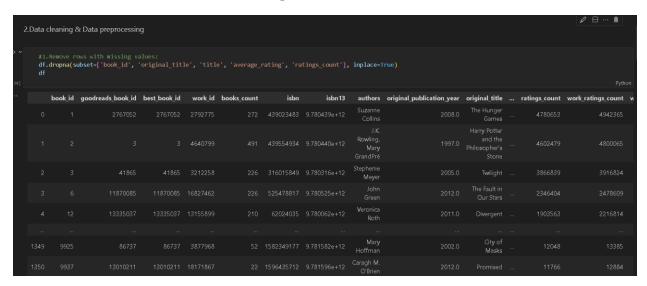
1. Load the data set

```
import pandas as pd
    # Load the dataset

f = pd.read_csv('books.csv')
    print(df.columns.tolist())

Python
['book_id', 'goodreads_book_id', 'best_book_id', 'work_id', 'books_count', 'isbn', 'isbn13', 'authors', 'original_publication_year', 'original_title', 'title', 'language_code',
```

- 2. Data cleaning & data processing
- 2.1. remove rows with missing values.



2.2. Convert 'ratings_count' to numeric (replace non-numeric values with NaN):

2.3. Convert 'Rating' to numeric:

3. Filter for Harry Potter Series:

harry_potter_books Python													
b	ook_id	goodreads_book_id	best_book_id	work_id	books_count	isbn	isbn 13		original_publication_year			ratings_count	work_ratings_count
				4640799		439554934	9.780440e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Philosopher's Stone			
							9.780440e+12	J.K. Rowling, Mary GrandPré, Rufus Beck		Harry Potter and the Prisoner of Azkaban			
							9.780439e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Order of the Phoenix			1840548
								J.K. Rowling, Mary GrandPré		Harr y Potter and the Chamber of Secrets			
				3046572				J.K. Rowling, Mary GrandPré		Harr y Potter and the Goblet of Fire			1868642
11	25	136251	136251	2963218	263	545010225	9.780545e+12	J.K. Rowling, Mary	2007.0	Harry Potter and the Deathly		1746574	1847395
11	25	136251	136251	2963218	263	545010225	9.7805 4 5e+12	J.K. Rowling, Mary GrandPré	2007.0	Harry Potter and the Deathly Hallows		1746574	1847395
							9.780440e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Half- Blood Prince			
						545044251	9.780545e+12	J.K. Rowling		Complete Harry Potter Boxed Set			
							9.780440e+12	J.K. Rowling		Harry Potter Collection (Harry Potter, #1-6)			
		483445	483445					Da v id Colbert		The Magical Worlds of Harry Potter: A Treasury			

4. Find the Most selling Harry Potter Book:

<pre>most_selling_hp = harry_potter_books.nlargest(10, 'ratings_count') most_selling_hp Pribon</pre>												
	book_id	goodreads_book_id	best_book_id	work_id	books_count	isbn	isbn13	authors	original_publication_year	original_title	ratings_count	work_ratings_count
				4640799		439554934	9.780440e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Philosopher's Stone		
						043965548 X	9.780440e+12	J.K. Rowling, Mary Grand Pré, Rufus Beck		Harry Potter and the Prisoner of Azkaban		
						439064864	9.780439e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Chamber of Secrets		
								J.K. Rowling, Mary GrandPré		Harry Potter and the Goblet of Fire		
							9.780545e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Deathly Hallows		

4. the rest of the table

11	25	136251	136251	2963218	263	545010225	9.780545e+12	J.K. Rowling, Mary GrandPré	2007.0	Harry Potter and the Deathly Hallows	1746574	1847395
8								J.K. Rowling, Mary GrandPré		Harry Potter and the Order of the Phoenix		1840548
12				41335427			9.780440e+12	J.K. Rowling, Mary GrandPré		Harry Potter and the Half- Blood Prince		1785676
96						545044251	9.780545e+12	J.K. Rowling		Complete Harry Potter Boxed Set		204125
613							9.780440e+12	J.K. Rowling		Harry Potter Collection (Harry Potter, #1-6)		26274
1036		483445	483445				9.780425e+12	David Colbert		The Magical Worlds of Harry Potter: A Treasury		15145
10 rows ×												

5. Calculate the Average Rating of Harry Potter Books:

```
5.Calculate the Average Rating of Harry Potter Books:

average_rating_hp = harry_potter_books['average_rating'].mean()
print(average_rating_hp)

4.491000000000000005
```

6. Display the Results:

```
print("Most selling Harry Potter book:")
print(most_selling.hp[['original_title', 'average_rating']])
print(f'\naverage rating of Harry Potter books: {average_rating.hp:.2f}")

**Most selling Harry Potter book:

**Original_title original_title or
```

7. Save the Cleaned and Processed Dataset:

```
7.Save the Cleaned and Processed Dataset:

df.to_csv('cleaned_books.csv', index=False)
print("\nCleaned_dataset_saved_as 'cleaned_books.csv'.")

Python

Cleaned_dataset_saved_as 'cleaned_books.csv'.
```