# moodshelf

May 24, 2025

## 1 MoodShelf

```python
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

pd.set_option("display.max_columns", None)

df = pd.read_csv("./goodreads-data-50k.csv")
df.head()
```

```
[1]:                                         bookId  \
     0                    2767052-the-hunger-games
     1  2.Harry_Potter_and_the_Order_of_the_Phoenix
     2                   2657.To_Kill_a_Mockingbird
     3                     1885.Pride_and_Prejudice
     4                              41865.Twilight

                                     title                   series  \
     0                        The Hunger Games    The Hunger Games #1
     1  Harry Potter and the Order of the Phoenix        Harry Potter #5
     2                    To Kill a Mockingbird  To Kill a Mockingbird
     3                      Pride and Prejudice                    NaN
     4                               Twilight    The Twilight Saga #1

                                     author  rating  \
     0                        Suzanne Collins    4.33
     1  J.K. Rowling, Mary GrandPré (Illustrator)    4.50
     2                            Harper Lee    4.28
     3  Jane Austen, Anna Quindlen (Introduction)    4.26
     4                        Stephenie Meyer    3.60

                                     description language          isbn  \
     0  WINNING MEANS FAME AND FORTUNE.LOSING MEANS CE…  English  9780439023481
     1  There is a door at the end of a silent corrido…  English  9780439358071
     2  The unforgettable novel of a childhood in a sl…  English  9999999999999
     3  Alternate cover edition of ISBN 9780679783268S…  English  9999999999999
```

```
4   About three things I was absolutely positive.\…   English   9780316015844

                                        genres  \
0   ['Young Adult', 'Fiction', 'Dystopia', 'Fantas…
1   ['Fantasy', 'Young Adult', 'Fiction', 'Magic',…
2   ['Classics', 'Fiction', 'Historical Fiction', …
3   ['Classics', 'Fiction', 'Romance', 'Historical…
4   ['Young Adult', 'Fantasy', 'Romance', 'Vampire…

                                     characters bookFormat  \
0   ['Katniss Everdeen', 'Peeta Mellark', 'Cato (H…  Hardcover
1   ['Sirius Black', 'Draco Malfoy', 'Ron Weasley'…  Paperback
2   ['Scout Finch', 'Atticus Finch', 'Jem Finch', …  Paperback
3   ['Mr. Bennet', 'Mrs. Bennet', 'Jane Bennet', '…  Paperback
4   ['Edward Cullen', 'Jacob Black', 'Laurent', 'R…  Paperback

                            edition pages                    publisher  \
0                     First Edition   374              Scholastic Press
1                        US Edition   870               Scholastic Inc.
2                               NaN   324  Harper Perennial Modern Classics
3   Modern Library Classics, USA / CAN   279               Modern Library
4                               NaN   501       Little, Brown and Company

  publishDate firstPublishDate  \
0    09/14/08              NaN
1    09/28/04         06/21/03
2    05/23/06         07/11/60
3    10/10/00         01/28/13
4    09/06/06         10/05/05

                                        awards  numRatings  \
0   ['Locus Award Nominee for Best Young Adult Boo…    6376780
1   ['Bram Stoker Award for Works for Young Reader…    2507623
2   ['Pulitzer Prize for Fiction (1961)', 'Audie A…    4501075
3                                            []    2998241
4   ['Georgia Peach Book Award (2007)', 'Buxtehude…    4964519

                                ratingsByStars  likedPercent  \
0   ['3444695', '1921313', '745221', '171994', '93…          96.0
1   ['1593642', '637516', '222366', '39573', '14526']          98.0
2   ['2363896', '1333153', '573280', '149952', '80…          95.0
3   ['1617567', '816659', '373311', '113934', '767…          94.0
4   ['1751460', '1113682', '1008686', '542017', '5…          78.0

                                        setting  \
0   ['District 12, Panem', 'Capitol, Panem', 'Pane…
1   ['Hogwarts School of Witchcraft and Wizardry (…
```

```
2                  ['Maycomb, Alabama (United States)']
3  ['United Kingdom', 'Derbyshire, England (Unite…
4  ['Forks, Washington (United States)', 'Phoenix…


                                          coverImg  bbeScore  bbeVotes  \
0  https://i.gr-assets.com/images/S/compressed.ph…   2993816     30516
1  https://i.gr-assets.com/images/S/compressed.ph…   2632233     26923
2  https://i.gr-assets.com/images/S/compressed.ph…   2269402     23328
3  https://i.gr-assets.com/images/S/compressed.ph…   1983116     20452
4  https://i.gr-assets.com/images/S/compressed.ph…   1459448     14874


   price  word_count                                          input_ids  \
0  5.09          147  [  101    160 11607 27451 11780 22157 14962   17…
1  7.38          136  [  101   1247  1110   170   1442  1120  1103  13…
2   NaN          137  [  101   1109  8362 14467 12272  5100  2165  22…
3   NaN          100  [  101 14234  2267  2596  1104  2282  4087  15…
4   2.1           51  [  101   3517  1210  1614   146  1108  7284  31…


                                     token_type_ids  \
0  [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0…
1  [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0…
2  [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0…
3  [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0…
4  [0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0…


                                     attention_mask
0  [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1…
1  [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1…
2  [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1…
3  [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1…
4  [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1…
```

## 1.1 Explanatory Data Analysis

```python
[2]: cols_to_keep = [
         "title",
         "author",
         "rating",
         "description",
         "language",
         "genres",
         "pages",
         "publisher",
         "publishDate",
         "numRatings",
         "likedPercent",
         "coverImg",
```

```
]

df_subset = df[cols_to_keep]
df_subset.head()
```

[2]:                                                title  \
0                                    The Hunger Games
1                  Harry Potter and the Order of the Phoenix
2                              To Kill a Mockingbird
3                                Pride and Prejudice
4                                           Twilight


                                            author  rating  \
0                           Suzanne Collins    4.33
1      J.K. Rowling, Mary GrandPré (Illustrator)    4.50
2                                Harper Lee    4.28
3      Jane Austen, Anna Quindlen (Introduction)    4.26
4                            Stephenie Meyer    3.60


                                        description language  \
0  WINNING MEANS FAME AND FORTUNE.LOSING MEANS CE…  English
1  There is a door at the end of a silent corrido…  English
2  The unforgettable novel of a childhood in a sl…  English
3  Alternate cover edition of ISBN 9780679783268S…  English
4  About three things I was absolutely positive.\…  English


                                             genres pages  \
0  ['Young Adult', 'Fiction', 'Dystopia', 'Fantas…    374
1  ['Fantasy', 'Young Adult', 'Fiction', 'Magic',…    870
2  ['Classics', 'Fiction', 'Historical Fiction', …    324
3  ['Classics', 'Fiction', 'Romance', 'Historical…    279
4  ['Young Adult', 'Fantasy', 'Romance', 'Vampire…    501


                      publisher publishDate  numRatings  likedPercent  \
0                Scholastic Press    09/14/08     6376780          96.0
1                 Scholastic Inc.    09/28/04     2507623          98.0
2  Harper Perennial Modern Classics    05/23/06     4501075          95.0
3                  Modern Library    10/10/00     2998241          94.0
4       Little, Brown and Company    09/06/06     4964519          78.0


                                           coverImg
0  https://i.gr-assets.com/images/S/compressed.ph…
1  https://i.gr-assets.com/images/S/compressed.ph…
2  https://i.gr-assets.com/images/S/compressed.ph…
3  https://i.gr-assets.com/images/S/compressed.ph…
4  https://i.gr-assets.com/images/S/compressed.ph…

```
[3]: duplicate_groups = df_subset[df_subset.duplicated(subset=cols_to_keep,␣
     ↪keep=False)]

     num_duplicate_groups = duplicate_groups.duplicated(subset=cols_to_keep).sum()

     print(f"Number of duplicate groups: {num_duplicate_groups}")
```

Number of duplicate groups: 11821

```
[4]: # View the first few duplicate rows
     duplicate_groups.sort_values(by=cols_to_keep).head()
```

[4]:                                                     title  \
     33714  "Evil" Arabs in American Popular Film: Orienta…
     33715  "Evil" Arabs in American Popular Film: Orienta…
     22310                           "The Mmmmmhh Thing"
     22311                           "The Mmmmmhh Thing"
     34580  "There Are Things I Want You to Know" about St…

                                                    author  rating  \
     33714                          Tim Jon Semmerling    4.50
     33715                          Tim Jon Semmerling    4.50
     22310                   K. Harvey (Goodreads Author)    4.67
     22311                   K. Harvey (Goodreads Author)    4.67
     34580  Eva Gabrielsson, Marie-Françoise Colombani, Li…    3.47

                                               description language  \
     33714  Runner-up, 2006 Arab American National Museum …  English
     33715  Runner-up, 2006 Arab American National Museum …  English
     22310  "The Mmmmmmhh Thing"… Is a provocative conte…  English
     22311  "The Mmmmmmhh Thing"… Is a provocative conte…  English
     34580  Here is the real inside story-not the one abou…  English

                                                 genres pages  \
     33714                                            []   303
     33715                                            []   303
     22310                                            []   350
     22311                                            []   350
     34580  ['Nonfiction', 'Biography', 'Memoir', 'Biograp…   209

                          publisher         publishDate  numRatings  \
     33714  University of Texas Press  September 1st 2006           4
     33715  University of Texas Press  September 1st 2006           4
     22310         Luminous Events LLC  November 11th 2014           3
     22311         Luminous Events LLC  November 11th 2014           3
     34580        Seven Stories Press       June 21st 2011        1348

            likedPercent                                        coverImg
```

```
33714          NaN  https://i.gr-assets.com/images/S/compressed.ph...
33715          NaN  https://i.gr-assets.com/images/S/compressed.ph...
22310        100.0  https://i.gr-assets.com/images/S/compressed.ph...
22311        100.0  https://i.gr-assets.com/images/S/compressed.ph...
34580         84.0  https://i.gr-assets.com/images/S/compressed.ph...
```

[5]:
```python
# Remove duplicates and keep the first occurrence
df_subset_cleaned = df_subset.drop_duplicates(subset=cols_to_keep, keep="first")

print(f"Before: {df_subset.shape}")
print(f"After: {df_subset_cleaned.shape}")
```

```
Before: (53944, 12)
After: (42123, 12)
```

[6]:
```python
def create_histogram(x, xlabel, title):
    plt.figure(figsize=(8, 5))
    plt.hist(df_subset_cleaned[x], bins=20, edgecolor="black")
    plt.title(title)
    plt.xlabel(xlabel)
    plt.ylabel("Count")
    plt.grid(axis="y", alpha=0.7)
    plt.tight_layout()
    plt.show()
```

[7]:
```python
create_histogram(x="rating", xlabel="Rating", title="Book Rating Distribution")
```
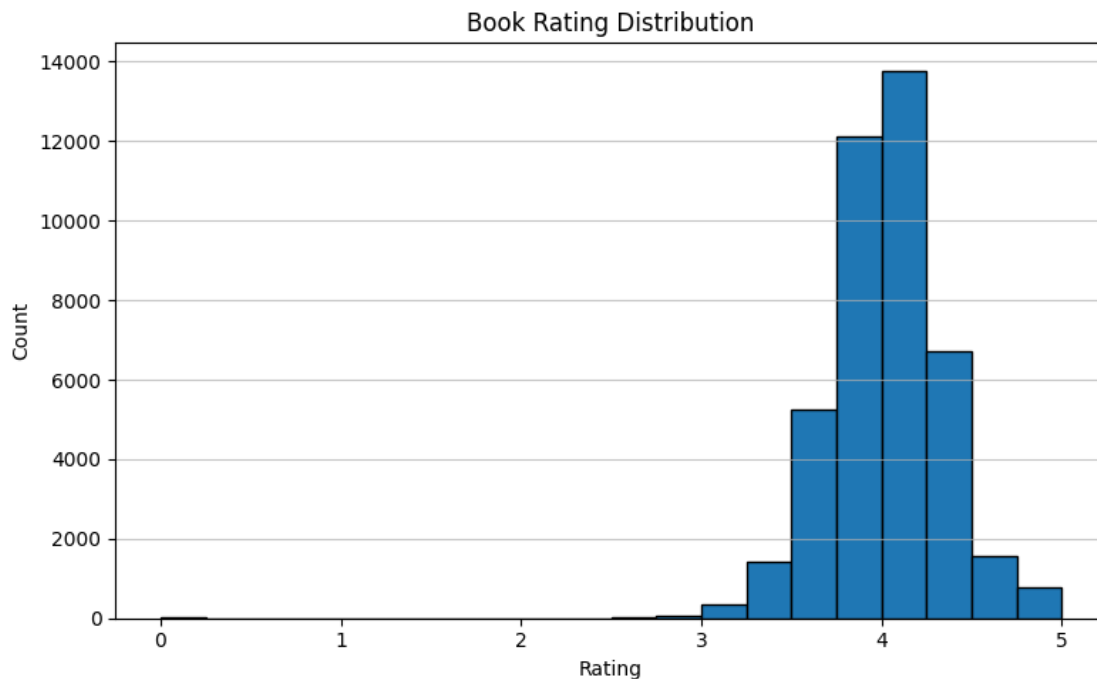
```
[8]: create_histogram(
         x="likedPercent", xlabel="Liked Percent", title="Book Liked Percent␣
     ↪Distribution"
     )
```

**Book Liked Percent Distribution**
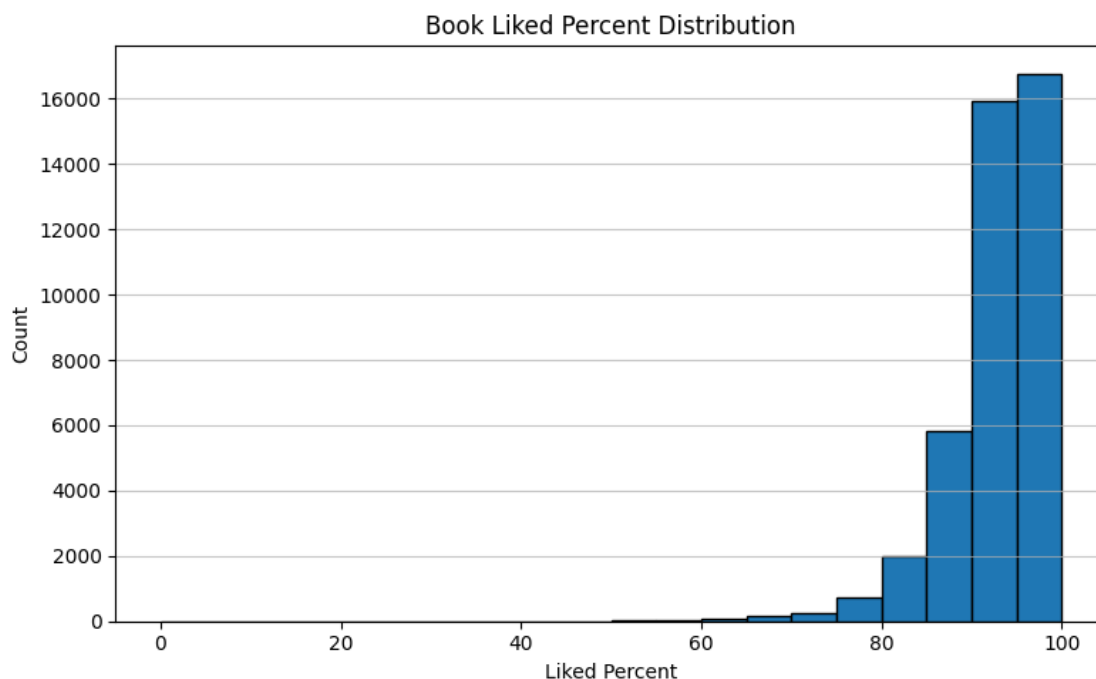


## 1.2 Preprocess the Description

```
[9]: import re
     import nltk
     from nltk.corpus import stopwords
     from nltk.tokenize import word_tokenize

     nltk.download("punkt_tab")

     stop_words = set(stopwords.words("english"))


     def preprocess_text(text):
         text = str(text).lower()
         text = re.sub(r"http\S+", "", text)
         text = re.sub(r"[^a-z\s]", "", text)
         tokens = word_tokenize(text)  # tokenize
         filtered_tokens = [word for word in tokens if word not in stop_words]
```

```
        cleaned_text = " ".join(filtered_tokens)
    return cleaned_text
```

```
[nltk_data] Downloading package punkt_tab to
[nltk_data]     /home/christiandarvs/nltk_data…
[nltk_data]   Package punkt_tab is already up-to-date!
```

[10]:
```
df_subset_cleaned["description_cleaned"] = df_subset_cleaned["description"].
 ↪apply(
    preprocess_text
)
```

```
/tmp/ipykernel_9573/3898192295.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df_subset_cleaned["description_cleaned"] =
df_subset_cleaned["description"].apply(
```

[11]:
```
df_subset_cleaned.head()
```

[11]:
```
                                       title  \
0                            The Hunger Games
1  Harry Potter and the Order of the Phoenix
2                       To Kill a Mockingbird
3                         Pride and Prejudice
4                                    Twilight


                               author  rating  \
0                      Suzanne Collins    4.33
1  J.K. Rowling, Mary GrandPré (Illustrator)    4.50
2                           Harper Lee    4.28
3  Jane Austen, Anna Quindlen (Introduction)    4.26
4                      Stephenie Meyer    3.60


                                         description language  \
0  WINNING MEANS FAME AND FORTUNE.LOSING MEANS CE…  English
1  There is a door at the end of a silent corrido…  English
2  The unforgettable novel of a childhood in a sl…  English
3  Alternate cover edition of ISBN 9780679783268S…  English
4  About three things I was absolutely positive.\…  English


                                         genres  pages  \
0  ['Young Adult', 'Fiction', 'Dystopia', 'Fantas…    374
1  ['Fantasy', 'Young Adult', 'Fiction', 'Magic',…    870
2  ['Classics', 'Fiction', 'Historical Fiction', …    324
```

```
3  ['Classics', 'Fiction', 'Romance', 'Historical…    279
4  ['Young Adult', 'Fantasy', 'Romance', 'Vampire…    501


                            publisher publishDate  numRatings  likedPercent  \
0                    Scholastic Press    09/14/08     6376780          96.0
1                     Scholastic Inc.    09/28/04     2507623          98.0
2   Harper Perennial Modern Classics    05/23/06     4501075          95.0
3                      Modern Library    10/10/00     2998241          94.0
4            Little, Brown and Company    09/06/06     4964519          78.0


                                      coverImg  \
0  https://i.gr-assets.com/images/S/compressed.ph…
1  https://i.gr-assets.com/images/S/compressed.ph…
2  https://i.gr-assets.com/images/S/compressed.ph…
3  https://i.gr-assets.com/images/S/compressed.ph…
4  https://i.gr-assets.com/images/S/compressed.ph…


                                 description_cleaned
0  winning means fame fortunelosing means certain…
1  door end silent corridor haunting harry pottte…
2  unforgettable novel childhood sleepy southern …
3  alternate cover edition isbn since immediate s…
4  three things absolutely positive first edward …
```

```python
"""
Return books that have:
    Ratings: 4 - 5
    Liked Percentage: 90% - 100%
    Number of Ratings: >= 250,000
"""

high_rating_and_percent = df_subset_cleaned[
    (df_subset_cleaned["rating"] >= 4)
    & (df_subset_cleaned["rating"] <= 5)
    & (df_subset_cleaned["likedPercent"] >= 90)
    & (df_subset_cleaned["likedPercent"] <= 100)
    & (df_subset_cleaned["numRatings"] >= 150_000)
]

high_rating_and_percent.sort_values(
    ascending=[False, False], by=["rating", "likedPercent"]
)
high_rating_and_percent.reset_index(drop=True)
high_rating_and_percent.to_csv("highly_rated_popular_books.csv", index=False)
```

## 1.3 Emotion Mapping

```python
[13]: import torch
      from transformers import RobertaTokenizer, RobertaForSequenceClassification

      device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

      model_name = "j-hartmann/emotion-english-distilroberta-base"

      # Use RobertaTokenizer and RobertaForSequenceClassification, NOT BertTokenizer
      tokenizer = RobertaTokenizer.from_pretrained(model_name)
      model = RobertaForSequenceClassification.from_pretrained(model_name)
      model = model.to(device)


      # Get labels
      emotion_labels = [model.config.id2label[i] for i in range(model.config.
        ↪num_labels)]
      print("Model's Actual Emotion Labels:", emotion_labels)
```

/home/christiandarvs/pytorch-venv/lib/python3.12/site-packages/tqdm/auto.py:21:
TqdmWarning: IProgress not found. Please update jupyter and ipywidgets. See
https://ipywidgets.readthedocs.io/en/stable/user_install.html
  from .autonotebook import tqdm as notebook_tqdm

Model's Actual Emotion Labels: ['anger', 'disgust', 'fear', 'joy', 'neutral',
'sadness', 'surprise']

```python
[14]: def classify_emotions_in_batch(texts, batch_size=32):
          results = []
          for i in range(0, len(texts), batch_size):
              batch_texts = texts[i : i + batch_size]

              inputs = tokenizer(
                  batch_texts,
                  return_tensors="pt",
                  truncation=True,
                  padding=True,
                  max_length=128,
              )
              inputs = {k: v.to(device) for k, v in inputs.items()}

              with torch.no_grad():
                  outputs = model(**inputs)
                  probs = torch.nn.functional.softmax(outputs.logits, dim=1)

              for prob in probs:
                  # Get index of max probability
```

```
            max_idx = prob.argmax().item()
            predicted_label = emotion_labels[max_idx]
            results.append(predicted_label)
    return results
```

[15]:
```
high_rating_and_percent["description"] = high_rating_and_percent["description"].
 ↪astype(
    str
)

# Run classification
high_rating_and_percent["predicted_emotion"] = classify_emotions_in_batch(
    high_rating_and_percent["description"].tolist()
)

# Save output
high_rating_and_percent.to_csv(
    "highly_rated_popular_books_with_emotions.csv", index=False
)
print("Saved to highly_rated_popular_books_with_emotions.csv")
```

/tmp/ipykernel_9573/287084868.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  high_rating_and_percent["description"] =
high_rating_and_percent["description"].astype(

Saved to highly_rated_popular_books_with_emotions.csv

/tmp/ipykernel_9573/287084868.py:6: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  high_rating_and_percent["predicted_emotion"] = classify_emotions_in_batch(

[16]:
```
df = pd.read_csv("./highly_rated_popular_books_with_emotions.csv")
df
```

[16]:
```
                                             title  \
0                                 The Hunger Games
1      Harry Potter and the Order of the Phoenix
2                            To Kill a Mockingbird
3                              Pride and Prejudice
4                                   The Book Thief
```

```
..                                      …
585                        The Way of Kings
586                       The Giver of Stars
587                           City of Girls
588                        The Way of Kings
589                          Norwegian Wood

                                                author  rating  \
0                                       Suzanne Collins    4.33
1              J.K. Rowling, Mary GrandPré (Illustrator)    4.50
2                                            Harper Lee    4.28
3              Jane Austen, Anna Quindlen (Introduction)    4.26
4                     Markus Zusak (Goodreads Author)    4.37
..                                                   …       …
585              Brandon Sanderson (Goodreads Author)    4.63
586                     Jojo Moyes (Goodreads Author)    4.28
587             Elizabeth Gilbert (Goodreads Author)    4.04
588             Brandon Sanderson (Goodreads Author)    4.63
589                               Haruki Murakami    4.03

                                           description language  \
0      WINNING MEANS FAME AND FORTUNE.LOSING MEANS CE…  English
1      There is a door at the end of a silent corrido…  English
2      The unforgettable novel of a childhood in a sl…  English
3      Alternate cover edition of ISBN 9780679783268S…  English
4      Librarian's note: An alternate cover edition c…  English
..                                                  …        …
585    I long for the days before the Last Desolation…  English
586    From the author of Me Before You, set in Depre…  English
587    "Life is both fleeting and dangerous, and ther…  English
588    From #1 New York Times bestselling author Bran…  English
589    When he hears her favourite Beatles song, Toru…  English

                                           genres  pages  \
0      ['Young Adult', 'Fiction', 'Dystopia', 'Fantas…    374
1      ['Fantasy', 'Young Adult', 'Fiction', 'Magic',…    870
2      ['Classics', 'Fiction', 'Historical Fiction', …    324
3      ['Classics', 'Fiction', 'Romance', 'Historical…    279
4      ['Historical Fiction', 'Fiction', 'Young Adult…    552
..                                                  …      …
585    ['Fantasy', 'Fiction', 'Epic Fantasy', 'High F…   1258
586    ['Historical Fiction', 'Fiction', 'Audiobook',…    388
587    ['Historical Fiction', 'Fiction', 'Audiobook',…    496
588    ['Fantasy', 'Fiction', 'Epic Fantasy', 'High F…   1137
589    ['Fiction', 'Japan', 'Romance', 'Japanese Lite…    388

                      publisher       publishDate  numRatings  \
```

```
0                 Scholastic Press           09/14/08    6376780
1                 Scholastic Inc.            09/28/04    2507623
2    Harper Perennial Modern Classics        05/23/06    4501075
3                 Modern Library             10/10/00    2998241
4                 Alfred A. Knopf            03/14/06    1834276
..                      ...                     ...         ...
585               Tom Doherty         May 24th 2011        303454
586       Pamela Dorman Books      October 8th 2019        183231
587          Riverhead Books        April 7th 2020        157326
588               Tor Books      August 31st 2010        304529
589                 Vintage    November 8th 2012        348525

     likedPercent                                        coverImg  \
0            96.0  https://i.gr-assets.com/images/S/compressed.ph…
1            98.0  https://i.gr-assets.com/images/S/compressed.ph…
2            95.0  https://i.gr-assets.com/images/S/compressed.ph…
3            94.0  https://i.gr-assets.com/images/S/compressed.ph…
4            96.0  https://i.gr-assets.com/images/S/compressed.ph…
..            ...                                              ...
585          98.0  https://i.gr-assets.com/images/S/compressed.ph…
586          97.0  https://i.gr-assets.com/images/S/compressed.ph…
587          95.0  https://i.gr-assets.com/images/S/compressed.ph…
588          98.0  https://i.gr-assets.com/images/S/compressed.ph…
589          94.0  https://i.gr-assets.com/images/S/compressed.ph…

                           description_cleaned predicted_emotion
0    winning means fame fortunelosing means certain…            anger
1    door end silent corridor haunting harry pottte…             fear
2    unforgettable novel childhood sleepy southern …              joy
3    alternate cover edition isbn since immediate s…              joy
4    librarians note alternate cover edition found …          disgust
..                                           …                    …
585  long days last desolation heralds abandoned us…          disgust
586  author set depressionera america breathtaking …             fear
587  life fleeting dangerous point denying pleasure…          sadness
588  new york times bestselling author brandon sand…          neutral
589  hears favourite beatles song toru watanabe rec…             fear
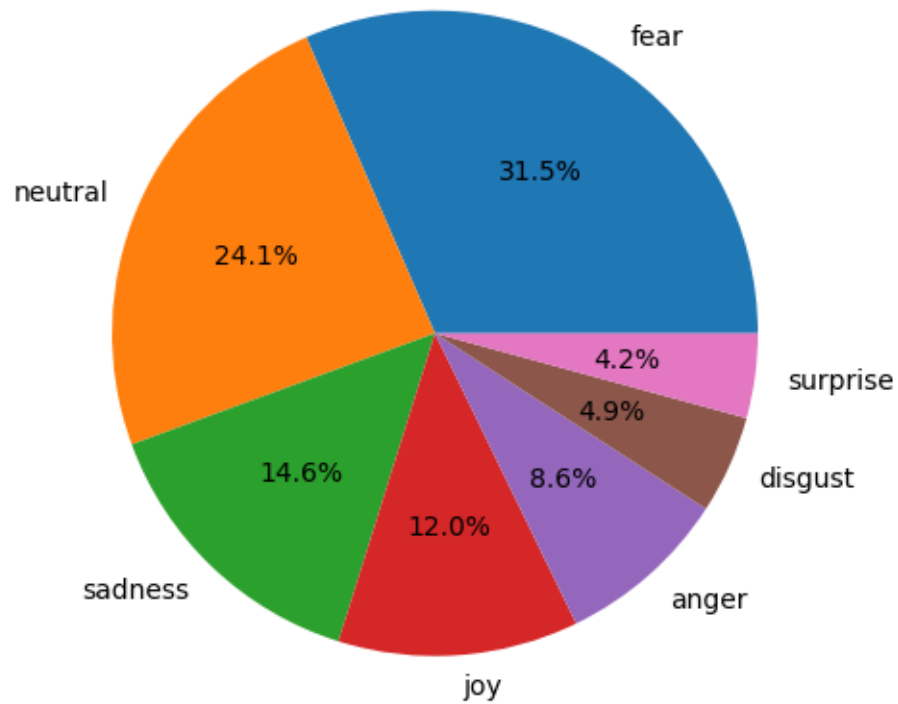
[590 rows x 14 columns]
```

```python
emotion_counts = df["predicted_emotion"].value_counts()

# Plot with correct label order
plt.pie(emotion_counts, labels=emotion_counts.index, autopct="%1.1f%%")
plt.axis("equal")
plt.title("Predicted Emotions from Book Descriptions")
plt.show()
```

```
# Print raw counts
print(emotion_counts)
```

## Predicted Emotions from Book Descriptions



```
predicted_emotion
fear        186
neutral     142
sadness      86
joy          71
anger        51
disgust      29
surprise     25
Name: count, dtype: int64
```

[31]: 
```
sns.countplot(data=df, x="predicted_emotion")
plt.title("Predicted Emotions from Book Descriptions")
plt.show()
```

14

## Predicted Emotions from Book Descriptions
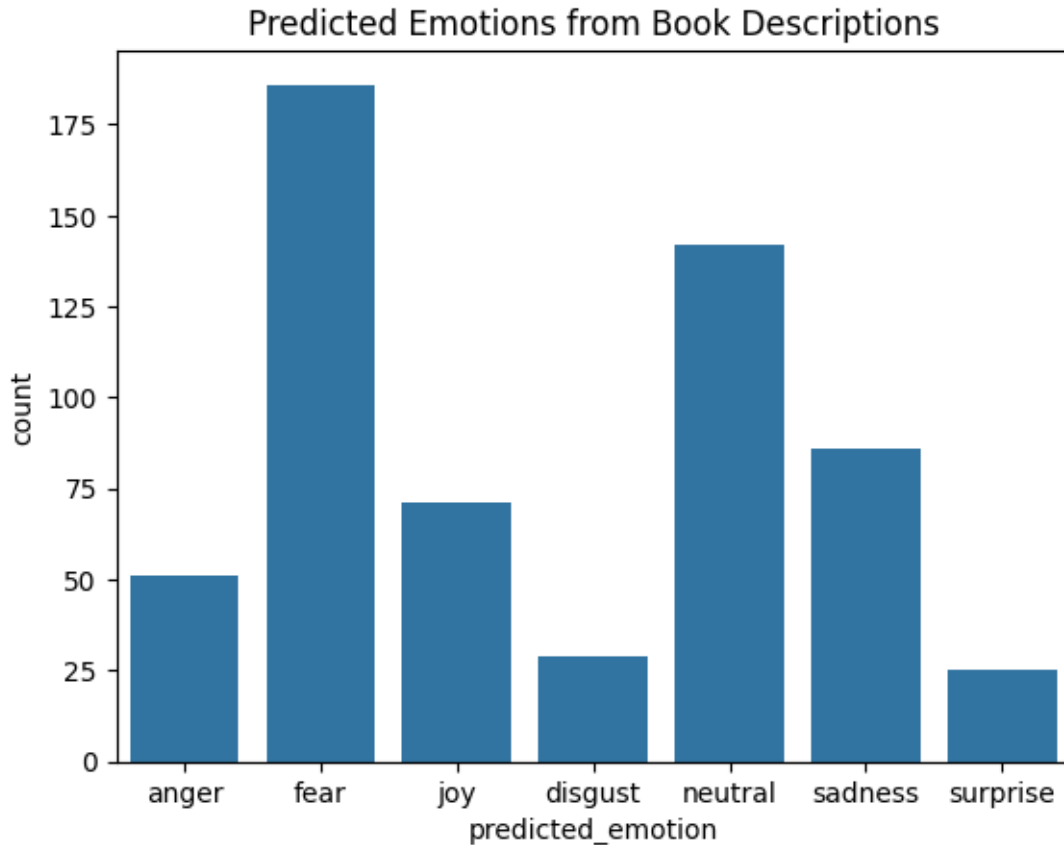


```
[32]: print(
          f'The Shining Description: {df[df["title"] == "The Shining"].
       ↪description_cleaned.values}'
      )

      print(
          f'\nThe Shining Prediction Emotion: {df[df["title"] == "The Shining"].
       ↪predicted_emotion}'
      )
```

The Shining Description: ['jack torrances new job overlook hotel perfect chance
fresh start offseason caretaker atmospheric old hotel hell plenty time spend
reconnecting family working writing harsh winter weather sets idyllic location
feels ever remoteand sinister one notice strange terrible forces gathering
around overlook danny torrance uniquely gifted fiveyearoldthe shiningstephen
king goodreads authorthe shining shiningstephen king goodreads authorthe
shining']

The Shining Prediction Emotion: 80    fear
Name: predicted_emotion, dtype: object

```
[33]:  book_emotions = (
           df.groupby("title")["predicted_emotion"]
           .value_counts(normalize=True)
           .unstack()
           .fillna(0)
       )


       book_emotions = book_emotions.reindex(columns=emotion_labels, fill_value=0)

       print("Emotional Fingerprints:")
       display(book_emotions)
```

Emotional Fingerprints:

```
predicted_emotion                                      anger  disgust  fear  joy  \
title
'Salem's Lot                                            0.0      0.0   1.0  0.0
11/22/63                                                0.0      1.0   0.0  0.0
1776                                                    0.0      0.0   0.0  0.0
1984                                                    0.0      0.0   1.0  0.0
1st to Die                                              0.0      0.0   1.0  0.0
...                                                      ...      ...   ...  ...
Wool Omnibus                                            1.0      0.0   0.0  0.0
Words of Radiance                                       0.0      0.0   0.0  0.0
World War Z: An Oral History of the Zombie War          0.0      0.0   1.0  0.0
World Without End                                       0.0      0.0   0.0  0.0
Zero to One: Notes on Startups, or How to Build…  0.0      0.0   0.0  0.0

predicted_emotion                                      neutral  sadness  surprise
title
'Salem's Lot                                            0.0      0.0       0.0
11/22/63                                                0.0      0.0       0.0
1776                                                    1.0      0.0       0.0
1984                                                    0.0      0.0       0.0
1st to Die                                              0.0      0.0       0.0
...                                                      ...      ...       ...
Wool Omnibus                                            0.0      0.0       0.0
Words of Radiance                                       0.0      1.0       0.0
World War Z: An Oral History of the Zombie War          0.0      0.0       0.0
World Without End                                       1.0      0.0       0.0
Zero to One: Notes on Startups, or How to Build…  1.0      0.0       0.0

[587 rows x 7 columns]
```

```
[34]:  from sklearn.metrics.pairwise import cosine_similarity
       import numpy as np
```

```python
def recommend_books(user_emotion_vector, book_emotions_df, full_df, top_n=10):
    book_vectors = book_emotions_df[emotion_labels].values
    similarities = cosine_similarity([user_emotion_vector], book_vectors)

    similar_books_idx = np.argsort(similarities[0])[::-1]
    top_indices = similar_books_idx[:top_n]
    top_titles = book_emotions_df.index[top_indices]

    # Filter the original DataFrame and reset index
    recommended_books_df = (
        full_df[full_df["title"].isin(top_titles)]
        .drop_duplicates(subset="title")
        .reset_index(drop=True)
    )

    # Sort by recommendation order
    recommended_books_df["title_order"] = recommended_books_df["title"].apply(
        lambda x: list(top_titles).index(x)
    )
    recommended_books_df = (
        recommended_books_df.sort_values("title_order")
        .drop(columns="title_order")
        .reset_index(drop=True)
    )

    return recommended_books_df
```

```python
[35]: user_input = [
    0.0,   # anger
    0.0,   # disgust
    1.0,   # fear
    0.0,   # joy
    0.0,   # neutral
    0.0,   # sadness
    0.0,   # surprise
]


top_10_books = recommend_books(user_input, book_emotions, df, top_n=10)

for idx, row in top_10_books.iterrows():
    print(f"\nTitle: {row['title']}")
    print(f"Author: {row['author']}")
    print(f"Rating: {row['rating']}")
    print(f"Description: {row['description']}")
    print(f"Language: {row['language']}")
```

```
    print(f"Genres: {row['genres']}")
    print(f"Pages: {row['pages']}")
    print(f"Publisher: {row['publisher']}")
    print(f"Publish Date: {row['publishDate']}")
    print(f"Number of Ratings: {row['numRatings']}")
    print(f"Liked Percent: {row['likedPercent']}")
    print(f"Cover Image URL: {row['coverImg']}")
```

Title: 'Salem's Lot
Author: Stephen King (Goodreads Author)
Rating: 4.03
Description: Thousands of miles away from the small township of 'Salem's Lot,
two terrified people, a man and a boy, still share the secrets of those
clapboard houses and tree-lined streets. They must return to 'Salem's Lot for a
final confrontation with the unspeakable evil that lives on in the town.'Salem's
LotStephen King (Goodreads Author)'Salem's LotStephen King (Goodreads Author)
Language: English
Genres: ['Horror', 'Fiction', 'Vampires', 'Fantasy', 'Thriller', 'Paranormal',
'Supernatural', 'Audiobook', 'Adult', 'Mystery']
Pages: 483
Publisher: New English Library
Publish Date: 10/28/91
Number of Ratings: 334051
Liked Percent: 94.0
Cover Image URL: https://i.gr-
assets.com/images/S/compressed.photo.goodreads.com/books/1327891565l/11590.jpg


Title: 1984
Author: George Orwell
Rating: 4.19
Description: Among the seminal texts of the 20th century, Nineteen Eighty-Four
is a rare work that grows more haunting as its futuristic purgatory becomes more
real. Published in 1949, the book offers political satirist George Orwell's
nightmarish vision of a totalitarian, bureaucratic world and one poor stiff's
attempt to find individuality. The brilliance of the novel is Orwell's
prescience of modern life-the ubiquity of television, the distortion of the
language-and his ability to construct such a thorough version of hell. Required
reading for students since it was published, it ranks among the most terrifying
novels ever written.1984George Orwell1984George Orwell
Language: English
Genres: ['Classics', 'Fiction', 'Science Fiction', 'Dystopia', 'Literature',
'Novels', 'Politics', 'School', 'Fantasy', 'Adult']
Pages: 237
Publisher: Houghton Mifflin Harcourt
Publish Date: 09/03/13
Number of Ratings: 3140442
Liked Percent: 94.0

Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1532714506l/40961427._SX318_.jpg

Title: 1st to Die
Author: James Patterson (Goodreads Author)
Rating: 4.08
Description: James Patterson, bestselling author of the Alex Cross novels Along Came a Spider, Kiss the Girls, and Pop Goes the Weasel, offers the first of a new series dubbed The Women's Murder Club, featuring a four-woman team that occasionally works outside the system. None of the gritty darkness or frenzied action is lost in 1st to Die, although the female protagonists offer an even deeper emotional context to this suspense thriller.Inspector Lindsay Boxer of the San Francisco Police Department suddenly finds herself in the middle of two horrifying situations: The first is that she's just learned she has an often-fatal blood disease. The second is a double homicide case she is now heading up that involves the murder of newlyweds on their wedding night. Burdened with Chris Raleigh, a new partner reassigned from the mayor's office, Lindsay finds that she has too much to deal with and turns to her best friend, Claire, the head ME on the case. Claire offers helpful advice and human, friendly contact amid a job filled with violence, cruelty, and fear.Soon a fledgling newspaper reporter, Cindy, makes contact with Lindsay looking for a career-making story. Although Lindsay can't officially comment on the case, the two women form a rapport, and Cindy joins Lindsay and Claire for their weekly meeting. When a second pair of newlyweds is murdered, and later a third, the investigation leads to a prominent crime writer, Nicholas Jenks, who has a history of spousal abuse and a predilection for kinky, dangerous sex games. With the help of an understanding assistant D.A., Jill Bernhardt, Lindsay tries to make a case against Jenks, who even had an affair with one of the slain women. Eventually Jill joins the Murder Club, and the four ladies share private interdepartmental information in an effort to track and stop the killer before he strikes again.The major subplot -- Lindsay's facing up to her illness even while she learns to fall in love again -- carefully compensates for the novel's coarse scenes of brutality. Lindsay Boxer is't merely an obsessed cop trailing a maniac; she's also a terrified woman confronting the onslaught of disease. The story lines balance out to show us the true mettle of someone who puts the safety of others before her own.Again, Patterson's skill for producing furiously paced fiction are evident as the novel breezes by rapidly. The short chapters keep the narrative leaping with increasingly taut plot elements, but there's an emotional commitment that makes our protagonist even more amiable and involving. 1st to Die is a novel that works as an intense series of character portraits that will leave the reader touched and delighted.--Tom Piccirilli1st to DieJames Patterson (Goodreads Author)Women's Murder Club #11st to DieJames Patterson (Goodreads Author)Women's Murder Club #1
Language: English
Genres: ['Mystery', 'Fiction', 'Crime', 'Thriller', 'Suspense', 'Mystery Thriller', 'Murder Mystery', 'Audiobook', 'Detective', 'Adult']
Pages: 424
Publisher: Grand Central Publishing

Publish Date: May 20th 2005
Number of Ratings: 296188
Liked Percent: 94.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com
/books/1522906581l/13137._SY475_.jpg


Title: World War Z: An Oral History of the Zombie War
Author: Max Brooks (Goodreads Author)
Rating: 4.01
Description: The Zombie War came unthinkably close to eradicating humanity. Max
Brooks, driven by the urgency of preserving the acid-etched first-hand
experiences of the survivors from those apocalyptic years, traveled across the
United States of America and throughout the world, from decimated cities that
once teemed with upwards of thirty million souls to the most remote and
inhospitable areas of the planet. He recorded the testimony of men, women, and
sometimes children who came face-to-face with the living, or at least the
undead, hell of that dreadful time. World War Z is the result. Never before have
we had access to a document that so powerfully conveys the depth of fear and
horror, and also the ineradicable spirit of resistance, that gripped human
society through the plague years.Ranging from the now infamous village of New
Dachang in the United Federation of China, where the epidemiological trail began
with the twelve-year-old Patient Zero, to the unnamed northern forests where
untold numbers sought a terrible and temporary refuge in the cold, to the United
States of Southern Africa, where the Redeker Plan provided hope for humanity at
an unspeakable price, to the west-of-the-Rockies redoubt where the North
American tide finally started to turn, this invaluable chronicle reflects the
full scope and duration of the Zombie War.Most of all, the book captures with
haunting immediacy the human dimension of this epochal event. Facing the often
raw and vivid nature of these personal accounts requires a degree of courage on
the part of the reader, but the effort is invaluable because, as Mr. Brooks says
in his introduction, "By excluding the human factor, aren't we risking the kind
of personal detachment from history that may, heaven forbid, lead us one day to
repeat it? And in the end, isn't the human factor the only true difference
between us and the enemy we now refer to as 'the living dead'?"Note: Some of the
numerical and factual material contained in this edition was previously
published under the auspices of the United Nations Postwar Commission.World War
Z: An Oral History of the Zombie WarMax Brooks (Goodreads Author)World War
ZWorld War Z: An Oral History of the Zombie WarMax Brooks (Goodreads
Author)World War Z
Language: English
Genres: ['Horror', 'Fiction', 'Zombies', 'Science Fiction', 'Fantasy', 'Post
Apocalyptic', 'Dystopia', 'Audiobook', 'Apocalyptic', 'War']
Pages: 342
Publisher: Crown
Publish Date: 09/12/06
Number of Ratings: 439825
Liked Percent: 92.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com

/books/1528312647l/8908._SX318_.jpg

Title: 2001: A Space Odyssey
Author: Arthur C. Clarke
Rating: 4.15
Description: On the Moon, an enigma is uncovered.

So great are the implications of this discovery that for the first time men are
sent out deep into our solar system.

But long before their destination is reached, things begin to go horribly,
inexplicably wrong…

One of the greatest-selling science fiction novels of our time, this classic
book will grip you to the very end.2001: A Space OdysseyArthur C. ClarkeSpace
Odyssey #12001: A Space OdysseyArthur C. ClarkeSpace Odyssey #1
Language: English
Genres: ['Science Fiction', 'Fiction', 'Classics', 'Science Fiction Fantasy',
'Space', 'Novels', 'Fantasy', 'Audiobook', 'Literature', 'Speculative Fiction']
Pages: 297
Publisher: Roc
Publish Date: 09/01/00
Number of Ratings: 259477
Liked Percent: 95.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com
/books/1432468943l/70535._SY475_.jpg

Title: A Long Way Gone: Memoirs of a Boy Soldier
Author: Ishmael Beah
Rating: 4.16
Description: The devastating story of war through the eyes of a child soldier.
Beah tells how, at the age of twelve, he fled attacking rebels and wandered a
land rendered unrecognizable by violence. By thirteen, he'd been picked up by
the government army, and became a soldier.My new friends have begun to suspect I
haven't told them the full story of my life."Why did you leave Sierra
Leone?""Because there is a war.""You mean, you saw people running around with
guns and shooting each other?""Yes, all the time.""Cool."I smile a little."You
should tell us about it sometime.""Yes, sometime." This is how wars are fought
now: by children, hopped-up on drugs and wielding AK-47s. Children have become
soldiers of choice. In the more than fifty conflicts going on worldwide, it is
estimated that there are some 300,000 child soldiers. Ishmael Beah used to be
one of them. What is war like through the eyes of a child soldier? How does one
become a killer? How does one stop? Child soldiers have been profiled by
journalists, and novelists have struggled to imagine their lives. But until now,
there has not been a first-person account from someone who came through this
hell and survived.In A Long Way Gone, Beah, now twenty-five years old, tells a
riveting story: how at the age of twelve, he fled attacking rebels and wandered
a land rendered unrecognizable by violence. By thirteen, he'd been picked up by

the government army, and Beah, at heart a gentle boy, found that he was capable of truly terrible acts. This is a rare and mesmerizing account, told with real literary force and heartbreaking honesty.A Long Way Gone: Memoirs of a Boy SoldierIshmael BeahA Long Way Gone: Memoirs of a Boy SoldierIshmael Beah
Language: English
Genres: ['Nonfiction', 'Memoir', 'Africa', 'Biography', 'War', 'History', 'Biography Memoir', 'Autobiography', 'School', 'Sierra Leone']
Pages: 229
Publisher: Sarah Crichton Books
Publish Date: February 13th 2007
Number of Ratings: 163098
Liked Percent: 95.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1524839125l/43015._SY475_.jpg

Title: A Monster Calls
Author: Patrick Ness
Rating: 4.35
Description: The bestselling novel about love, loss and hope from the twice Carnegie Medal-winning Patrick Ness. Conor has the same dream every night, ever since his mother first fell ill, ever since she started the treatments that don't quite seem to be working. But tonight is different. Tonight, when he wakes, there's a visitor at his window. It's ancient, elemental, a force of nature. And it wants the most dangerous thing of all from Conor. It wants the truth. Patrick Ness takes the final idea of the late, award-winning writer Siobhan Dowd and weaves an extraordinary and heartbreaking tale of mischief, healing and above all, the courage it takes to survive.A Monster CallsPatrick NessA Monster CallsPatrick Ness
Language: English
Genres: ['Fantasy', 'Young Adult', 'Fiction', 'Horror', 'Contemporary', 'Middle Grade', 'Magical Realism', 'Childrens', 'Death', 'Paranormal']
Pages: 237
Publisher: Walker Books
Publish Date: 05/07/15
Number of Ratings: 208089
Liked Percent: 96.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1430763000l/25480342.jpg

Title: A Prayer for Owen Meany
Author: John Irving (Goodreads Author)
Rating: 4.23
Description: Eleven-year-old Owen Meany, playing in a Little League baseball game in Gravesend, New Hampshire, hits a foul ball and kills his best friend's mother. Owen doesn't believe in accidents; he believes he is God's instrument. What happens to Owen after that 1953 foul is both extraordinary and terrifying. At moments a comic, self-deluded victim, but in the end the principal, tragic actor in a divine plan, Owen Meany is the most heartbreaking hero John Irving

has yet created.A Prayer for Owen MeanyJohn Irving (Goodreads Author)A Prayer for Owen MeanyJohn Irving (Goodreads Author)
Language: English
Genres: ['Fiction', 'Classics', 'Contemporary', 'Literature', 'Literary Fiction', 'Novels', 'Historical Fiction', 'Adult Fiction', 'Coming Of Age', 'American']
Pages: 637
Publisher: Black Swan
Publish Date: 10/28/90
Number of Ratings: 285954
Liked Percent: 94.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1260470010l/4473.jpg

Title: A Study in Scarlet
Author: Arthur Conan Doyle
Rating: 4.16
Description: "A Study in Scarlet" is the first published story of one of the most famous literary detectives of all time, Sherlock Holmes. Here Dr. Watson, who has just returned from a war in Afghanistan, meets Sherlock Holmes for the first time when they become flat-mates at the famous 221 B Baker Street. In "A Study in Scarlet" Sherlock Holmes investigates a murder at Lauriston Gardens as Dr. Watson tags along with Holmes while narratively detailing his amazing deductive abilities.A Study in ScarletArthur Conan DoyleSherlock Holmes #1A Study in ScarletArthur Conan DoyleSherlock Holmes #1
Language: English
Genres: ['Classics', 'Mystery', 'Fiction', 'Crime', 'Detective', 'Mystery Thriller', 'Historical Fiction', 'Audiobook', '19th Century', 'British Literature']
Pages: 123
Publisher: Digireads.com
Publish Date: January 1st 2005
Number of Ratings: 324483
Liked Percent: 96.0
Cover Image URL: https://i.gr-assets.com/images/S/compressed.photo.goodreads.com/books/1519031842l/102868._SX318_.jpg

Title: A Thousand Splendid Suns
Author: Khaled Hosseini (Goodreads Author)
Rating: 4.38
Description: A Thousand Splendid Suns is a breathtaking story set against the volatile events of Afghanistan's last thirty years-from the Soviet invasion to the reign of the Taliban to post-Taliban rebuilding-that puts the violence, fear, hope, and faith of this country in intimate, human terms. It is a tale of two generations of characters brought jarringly together by the tragic sweep of war, where personal lives-the struggle to survive, raise a family, find happiness-are inextricable from the history playing out around them.Propelled by the same storytelling instinct that made The Kite Runner a beloved classic, A

Thousand Splendid Suns is at once a remarkable chronicle of three decades of
Afghan history and a deeply moving account of family and friendship. It is a
striking, heart-wrenching novel of an unforgiving time, an unlikely friendship,
and an indestructible love-a stunning accomplishment.--front flapA Thousand
Splendid SunsKhaled Hosseini (Goodreads Author)A Thousand Splendid SunsKhaled
Hosseini (Goodreads Author)
Language: English
Genres: ['Fiction', 'Historical Fiction', 'Contemporary', 'Novels',
'Historical', 'War', 'Adult Fiction', 'Adult', 'Drama', 'Literature']
Pages: 372
Publisher: Riverhead Books
Publish Date: 06/01/07
Number of Ratings: 1151288
Liked Percent: 97.0
Cover Image URL: https://i.gr-
assets.com/images/S/compressed.photo.goodreads.com/books/1345958969l/128029.jpg