

# Data cleaning and preprocessing :

Data cleaning and preprocessing are crucial steps in preparing data for analysis or modeling. This involves tasks like handling missing values, removing duplicates, transforming data types, scaling features

Unique Harry Potter Books:

["Harry Potter and the Philosopher's Stone"

'Harry Potter and the Prisoner of Azkaban'

'Harry Potter and the Order of the Phoenix'

'Harry Potter and the Chamber of Secrets'

'Harry Potter and the Goblet of Fire'

'Harry Potter and the Deathly Hallows'

'Harry Potter and the Half-Blood Prince'

'Complete Harry Potter Boxed Set'

'Harry Potter Collection (Harry Potter, #1-6)'

'The Magical Worlds of Harry Potter: A Treasury of Myths, Legends, and Fascinating Facts']

## #Missing values:

Handling missing values is essential for ensuring the quality and reliability of your data

```

Missing Values:
book_id                0
goodreads_book_id     0
best_book_id          0
work_id               0
books_count           0
isbn                  0
isbn13                0
authors               0
original_publication_year 0
original_title        0
title                 0
language_code         0
average_rating         0
ratings_count         0
work_ratings_count    0
work_text_reviews_count 0
ratings_1             0
ratings_2             0
ratings_3             0
ratings_4             0
ratings_5             0
image_url             0
small_image_url       0
dtype: int64

```

## The most selling books :

Most Selling Books in Harry Potter Series:

original\_title

Harry Potter and the Philosopher's Stone	4602479
Harry Potter and the Prisoner of Azkaban	1832823
Harry Potter and the Chamber of Secrets	1779331
Harry Potter and the Goblet of Fire	1753043
Harry Potter and the Deathly Hallows	1746574

Name: ratings\_count, dtype: int64

## The average rating :

---

Average rating of the Harry Potter books: 4.49