

```
In [1]: import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [2]: data=pd.read_csv('books.csv')
print(data.describe())
```

|       | book_id   | goodreads_book_id | best_book_id | work_id      | books_count \ |
|-------|-----------|-------------------|--------------|--------------|---------------|
| count | 99.000000 | 9.900000e+01      | 9.900000e+01 | 9.900000e+01 | 99.000000     |
| mean  | 50.000000 | 1.653979e+06      | 2.020180e+06 | 4.388198e+06 | 532.383838    |
| std   | 28.722813 | 4.030956e+06      | 4.681734e+06 | 6.465936e+06 | 700.853630    |
| min   | 1.000000  | 1.000000e+00      | 1.000000e+00 | 5.397000e+03 | 14.000000     |
| 25%   | 25.500000 | 3.925000e+03      | 4.297500e+03 | 1.681676e+06 | 172.500000    |
| 50%   | 50.000000 | 1.813500e+04      | 1.906300e+04 | 3.036731e+06 | 226.000000    |
| 75%   | 74.500000 | 1.927180e+05      | 3.135880e+05 | 3.357144e+06 | 480.500000    |
| max   | 99.000000 | 2.255727e+07      | 2.255727e+07 | 4.133543e+07 | 3455.000000   |

|       | isbn13       | original_publication_year | average_rating | ratings_count \ |
|-------|--------------|---------------------------|----------------|-----------------|
| count | 9.900000e+01 | 99.000000                 | 99.000000      | 9.900000e+01    |
| mean  | 9.780473e+12 | 1943.171717               | 4.055051       | 1.263703e+06    |
| std   | 3.961477e+08 | 277.288365                | 0.245059       | 7.832598e+05    |
| min   | 9.780007e+12 | -720.000000               | 3.510000       | 3.872900e+05    |
| 25%   | 9.780150e+12 | 1952.500000               | 3.870000       | 7.388255e+05    |
| 50%   | 9.780385e+12 | 1997.000000               | 4.060000       | 1.053403e+06    |
| 75%   | 9.780553e+12 | 2005.000000               | 4.245000       | 1.653671e+06    |
| max   | 9.781612e+12 | 2015.000000               | 4.610000       | 4.780653e+06    |

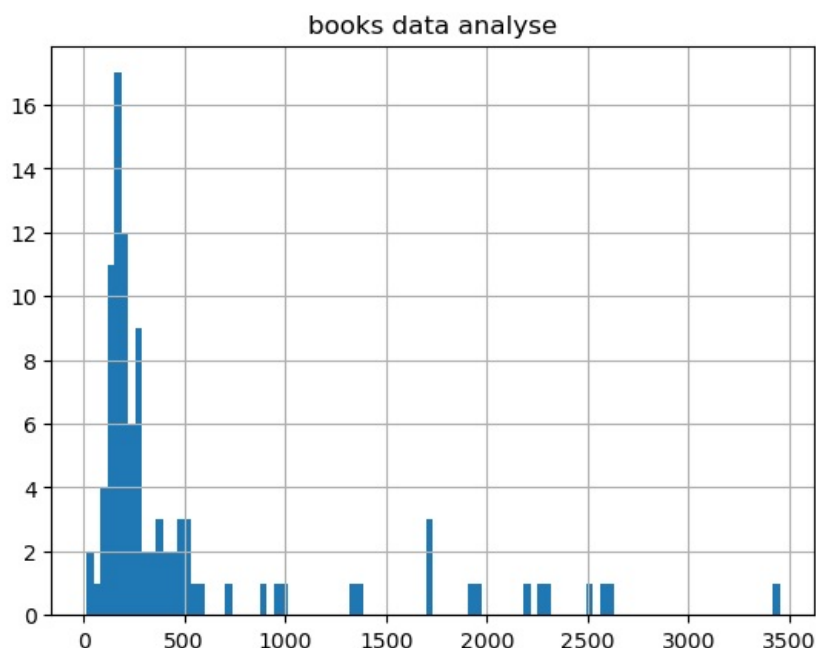
  

|       | work_ratings_count | work_text_reviews_count | ratings_1 \   |
|-------|--------------------|-------------------------|---------------|
| count | 9.900000e+01       | 99.000000               | 99.000000     |
| mean  | 1.361989e+06       | 39012.222222            | 39065.565657  |
| std   | 7.978995e+05       | 28693.522227            | 50901.940472  |
| min   | 5.493010e+05       | 4239.000000             | 4623.000000   |
| 25%   | 8.051315e+05       | 17789.500000            | 15337.500000  |
| 50%   | 1.125231e+06       | 31212.000000            | 27340.000000  |
| 75%   | 1.729282e+06       | 48035.000000            | 45624.000000  |
| max   | 4.942365e+06       | 155254.000000           | 456191.000000 |

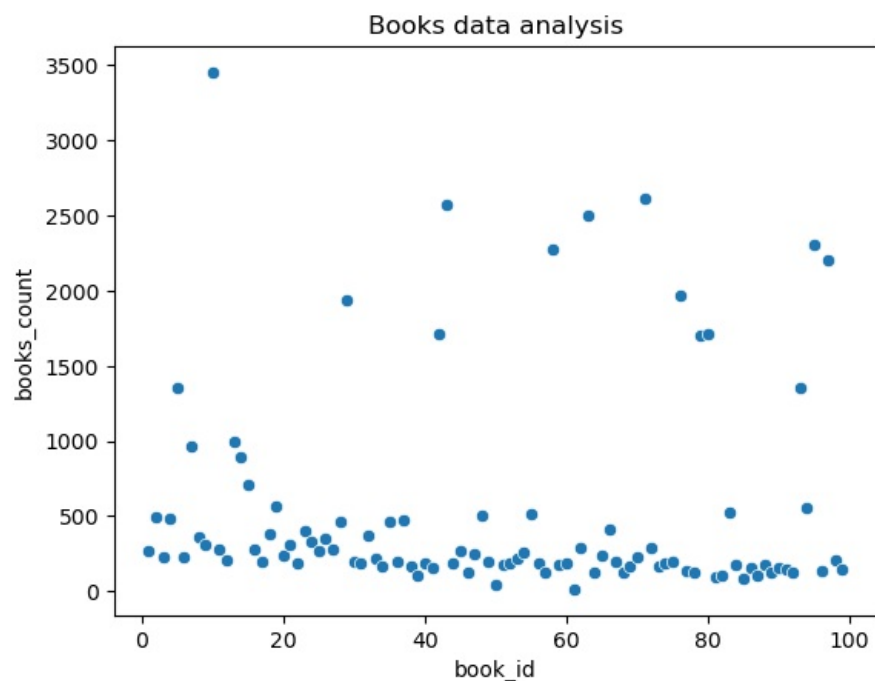
  

|       | ratings_2     | ratings_3     | ratings_4    | ratings_5    |
|-------|---------------|---------------|--------------|--------------|
| count | 99.000000     | 99.000000     | 9.900000e+01 | 9.900000e+01 |
| mean  | 70538.828283  | 228209.626263 | 4.204259e+05 | 6.037495e+05 |
| std   | 55694.613546  | 124117.176087 | 2.231151e+05 | 4.628464e+05 |
| min   | 15781.000000  | 76071.000000  | 1.403040e+05 | 1.760720e+05 |
| 25%   | 35029.500000  | 138675.500000 | 2.635695e+05 | 3.130130e+05 |
| 50%   | 58323.000000  | 200154.000000 | 3.475400e+05 | 4.628290e+05 |
| 75%   | 86238.000000  | 277074.000000 | 5.184895e+05 | 7.283425e+05 |
| max   | 436802.000000 | 793319.000000 | 1.481305e+06 | 3.011543e+06 |

```
In [9]: data.head(10)
data['books_count'].hist(bins=100)
plt.title("books data analyse")
plt.show()
```



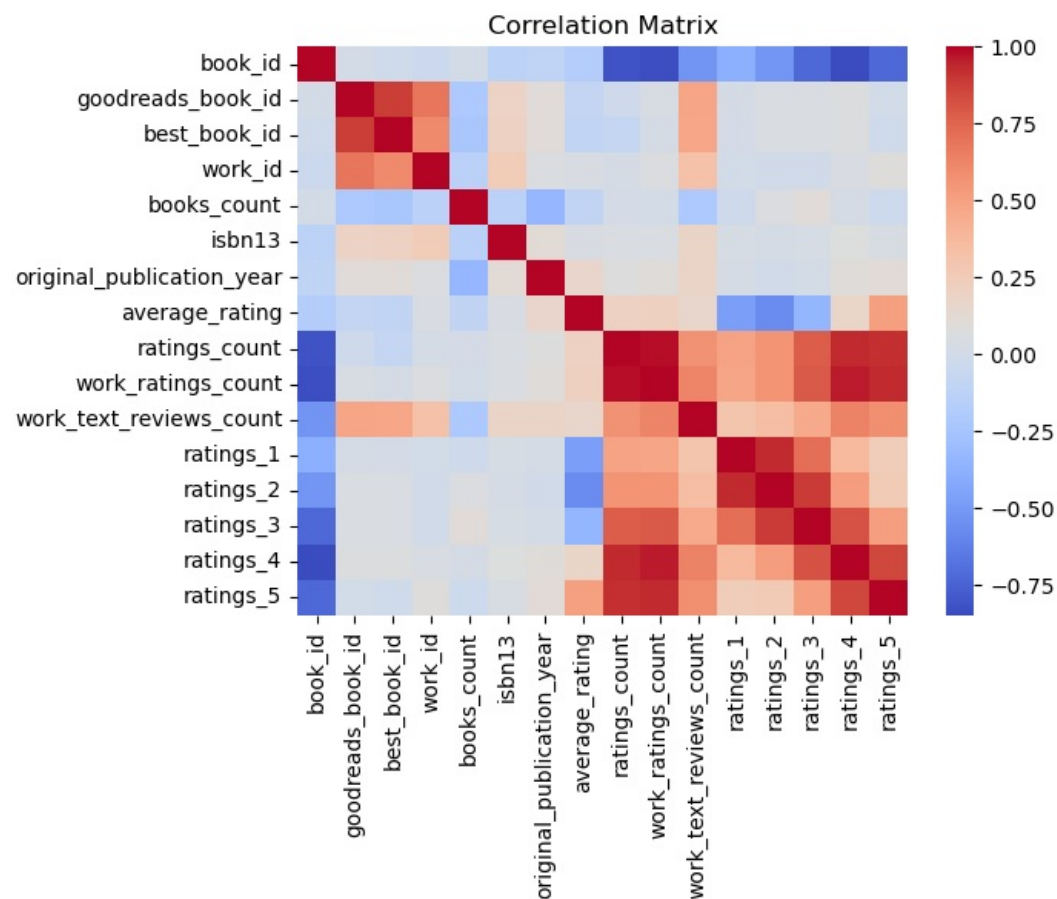
```
In [4]: sns.scatterplot(x='book_id', y='books_count', data=data)
plt.title('Books data analysis')
plt.show()
```



```
In [14]: sns.heatmap(data.corr(), annot=False, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

C:\Users\DELL\AppData\Local\Temp\ipykernel\_11536\1628057676.py:1: FutureWarning: The default value of numeric\_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric\_only to silence this warning.

```
sns.heatmap(data.corr(), annot=False, cmap='coolwarm')
```



```
In [ ]:
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js