- Focus your analysis on the Harry Potter book series.

```
In [5]: Harry_Potter = df[df['title'].str.contains('Harry Potter', case=False)]
        Harry_Potter
```

Out[5]:

books_count	isbn	isbn13	authors	original_publication_year	original_title	...	ratings_count	work_ratings_count	work_text_reviews_count	ratin
491	439554934	9.780440e+12	J.K. Rowling, Mary GrandPré	1997.0	Harry Potter and the Philosopher's Stone	...	4602479	4800065	75867	7
376	043965548X	9.780440e+12	J.K. Rowling, Mary GrandPré, Rufus Beck	1999.0	Harry Potter and the Prisoner of Azkaban	...	1832823	1969375	36099	
307	439358078	9.780439e+12	J.K. Rowling, Mary GrandPré	2003.0	Harry Potter and the Order of the Phoenix	...	1735368	1840548	28685	
398	439064864	9.780439e+12	J.K. Rowling, Mary GrandPré	1998.0	Harry Potter and the Chamber of Secrets	...	1779331	1906199	34172	
332	439139600	9.780439e+12	J.K. Rowling, Mary GrandPré	2000.0	Harry Potter and the Goblet of Fire	...	1753043	1868642	31084	
263	545010225	9.780545e+12	J.K. Rowling, Mary GrandPré	2007.0	Harry Potter and the Deathly Hallows	...	1746574	1847395	51942	
275	439785960	9.780440e+12	J.K. Rowling, Mary GrandPré	2005.0	Harry Potter and the Half-Blood Prince	...	1678823	1785676	27520	
76	545044251	9.780545e+12	J.K. Rowling	1998.0	Complete Harry Potter Boxed Set	...	190050	204125	6508	
6	439827604	9.780440e+12	J.K. Rowling	2005.0	Harry Potter Collection (Harry Potter	...	24618	26274	882	

```
In [6]: #detect the importatnt columns for analysis
        columns_to_keep = ['books_count','average_rating','ratings_count','work_ratings_count','work_text_reviews_count']
        columns_to_drop = df.columns.difference(columns_to_keep)

        # Drop all columns except the specified ones
        Harry_Potter = Harry_Potter.drop(columns=columns_to_drop)
        Harry_Potter
```

Out[6]:

	books_count	average_rating	ratings_count	work_ratings_count	work_text_reviews_count
1	491	4.44	4602479	4800065	75867
6	376	4.53	1832823	1969375	36099
8	307	4.46	1735368	1840548	28685
9	398	4.37	1779331	1906199	34172
10	332	4.53	1753043	1868642	31084
11	263	4.61	1746574	1847395	51942
12	275	4.54	1678823	1785676	27520
96	76	4.74	190050	204125	6508
613	6	4.73	24618	26274	882
1036	42	3.96	13820	15145	267
1266	5	4.40	10736	11732	185

- Find the most selling books within the Harry Potter series.

```
In [9]: #To find the most selling books from the dataset,
        #we can use the work_ratings_count field, which represents the total number of ratings for the work.
        #We can sort the dataset based on this field in descending order to find the books with the highest number of ratings
        the_most_selling_books = Harry_Potter.sort_values(by='work_ratings_count', ascending=False)
        the_most_selling_books
```

Out[9]:

	books_count	average_rating	ratings_count	work_ratings_count	work_text_reviews_count
1	491	4.44	4602479	4800065	75867
6	376	4.53	1832823	1969375	36099
9	398	4.37	1779331	1906199	34172
10	332	4.53	1753043	1868642	31084
11	263	4.61	1746574	1847395	51942
8	307	4.46	1735368	1840548	28685
12	275	4.54	1678823	1785676	27520
96	76	4.74	190050	204125	6508
613	6	4.73	24618	26274	882
1036	42	3.96	13820	15145	267
1266	5	4.40	10736	11732	185

- Calculate the average rating of the Harry Potter books.

```
In [10]: avg_rate=the_most_selling_books['average_rating'].mean()
         avg_rate
```

Out[10]: 4.482727272727273