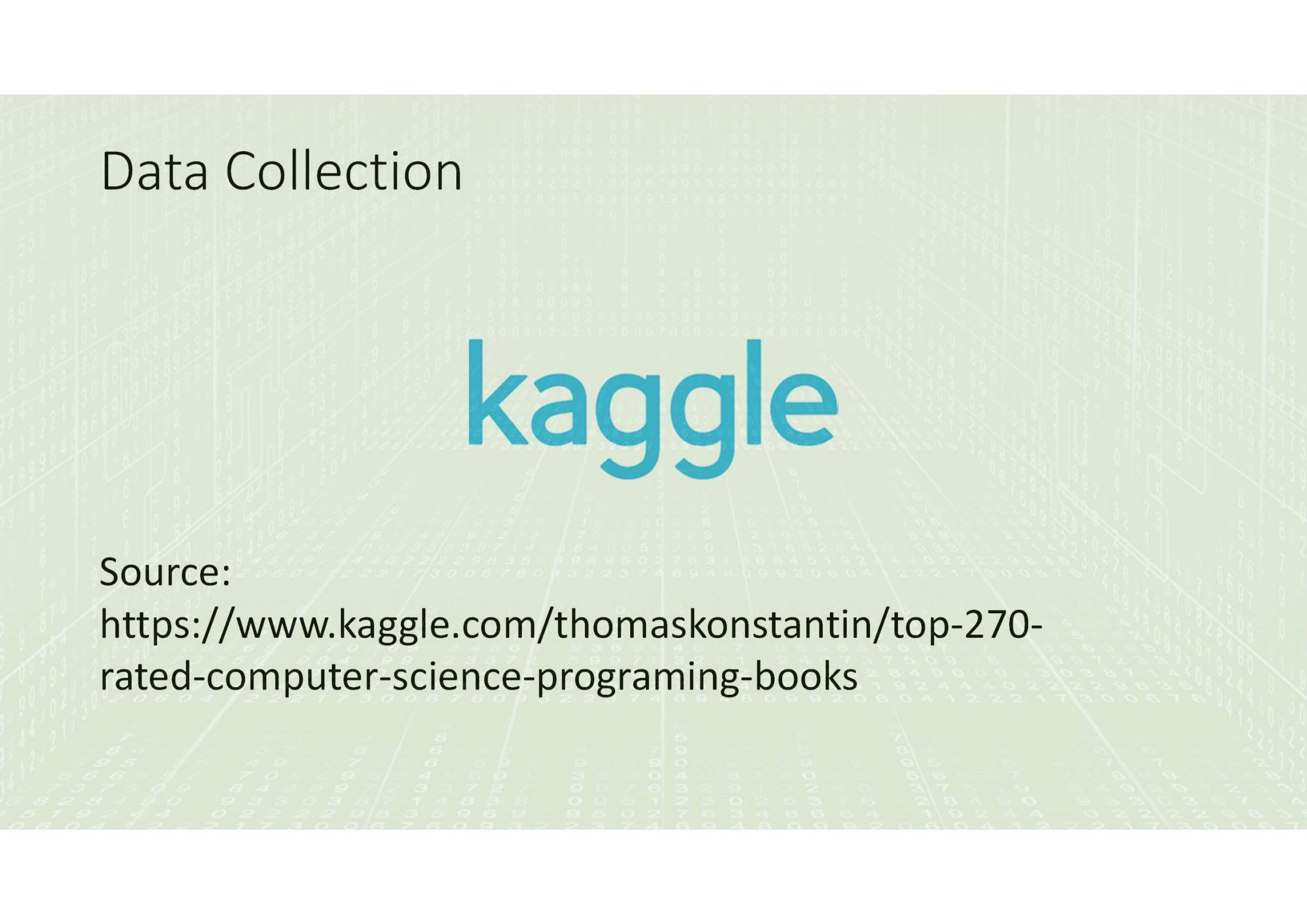


# Computer Science & Programming Books

# Question

As a Data Analyst (student), which book, within the computer science & programming category, should I read next ?

# Data Collection

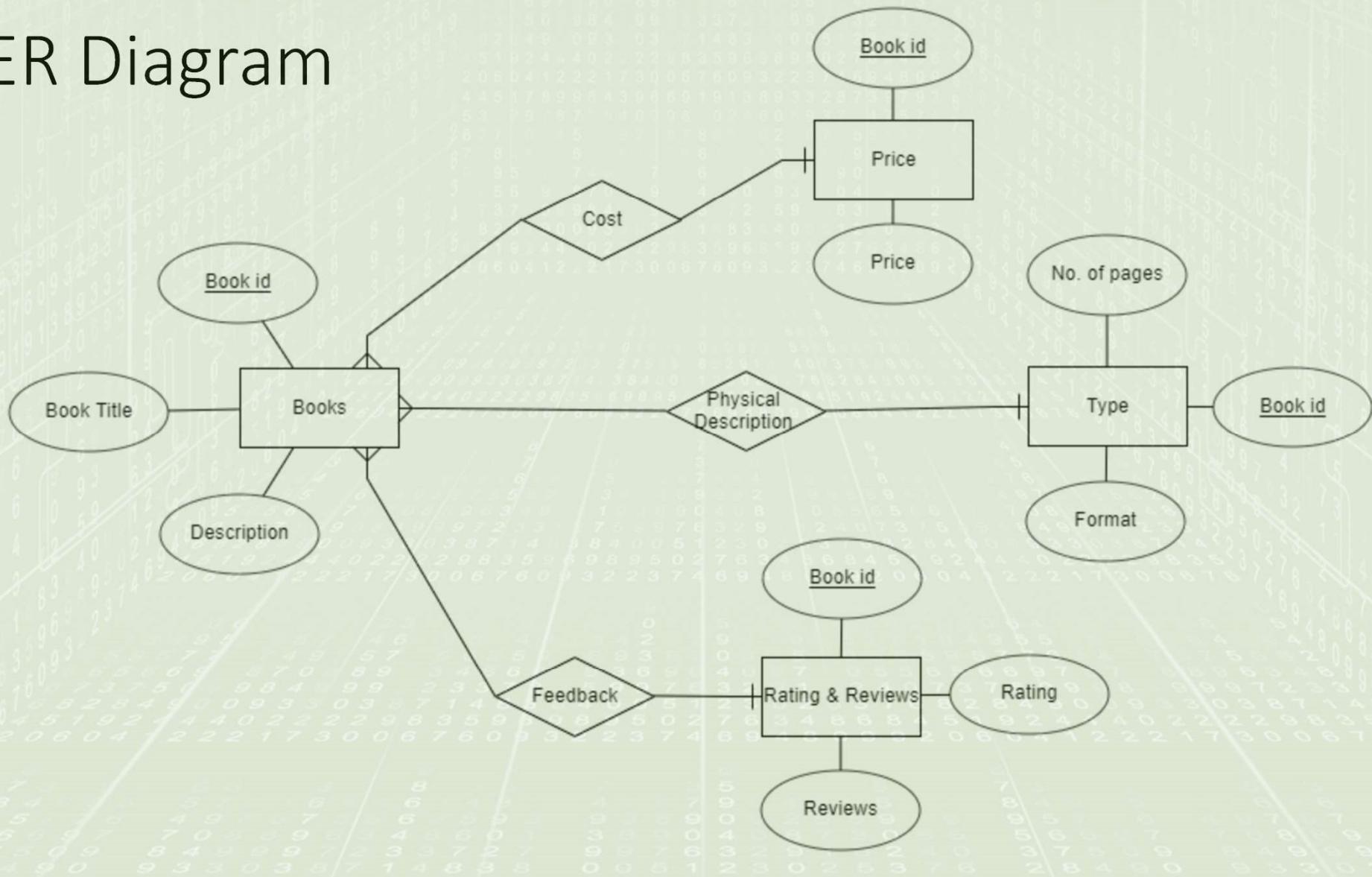


kaggle

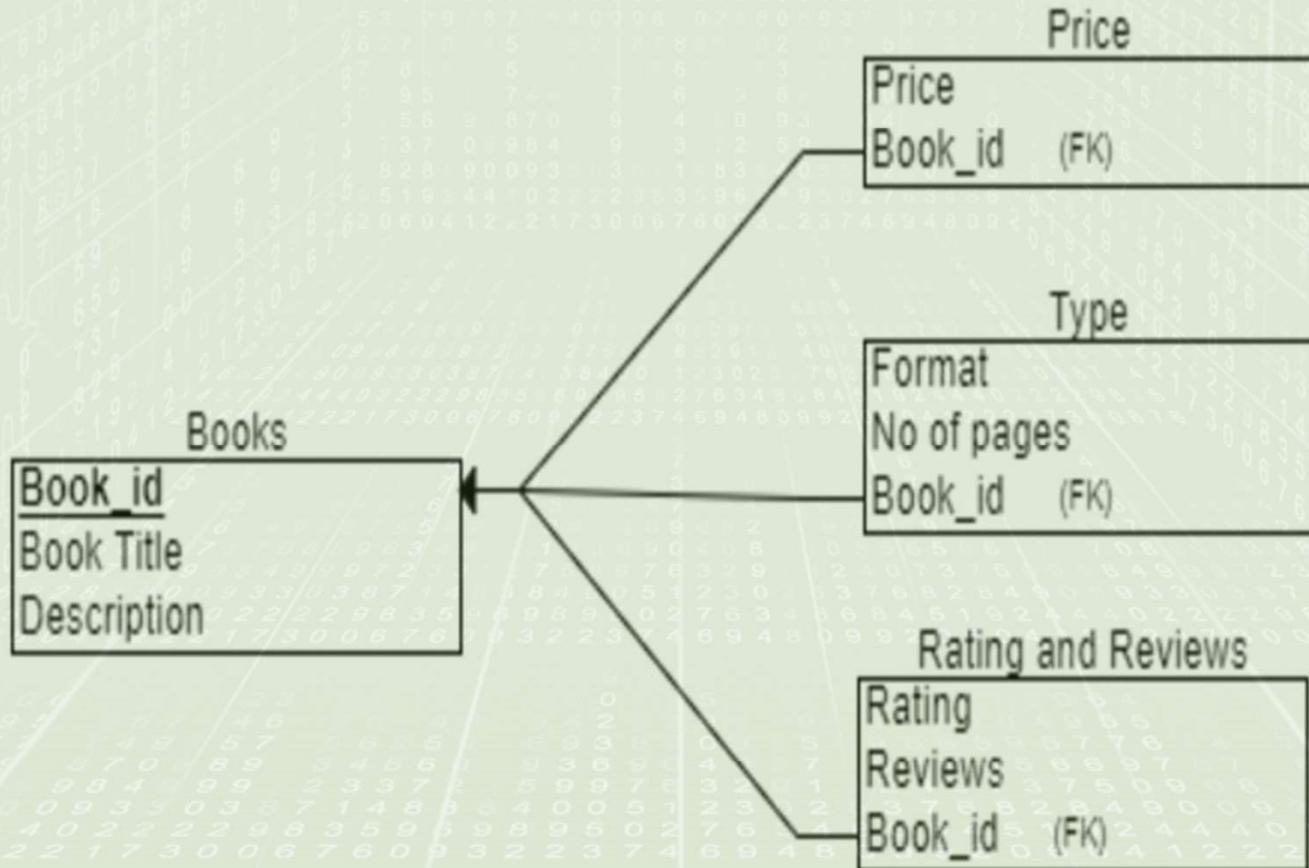
Source:

<https://www.kaggle.com/thomaskonstantin/top-270-rated-computer-science-programing-books>

# ER Diagram



# Relational Schema



# Data Preparation

```
---Round the Price column to 2 decimals
update Price
set Price = format (Price, 'N2')

---Alter the column Reviews from NVARCHAR to INT
alter table Rating_and_Reviews
alter column Reviews int;

---Remove books with no reviews
delete from Rating_and_Reviews
where Reviews = '0';

delete from Books
where Book_id in (3, 10, 25, 42, 45, 55, 65, 149, 178, 179, 183, 188, 200, 213, 227, 239, 240, 247, 262, 264, 265, 267);

delete from Price
where Book_id in (3, 10, 25, 42, 45, 55, 65, 149, 178, 179, 183, 188, 200, 213, 227, 239, 240, 247, 262, 264, 265, 267);

delete from Type
where Book_id in (3, 10, 25, 42, 45, 55, 65, 149, 178, 179, 183, 188, 200, 213, 227, 239, 240, 247, 262, 264, 265, 267);

---Remove the comma in Reviews that separates the thousands
update Rating_and_Reviews
set Reviews = replace(Reviews, ',', '');

---Combine ebook & Kindle
update Type
set Format='Kindle / ebook'
where Format='ebook';

update Type
set Format='Kindle / ebook'
where Format='Kindle Edition';

---Box set
update Type
set Format='Boxed Set'
where Book_id=269;

update Type
set Format='Boxed Set'
where Book_id=270;
```

# Data Analysis

```
--create view tables
create view View_All as
(select b.Book_id, b.Book_title, b.Description, p.Price, t.Number_of_Pages, t.Format, r.Rating, r.Reviews
from Books b
join Type t
on b.Book_id = t.Book_id
join Price p
on p.Book_id = b.Book_id
join Rating_and_Reviews r
on r.Book_id = b.Book_id);

--Average Rating & Average Reviews per Keyword in Book Title and Description
select avg(rating) as 'avg rating', avg(reviews) as 'avg reviews'
from View_All
where book_title like '%Game%' and Description like '%Game%';

--Quantity per book format
select Format, count(*) as 'Quantity'
from Type
group by Format
order by Count(*) desc;

--Average book rating
select round(avg(Rating),2) 'Average Rating'
from Rating_and_Reviews;

--Price in relation to format
select t.Format, avg(p.Price) as 'avg price'
from Books b
join Type t
on b.Book_id = t.Book_id
join Price p
on p.Book_id = b.Book_id
join Rating_and_Reviews r
on r.Book_id = b.Book_id
group by t.Format
order by 'avg price' desc;

--Finding the finalists
select *
from View_All
where
Description like '%Information%' and Description like '%Software%'
or
Description like '%Algorithm%' and Description like '%Information%'
or
Description like '%Algorithm%' and Description like '%Software%';
```

# Data Visualization & Insights

## Computer Science and Programming Books

Total  
**247**  
Books

Thickest  
**3168**  
Pages

Most  
**2629**  
Reviews

Top  
**4.77**  
Rating

Most Expensive  
**\$235.65**  
Price

Most common  
**Format**  
Paperback

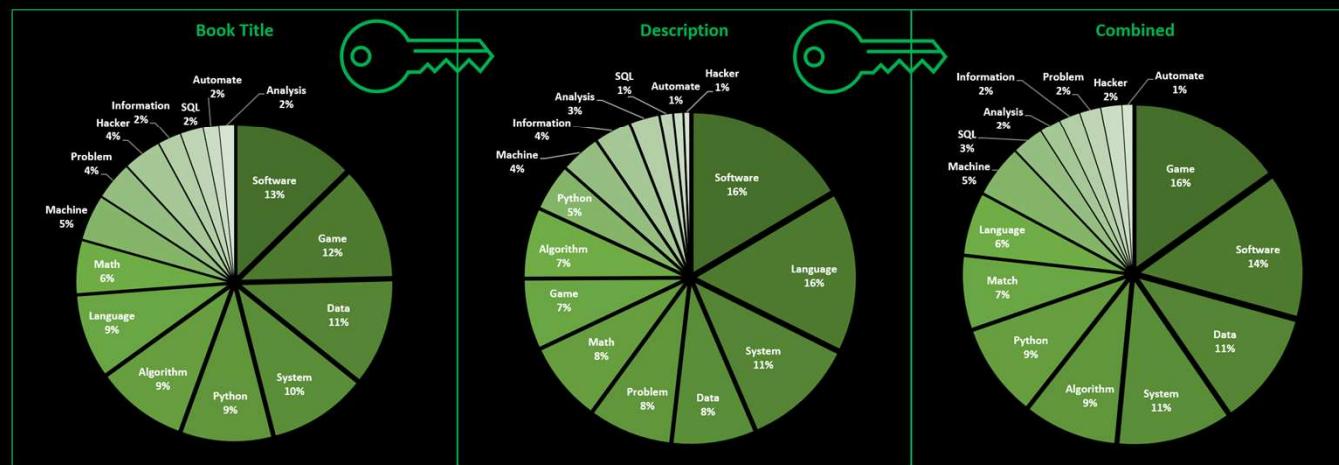
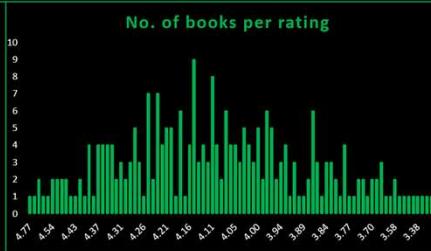
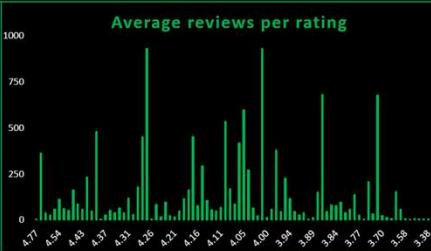
Format  
**Average Price**  
Boxed Set



Book Title
3D Game Engine Design: A Practical Approach to Real-Time Computer Graphics (The Morgan Kaufmann Series in Computer Graphics)
3D Math Primer for Graphics and Game Development
97 Things Every Programmer Should Know: Collective Wisdom from the Experts
A Discipline for Software Engineering
A Smarter Way to Learn JavaScript: The new approach that uses technology to cut your effort in half
A Tour of C++
Accelerate: Building and Scaling High-Performing Technology Organizations
Accelerated C++: Practical Programming by Example
Advanced Programming in the UNIX Environment
Adventures in Minecraft
Agile Retrospectives: Making Good Teams Great

Description
A fascinating exploration of how insights from computer algorithms can be applied to our everyday lives.

Book ID	Average Rating	Average Reviews
271	4.27	153
421	4.21	83
421	4.21	703
417	4.17	88
416	4.16	295
412	4.12	220
409	4.09	316
407	4.07	10
405	4.05	17
403	4.03	65
403	4.03	74
399	3.99	23
397	3.97	113
397	3.97	112
396	3.96	1037



Information	Algorithm	Software
Automate	Analysis	Software
Information	SQL	Game
Hacker	Machine	12%
Problem	Math	13%
Machine	Language	11%
Math	Algorithm	12%
Language	Python	10%
Algorithm	System	9%
Python		9%
System		10%



Information

Algorithm

Software



Results **244**

Rating **270**

Reviews **132**

# Conclusion

## Appears in most results

**Book ID: 244**

**Title:**

The Art of Computer Programming, Volume 1: Fundamental Algorithms

**Rating: 4.37**

**Reviews: 44**

**Price: \$93.46**

## The Finalists

### Best rating

**Book ID: 270**

**Title:**

The Art of Computer Programming, Volumes 1-4a. Boxed Set

**Rating: 4.77**

**Reviews: 4**

**Price: \$220.39**

### Most reviews

**Book ID: 132**

**Title:**

Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems

**Rating: 4.72**

**Reviews: 362**

**Price: \$45.56**

Q & A