

Data Mining Cup 2021

Session 1



TASK

Create a Recommendation Model for a Bookstore

- Provide the top-5 recommendations for **1,000 books**
- Use book descriptions (title, author, publisher, topics)
- Use existing transaction data (clicks, basket, orders)
- **Use external data!**

EVALUATION

Recommendations from Amazon

- **Our proposal:** Use book recommendations from Amazon for the evaluation of our approaches
- **Problem:** Amazon has good scraping protection, i.e. we have to gather the data manually (or we find a volunteer to invest some time into an automation of the process)
- **Task until next session:** Each team gathers book recommendations from Amazon for ~140 books (~17 books per student)
 - > enter them into [this google sheet](#)

Amazon Recommendations How-To

(1) Find book on Amazon

use amazon.de

use category „Bücher“

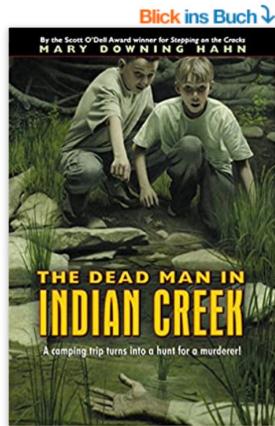
The screenshot shows the Amazon.de homepage with a search bar containing 'The Dead Man in Indian Creek'. A red box highlights the 'Bücher' dropdown menu. Below the search bar, the main navigation menu includes 'Alle', 'Bestseller', 'Amazon Basics', 'Angebote', 'Prime Video', 'Neuerscheinungen', 'Prime', 'Kundenservice', 'Bücher', 'Küche, Haushalt & Wohnen', 'Elektronik & Foto', 'amazon music', and '3 Monate gratis'. The left sidebar features sections for 'Berechtigt zum kostenfreien Versand', 'Kategorie' (with 'Bücher' selected), 'Durchschn. Kundenrezension' (with five star icons), and 'Preis' (with categories: 0 - 5 EUR, 5 - 10 EUR, 10 - 20 EUR, 20 - 50 EUR, Über 50 EUR). The main content area displays search results for 'The Dead Man in Indian Creek' by Mary Downing Hahn. The first result is highlighted with a red box around the title and author information. It shows a cover image of the book, the title 'Dead Man in Indian Creek', the subtitle 'Englisch Ausgabe | von Mary Downing Hahn', a price of 13,18€, and a rating of 4,86€. Other formats listed include Kindle (4,86€), Audible Hörbuch (0,00€), and Taschenbuch (2,50€). A note indicates that the item is not available in stock.



use private
browsing!

Amazon Recommendations How-To

(2) Locate Recommendations



The Dead Man in Indian Creek (Englisch) Taschenbuch – 9. Oktober 2001

von [Mary Downing Hahn](#) (Autor)

★★★★★ 76 Sternbewertungen

> Alle Formate und Ausgaben anzeigen

Kindle

4,86 €

Hörbuch

0,00 €

Gebundenes Buch

13,18 €

Taschenbuch

6,00 €

Lesen Sie mit unserer [kostenfreien App](#)

Gratis im Audible-Probemonat

4 Gebraucht ab 2,50 €

2 Neu ab 13,18 €

4 Gebraucht ab 5,40 €

2 Neu ab 6,00 €

When Parker Pettengill and Matt Armentrout find a dead man at the edge of Indian Creek, their camping trip comes to a close -- and a detective hunt begins. Evans, the creepy new boyfriend of Parker's mom, Pam, was seen at -- the place of the crime -- but there's no solid evidence that he committed the murder. Parker is convinced that Pam's boyfriend is to blame, and he will not rest until he has proof. And when Parker and Armentrout hide out in Pam's antique store and spy on Pam and Evans, they find out more than they bargained for....

Lesealter



8 - 12 Jahre

Seitenzahl der ...



144 Seiten

Sprache



Englisch

Schuljahr



3 - 7

Lexile-Bewertu...



820L

Abmessungen



13.11 x 1.02 x

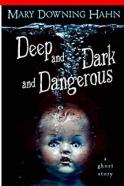
19.08 cm



Beliebte Taschenbuch-Empfehlungen des Monats

Stöbern Sie jetzt durch unsere Auswahl beliebter Bücher aus verschiedenen Genres wie Krimi, Thriller, historische Romane oder Liebesromane [Hier stöbern](#)

Kunden, die diesen Artikel gekauft haben, kauften auch



Deep and Dark and

Dangerous: A Ghost

Story

Mary Downing Hahn

Find recommendations under this heading!

If you can't find the box, try a different book version.

If you still can't find the box, look for recommendations with other headings.

Amazon Recommendations How-To

(3) Resolve Recommendations

DMC2021_Amazon-Recommendations																		
TEAM	itemID	title	author	rec1_title	rec1_ID	rec2_title	rec2_ID	rec3_title	rec3_ID	rec4_title	rec4_ID	rec5_title	rec5_ID	rec6_title	rec6_ID	rec7_title	rec7_ID	comment
Data Minerz	12	Breathtaking	Reva Ryann Thompson															
	45274	The Dead Man in Indian Creek	Mary Downing Hahn															
	10104	The Humble Seed	Linda Appleby															
	41371	Dryadenhain & Dschinnenzauber (Märchenanalogie)	Philipp Busch, Jonathan Driedger, Lena Kalupner, C															
	14015	The Man of Steel: Superman vs. the Moon Bandits	Scott Sonneborn															
	28515	Singing Innocence and Experience	Sonya Taaffe															
	70766	Ein Lied, mein Leben und was sonst noch schiefgehen kann	Tini Wider															
	1822	School of Gitty																

First add recommendation titles in the respective columns (collect up to 7).

Then use the **items.csv** file to find the respective item ID (if available) and add it to the sheet.

SUMMING UP

Your tasks for Session 2 (May, 4th)

- Acquire some knowledge about Recommender Systems
 - A good starting point is [this talk](#)
- Collect the recommendations from Amazon
- Develop a first baseline approach
 - E.g. start with something simple like recommending books with the same author
 - and/or recommend sequels (if available)
 - and/or use the transactions
 - and/or ...
- Think about suitable sources for external data