

search_engine_for_ecommerce

January 30, 2026

```
[1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics.pairwise import cosine_similarity
from sklearn.feature_extraction.text import TfidfVectorizer
import os
from scipy.sparse import coo_matrix

[2]: train_data = pd.read_csv(r'C:\Users\himan\Downloads\amazon.csv.
˓→zip', quotechar='"', skipinitialspace=True)

[3]: train_data.columns

[3]: Index(['product_id', 'product_name', 'category', 'discounted_price',
       'actual_price', 'discount_percentage', 'rating', 'rating_count',
       'about_product', 'user_id', 'user_name', 'review_id', 'review_title',
       'review_content', 'img_link', 'product_link'],
       dtype='object')

[4]: train_data.head()

[4]: product_id                               product_name \
0 B07JW9H4J1    Wayona Nylon Braided USB to Lightning Fast Cha...
1 B098NS6PVG    Ambrane Unbreakable 60W / 3A Fast Charging 1.5...
2 B096MSW6CT    Sounce Fast Phone Charging Cable & Data Sync U...
3 B08HDJ86NZ    boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...
4 B08CF3B7N1    Portronics Konnect L 1.2M Fast Charging 3A 8 P...

                                              category discounted_price \
0  Computers&Accessories|Accessories&Peripherals|...           399
1  Computers&Accessories|Accessories&Peripherals|...           199
2  Computers&Accessories|Accessories&Peripherals|...           199
3  Computers&Accessories|Accessories&Peripherals|...           329
4  Computers&Accessories|Accessories&Peripherals|...           154

   actual_price discount_percentage rating rating_count \
0        1,099             64%      4.2      24,269
```

1	349	43%	4.0	43,994
2	1,899	90%	3.9	7,928
3	699	53%	4.2	94,363
4	399	61%	4.2	16,905

about_product \

0 High Compatibility : Compatible With iPhone 12...
 1 Compatible with all Type C enabled devices, be...
 2 Fast Charger& Data Sync -With built-in safet...
 3 The boAt Deuce USB 300 2 in 1 cable is compati...
 4 [CHARGE & SYNC FUNCTION]- This cable comes wit...

user_id \

0 AG3D604STAQKAY2UVGEUV46KN35Q ,AHMY5CWJMMK5BJRBB...
 1 AECPFYFQVRUWC3KGNLJI0REFP5LQ ,AGYYVPDD7YG7FYNBX...
 2 AGU3BBQ2V2DDAMOAKGFAWDDQ6QHA ,AESFLDV2PT363T2AQ...
 3 AEWAZDZZJLQUYVOVGBEUKSLXHQ5A ,AG5HTSFRRRE6NL3M5S...
 4 AE3Q6KSUK5P75D5HFYHCRAOLODSA ,AFUGIFH5ZAFXRDSZH...

user_name \

0 Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...
 1 ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...
 2 Kunal,Himanshu,viswanath,sai niharka,saqib mal...
 3 Omkar dhale,JD,HEMALATHA,Ajwadh a.,amar singh ...
 4 rahuls6099,Swasat Borah,Ajay Wadke,Pranali,RVK...

review_id \

0 R3HXWTOLRPONMF ,R2AJM3LFTLZHFO ,R6AQJGUP6P86 ,R1...
 1 RG1QE07R9HS2 ,R1SMWZQ86XIN8U ,R2J3Y1WL29GWDE ,RY...
 2 R3J3EQQ9TZI5ZJ ,R3E7WBGK7IDOKV ,RWU79XKQ6I1QF ,R2...
 3 R3EEUZKKK9J36I ,R3HJVYCLY0Y554 ,REDECAZ7AMPQC ,R1...
 4 R1BP4L2HH9TFUP ,R16PVJEXKV6QZS ,R2UPDB81N66T4P ,R...

review_title \

0 Satisfied,Charging is really fast,Value for mo...
 1 A Good Braided Cable for Your Type C Device,Go...
 2 Good speed for earlier versions,Good Product,W...
 3 Good product,Good one,Nice,Really nice product...
 4 As good as original,Decent,Good one for second...

review_content \

0 Looks durable Charging is fine tooNo complains...
 1 I ordered this cable to connect my phone to An...
 2 Not quite durable and sturdy,https://m.media-a...
 3 Good product,long wire,Charges good,Nice,I bou...
 4 Bought this instead of original apple, does th...

```
                                img_link \
0  https://m.media-amazon.com/images/W/WEBP_40237...
1  https://m.media-amazon.com/images/W/WEBP_40237...
2  https://m.media-amazon.com/images/W/WEBP_40237...
3  https://m.media-amazon.com/images/I/41V5FtEWPK...
4  https://m.media-amazon.com/images/W/WEBP_40237...

                                product_link
0  https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
1  https://www.amazon.in/Ambrane-Unbreakable-Char...
2  https://www.amazon.in/Sounce-iPhone-Charging-C...
3  https://www.amazon.in/Deuce-300-Resistant-Tang...
4  https://www.amazon.in/Portronics-Konnect-POR-1...
```

[5]: train_data['product_name']

```
[5]: 0      Wayona Nylon Braided USB to Lightning Fast Cha...
1      Ambrane Unbreakable 60W / 3A Fast Charging 1.5...
2      Sounce Fast Phone Charging Cable & Data Sync U...
3      boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...
4      Portronics Konnect L 1.2M Fast Charging 3A 8 P...
...
1460    Noir Aqua - 5pcs PP Spun Filter + 1 Spanner | ...
1461    Prestige Delight PRWO Electric Rice Cooker (1 ...
1462    Bajaj Majesty RX10 2000 Watts Heat Convector R...
1463    Havells Ventil Air DSP 230mm Exhaust Fan (Pist...
1464    Borosil Jumbo 1000-Watt Grill Sandwich Maker (...  
Name: product_name, Length: 1465, dtype: object
```

[6]: train_data.shape

[6]: (1465, 16)

[7]: train_data.isnull().sum()

```
[7]: product_id          0
product_name         0
category            0
discounted_price    0
actual_price        0
discount_percentage 0
rating              0
rating_count        2
about_product       0
user_id             0
user_name           0
review_id           0
```

```
review_title      0  
review_content    0  
img_link          0  
product_link      0  
dtype: int64
```

```
[8]: train_data['rating_count'].fillna(0,inplace=True)
```

C:\Users\himan\AppData\Local\Temp\ipykernel_18348\1988824257.py:1:
FutureWarning: A value is trying to be set on a copy of a DataFrame or Series
through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
train_data['rating_count'].fillna(0,inplace=True)
```

```
[9]: train_data['rating'].fillna(0,inplace=True)
```

C:\Users\himan\AppData\Local\Temp\ipykernel_18348\3276964957.py:1:
FutureWarning: A value is trying to be set on a copy of a DataFrame or Series
through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work
because the intermediate object on which we are setting values always behaves as
a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using
'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value)
instead, to perform the operation inplace on the original object.

```
train_data['rating'].fillna(0,inplace=True)
```

```
[10]: train_data.isnull().sum()
```

```
[10]: product_id      0  
product_name     0  
category         0  
discounted_price 0  
actual_price     0  
discount_percentage 0  
rating           0  
rating_count     0
```

```
about_product          0
user_id                0
user_name              0
review_id              0
review_title           0
review_content          0
img_link               0
product_link            0
dtype: int64
```

```
[11]: train_data.duplicated().sum()
```

```
[11]: 0
```

```
[12]: train_data['product_id']
```

```
[12]: 0      B07JW9H4J1
1      B098NS6PVG
2      B096MSW6CT
3      B08HDJ86NZ
4      B08CF3B7N1
...
1460    B08L7J3T31
1461    B01M6453MB
1462    B009P2LIL4
1463    B00J5DYCCA
1464    B01486F4G6
Name: product_id, Length: 1465, dtype: object
```

```
[13]: train_data['user_id']
```

```
[13]: 0      AG3D604STAQKAY2UVGEUV46KN35Q,AHMY5CWJMMK5BJRBB...
1      AECPFYFQVRUWC3KGNLJI0REFP5LQ,AGYYVPDD7YG7FYNBX...
2      AGU3BBQ2V2DDAMOAKGFAWDDQ6QHA,AESFLDV2PT363T2AQ...
3      AEWAZDZZJLQUYVOVGBEUKSLXHQ5A,AG5HTSFRRE6NL3M5S...
4      AE3Q6KSUK5P75D5HFYHCRAOLODSA,AFUGIFH5ZAFXRDSZH...
...
1460    AHITFY6AHALOFOHOZE0C6XBP4FEA,AFRABBODZJZQB6Z4U...
1461    AFG5FM3NEMOL6BNFRV2NK5FNJCHQ,AGEINTRN6Z563RMLH...
1462    AGVPWCMAHYQWJOQKMUJN4DW3KM5Q,AF4Q3E66MY4SR7YQZ...
1463    AF2JQCLSCY3QJATWUNNHUSVUPNQQ,AFDMLUXC5LS5RXDJS...
1464    AFGW5PT3R6ZAVQR4Y5MWVAKBZAYA,AG7QNJ2SCS5VS5VYY...
Name: user_id, Length: 1465, dtype: object
```

```
[14]: train_data['product_id'] = train_data['product_id'].str.extract(r'(\d+)')
train_data['user_id'] = train_data['user_id'].str.extract(r'(\d+)')
```

```
[15]: train_data[['product_id','user_id']]
```

```
[15]: product_id user_id
0          07      3
1          098     3
2          096     3
3          08      5
4          08      3
...
1460        08      6
1461        01      5
1462        009     4
1463        00      2
1464      01486     5
```

[1465 rows x 2 columns]

```
[16]: train_data.head()
```

```
[16]: product_id          product_name \
0          07  Wayona Nylon Braided USB to Lightning Fast Cha...
1          098  Ambrane Unbreakable 60W / 3A Fast Charging 1.5...
2          096  Sounce Fast Phone Charging Cable & Data Sync U...
3          08  boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...
4          08  Portronics Konnect L 1.2M Fast Charging 3A 8 P...

                                              category discounted_price \
0  Computers&Accessories|Accessories&Peripherals|...            399
1  Computers&Accessories|Accessories&Peripherals|...            199
2  Computers&Accessories|Accessories&Peripherals|...            199
3  Computers&Accessories|Accessories&Peripherals|...            329
4  Computers&Accessories|Accessories&Peripherals|...            154

actual_price discount_percentage rating rating_count \
0        1,099                 64%    4.2      24,269
1         349                 43%    4.0      43,994
2        1,899                 90%    3.9      7,928
3         699                 53%    4.2      94,363
4         399                 61%    4.2      16,905

                                              about_product user_id \
0  High Compatibility : Compatible With iPhone 12...       3
1  Compatible with all Type C enabled devices, be...       3
2  Fast Charger& Data Sync -With built-in safet...       3
3  The boAt Deuce USB 300 2 in 1 cable is compati...       5
4  [CHARGE & SYNC FUNCTION]- This cable comes wit...       3

                                              user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...
```

```

1 ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...
2 Kunal,Himanshu,viswanath,sai niharka,saqib mal...
3 Omkar dhale,JD,HEMALATHA,Ajwadh a.,amar singh ...
4 rahuls6099,Swasat Borah,Ajay Wadke,Pranali,RVK...

                    review_id \
0 R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1...
1 RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...
2 R3J3EQQ9TZI5ZJ,R3E7WBGK7IDOKV,RWU79XKQ6I1QF,R2...
3 R3EEUZKKK9J36I,R3HJVYCLY0Y554,REDECAZ7AMPQC,R1...
4 R1BP4L2HH9TFUP,R16PVJEXKV6QZS,R2UPDB81N66T4P,R...

                    review_title \
0 Satisfied,Charging is really fast,Value for mo...
1 A Good Braided Cable for Your Type C Device,Go...
2 Good speed for earlier versions,Good Product,W...
3 Good product,Good one,Nice,Really nice product...
4 As good as original,Decent,Good one for second...

                    review_content \
0 Looks durable Charging is fine tooNo complains...
1 I ordered this cable to connect my phone to An...
2 Not quite durable and sturdy,https://m.media-a...
3 Good product,long wire,Charges good,Nice,I bou...
4 Bought this instead of original apple, does th...

                    img_link \
0 https://m.media-amazon.com/images/W/WEBP_40237...
1 https://m.media-amazon.com/images/W/WEBP_40237...
2 https://m.media-amazon.com/images/W/WEBP_40237...
3 https://m.media-amazon.com/images/I/41V5FtEWPk...
4 https://m.media-amazon.com/images/W/WEBP_40237...

                    product_link
0 https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
1 https://www.amazon.in/Ambrane-Unbreakable-Char...
2 https://www.amazon.in/Sounce-iPhone-Charging-C...
3 https://www.amazon.in/Deuce-300-Resistant-Tang...
4 https://www.amazon.in/Portronics-Konnect-POR-1...

```

```
[17]: num_users=train_data['user_id'].nunique()
num_items=train_data['product_id'].nunique()
num_ratings=train_data['rating'].nunique()
print(f"Number of unique users: {num_users}")
print(f"Number of unique products: {num_items}")
print(f"Number of unique ratings: {num_ratings}")
```

Number of unique users: 88

```
Number of unique products: 122
Number of unique ratings: 28
```

```
[18]: train_data.head(2)
```

```
[18]:   product_id          product_name \
0      07  Wayona Nylon Braided USB to Lightning Fast Cha...
1     098  Ambrane Unbreakable 60W / 3A Fast Charging 1.5...

                           category discounted_price \
0  Computers&Accessories|Accessories&Peripherals|...      399
1  Computers&Accessories|Accessories&Peripherals|...      199

   actual_price discount_percentage rating rating_count \
0        1,099             64%      4.2       24,269
1         349             43%      4.0       43,994

                           about_product user_id \
0  High Compatibility : Compatible With iPhone 12...      3
1  Compatible with all Type C enabled devices, be...      3

                           user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...
1  ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...

                           review_id \
0  R3HXWTOLRPONMF,R2AJM3LFTLZHF0,R6AQJGUP6P86,R1K...
1  RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...

                           review_title \
0  Satisfied,Charging is really fast,Value for mo...
1  A Good Braided Cable for Your Type C Device,Go...

                           review_content \
0  Looks durable Charging is fine tooNo complains...
1  I ordered this cable to connect my phone to An...

                           img_link \
0  https://m.media-amazon.com/images/W/WEBP_40237...
1  https://m.media-amazon.com/images/W/WEBP_40237...

                           product_link
0  https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
1  https://www.amazon.in/Ambrane-Unbreakable-Char...
```

```
[19]: import spacy
from spacy.lang.en.stop_words import STOP_WORDS
nlp = spacy.load("en_core_web_sm")
```

```

def clean_and_extract_tags(text):
    doc = nlp(text.lower())
    tags = [token.text for token in doc if token.text.isalnum() and token.text.isalpha() and token.text not in STOP_WORDS]
    return ', '.join(tags)
columns_to_extract_tags_from = ['category', 'product_name', 'about_product']
for column in columns_to_extract_tags_from:
    train_data[column] = train_data[column].apply(clean_and_extract_tags)

```

[20]: train_data.head(2)

```

[20]:   product_id          product_name category \
0        07  wayona, nylon, braided, usb, lightning, fast, ...
1       098  ambrane, unbreakable, 60w, 3a, fast, charging, ...

      discounted_price actual_price discount_percentage rating rating_count \
0            399         1,099                 64%      4.2        24,269
1            199         349                 43%      4.0        43,994

      about_product user_id \
0  high, compatibility, compatible, iphone, 12, 1...      3
1  compatible, type, c, enabled, devices, android...      3

      user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...
1  ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...

      review_id \
0  R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...
1  RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...

      review_title \
0  Satisfied,Charging is really fast,Value for mo...
1  A Good Braided Cable for Your Type C Device,Go...

      review_content \
0  Looks durable Charging is fine tooNo complains...
1  I ordered this cable to connect my phone to An...

      img_link \
0  https://m.media-amazon.com/images/W/WEBP_40237...
1  https://m.media-amazon.com/images/W/WEBP_40237...

      product_link
0  https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
1  https://www.amazon.in/Ambrane-Unbreakable-Char...

```

```
[21]: train_data['Tags'] = train_data[columns_to_extract_tags_from].apply(lambda row:  
    ' '.join(row), axis=1)
```

```
[22]: train_data.head(2)
```

```
[22]:   product_id          product_name category  \  
0        07  wayona, nylon, braided, usb, lightning, fast, ...  
1       098  ambrane, unbreakable, 60w, 3a, fast, charging,...  
  
   discounted_price actual_price discount_percentage rating rating_count  \  
0            399        1,099                 64%      4.2      24,269  
1            199         349                 43%      4.0     43,994  
  
   about_product user_id  \  
0  high, compatibility, compatible, iphone, 12, 1...      3  
1  compatible, type, c, enabled, devices, android...      3  
  
   user_name  \  
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...  
1  ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...  
  
   review_id  \  
0  R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...  
1  RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...  
  
   review_title  \  
0  Satisfied,Charging is really fast,Value for mo...  
1  A Good Braided Cable for Your Type C Device,Go...  
  
   review_content  \  
0  Looks durable Charging is fine tooNo complains...  
1  I ordered this cable to connect my phone to An...  
  
   img_link  \  
0  https://m.media-amazon.com/images/W/WEBP_40237...  
1  https://m.media-amazon.com/images/W/WEBP_40237...  
  
   product_link  \  
0  https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...  
1  https://www.amazon.in/Ambrane-Unbreakable-Char...  
  
   Tags
```

```
0 , wayona, nylon, braided, usb, lightning, fast...  
1 , ambrane, unbreakable, 60w, 3a, fast, chargin...
```

```
[23]: train_data['rating'].dtype
```

```
[23]: dtype('O')

[24]: train_data['rating'] = pd.to_numeric(train_data['rating'], errors='coerce')

[25]: train_data['rating'].dtype

[25]: dtype('float64')

[26]: average_ratings = train_data.
    ↪groupby(['product_name', 'rating_count', 'img_link'])['rating'].mean().
    ↪reset_index()

[27]: average_ratings.head()

[27]:
      product_name rating_count \
0          1000, watt, room, heater, fan, india           11
1  10k, 8k, 4k, hdmi, cable, certified, 48gbps, 1...        3,664
2  10werun, bluetooth, smartwatch, wireless, fitn...         412
3  3, m, post, sticky, note, cube, 200, sheets, 4...        7,429
4  3, m, scotch, double, sided, heavy, duty, m, h...       14,778

      img_link  rating
0  https://m.media-amazon.com/images/W/WEBP_40237...     4.5
1  https://m.media-amazon.com/images/W/WEBP_40237...     4.6
2  https://m.media-amazon.com/images/I/41YwW+0-SK...     4.1
3  https://m.media-amazon.com/images/W/WEBP_40237...     4.4
4  https://m.media-amazon.com/images/W/WEBP_40237...     3.9

[28]: top_rated_items = average_ratings.sort_values(by='rating', ascending=False)

[29]: top_rated_items.head()

[29]:
      product_name rating_count \
82    amazon, basics, wireless, mouse, ghz, connecti...           23
1084  redtech, usb, c, lightning, cable, ft, apple, ...            0
1235  syncwire, ltg, usb, cable, fast, charging, com...            5
1233  swiffer, instant, electric, water, heater, fau...        53,803
624    instant, pot, air, fryer, vortex, 2qt, touch, ...        3,964

      img_link  rating
82  https://m.media-amazon.com/images/I/31+Rg6Z46d...     5.0
1084 https://m.media-amazon.com/images/I/31-q0xhaTA...     5.0
1235 https://m.media-amazon.com/images/I/3170oQfs1g...     5.0
1233 https://m.media-amazon.com/images/I/41+t2HWvwF...     4.8
624  https://m.media-amazon.com/images/W/WEBP_40237...     4.8

[30]: rating_base_recommendation=top_rated_items.head(10)
```

```
[31]: rating_base_recommendation
```

```
[31]:
```

```
          product_name rating_count \
82      amazon, basics, wireless, mouse, ghz, connecti...      23
1084     redtech, usb, c, lightning, cable, ft, apple, ...      0
1235     syncwire, ltg, usb, cable, fast, charging, com...      5
1233     swiffer, instant, electric, water, heater, fau...  53,803
624      instant, pot, air, fryer, vortex, 2qt, touch, ...  3,964
896      oratech, coffee, frother, electric, milk, frot...      28
462      figment, handheld, milk, frother, rechargeable...  1,729
832      multifunctional, 2, 1, electric, egg, boiling,...  2,300
320      campfire, spring, chef, prolix, instant, porta...  2,591
1183     sony, bravia, 164, cm, 65, inches, 4k, ultra, ...  5,935

          img_link rating
82      https://m.media-amazon.com/images/I/31+Rg6Z46d...      5.0
1084     https://m.media-amazon.com/images/I/31-q0xhaTA...      5.0
1235     https://m.media-amazon.com/images/I/3170oQfs1g...      5.0
1233     https://m.media-amazon.com/images/I/41+t2HWvwF...      4.8
624      https://m.media-amazon.com/images/W/WEBP_40237...      4.8
896      https://m.media-amazon.com/images/W/WEBP_40237...      4.8
462      https://m.media-amazon.com/images/I/51oN+8Zs5Y...      4.7
832      https://m.media-amazon.com/images/W/WEBP_40237...      4.7
320      https://m.media-amazon.com/images/W/WEBP_40237...      4.7
1183     https://m.media-amazon.com/images/I/51uVckL1jR...      4.7
```

```
[32]: train_data.head(1)
```

```
[32]:
```

```
          product_id          product_name category \
0            07    wayona, nylon, braided, usb, lightning, fast, ...
                                                     discounted_price actual_price discount_percentage rating rating_count \
0              399           1,099                 64%        4.2       24,269

          about_product user_id \
0  high, compatibility, compatible, iphone, 12, 1...        3

          user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...

          review_id \
0  R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...

          review_title \
0  Satisfied,Charging is really fast,Value for mo...

          review_content \
0  
```

```

0 Looks durable Charging is fine tooNo complains...
                                         img_link \
0 https://m.media-amazon.com/images/W/WEBP_40237...
                                         product_link \
0 https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
                                         Tags
0 , wayona, nylon, braided, usb, lightning, fast...

```

```
[34]: output_path = "clean_data.csv"

train_data.to_csv(output_path, index=False)

from IPython.display import FileLink
FileLink(output_path)
```

[34]: C:\Users\himan\clean_data.csv

```
[35]: train_data.head()
```

```

[35]:   product_id                  product_name category \
0      07  wayona, nylon, braided, usb, lightning, fast, ...
1      098 ambrane, unbreakable, 60w, 3a, fast, charging, ...
2      096 sounce, fast, phone, charging, cable, data, sy...
3      08  boat, deuce, usb, 300, 2, 1, type, c, micro, u...
4      08 portronics, konnect, 1, m, fast, charging, 3a, ...

   discounted_price actual_price discount_percentage  rating rating_count \
0            399       1,099             64%        4.2      24,269
1            199        349             43%        4.0      43,994
2            199       1,899             90%        3.9      7,928
3            329        699             53%        4.2     94,363
4            154        399             61%        4.2     16,905

   about_product user_id \
0  high, compatibility, compatible, iphone, 12, 1...      3
1  compatible, type, c, enabled, devices, android...      3
2  fast, charger, data, built, safety, proctectio...      3
3  boat, deuce, usb, 300, 2, 1, cable, compatible...      5
4  charge, sync, cable, comes, charging, data, sy...      3

   user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...
1  ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Plac...
2  Kunal,Himanshu,viswanath,sai niharka,saqib mal...
```

```
3 Omkar dhale,JD,HEMALATHA,Ajwadh a.,amar singh ...
4 rahuls6099,Swasat Borah,Ajay Wadke,Pranali,RVK...
```

```
review_id \
0 R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...
1 RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...
2 R3J3EQQ9T5JZJ,R3E7WBGK7IDOKV,RWU79XKQ6I1QF,R2...
3 R3EEUZKKK9J36I,R3HJVYCLY0Y554,REDECAZ7AMPQC,R1...
4 R1BP4L2HH9TFUP,R16PVJEXKV6QZS,R2UPDB81N66T4P,R...
```

```
review_title \
0 Satisfied,Charging is really fast,Value for mo...
1 A Good Braided Cable for Your Type C Device,Go...
2 Good speed for earlier versions,Good Product,W...
3 Good product,Good one,Nice,Really nice product...
4 As good as original,Decent,Good one for second...
```

```
review_content \
0 Looks durable Charging is fine tooNo complains...
1 I ordered this cable to connect my phone to An...
2 Not quite durable and sturdy,https://m.media-a...
3 Good product,long wire,Charges good,Nice,I bou...
4 Bought this instead of original apple, does th...
```

```
img_link \
0 https://m.media-amazon.com/images/W/WEBP_40237...
1 https://m.media-amazon.com/images/W/WEBP_40237...
2 https://m.media-amazon.com/images/W/WEBP_40237...
3 https://m.media-amazon.com/images/I/41V5FtEWpk...
4 https://m.media-amazon.com/images/W/WEBP_40237...
```

```
product_link \
0 https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
1 https://www.amazon.in/Ambrane-Unbreakable-Char...
2 https://www.amazon.in/Sounce-iPhone-Charging-C...
3 https://www.amazon.in/Deuce-300-Resistant-Tang...
4 https://www.amazon.in/Portronics-Konnect-POR-1...
```

```
Tags
0 , wayona, nylon, braided, usb, lightning, fast...
1 , ambrane, unbreakable, 60w, 3a, fast, chargin...
2 , sounce, fast, phone, charging, cable, data, ...
3 , boat, deuce, usb, 300, 2, 1, type, c, micro, ...
4 , portronics, konnect, 1, m, fast, charging, 3...
```

```
[43]: train_data['product_name'][45]
```

```
[43]: 'wecool, nylon, braided, multifunction, fast, charging, cable, android,  
smartphone, ios, type, c, usb, devices, 3, 1, charging, cable, 3a, 3, feet,  
black'
```

```
[33]: from sklearn.feature_extraction.text import TfidfVectorizer  
from sklearn.metrics.pairwise import cosine_similarity  
  
tfidf_vectorizer=TfidfVectorizer(stop_words='english')  
tfidf_matrix_content=tfidf_vectorizer.fit_transform(train_data['Tags'])  
↳  
cosine_similarities_content=cosine_similarity(tfidf_matrix_content,tfidf_matrix_content)
```

```
[34]: cosine_similarities_content
```

```
[34]: array([[1.          , 0.12528853, 0.26806542, ..., 0.03798299, 0.0422975 ,  
          0.00321638],  
         [0.12528853, 1.          , 0.13844415, ..., 0.          , 0.01001183,  
          0.03144304],  
         [0.26806542, 0.13844415, 1.          , ..., 0.02536216, 0.02880817,  
          0.          ],  
         ...,  
         [0.03798299, 0.          , 0.02536216, ..., 1.          , 0.11094301,  
          0.01908344],  
         [0.0422975 , 0.01001183, 0.02880817, ..., 0.11094301, 1.          ,  
          0.03234351],  
         [0.00321638, 0.03144304, 0.          , ..., 0.01908344, 0.03234351,  
          1.          ]])
```

```
[35]: train_data['product_name'][0]
```

```
[35]: 'wayona, nylon, braided, usb, lightning, fast, charging, data, sync, cable,  
compatible, iphone, 13, x, 8, 7, 6, 5, ipad, air, pro, mini, 3, ft, pack, 1,  
grey'
```

```
[36]: item_name='wayona, nylon, braided, usb, lightning, fast, charging, data, sync, cable,  
compatible, iphone, 13, x, 8, 7, 6, 5, ipad, air, pro, mini, 3, ft, pack, 1, grey'  
  
item_index=train_data[train_data['product_name']==item_name].index[0]
```

```
[37]: similar_items = list(enumerate(cosine_similarities_content[item_index]))
```

```
[38]: similar_items = sorted(similar_items, key=lambda x:x[1], reverse=True)  
top_similar_items = similar_items[1:10]  
  
recommended_items_indices = [x[0] for x in top_similar_items]
```

```
[39]: train_data.  
↳ iloc[recommended_items_indics][['product_name', 'rating_count', 'category']]
```

```
[39]:
```

	product_name	rating_count	category
369	wayona, nylon, braided, usb, lightning, fast, ...	24,270	
614	wayona, nylon, braided, usb, lightning, fast, ...	24,269	
220	wayona, nylon, braided, usb, syncing, charging...	24,269	
42	wayona, nylon, braided, 3a, lightning, usb, sy...	24,269	
89	wayona, nylon, braided, 2, pack, lightning, fa...	24,269	
80	wayona, usb, nylon, braided, data, sync, charg...	24,269	
106	wayona, nylon, braided, 2, m, 6, ft, fast, cha...	24,269	
166	wayona, nylon, braided, lightning, usb, data, ...	13,120	
208	myvn, ltg, usb, fast, charging, data, sync, us...	2,249	

```
[40]: from sklearn.feature_extraction.text import TfidfVectorizer  
from sklearn.metrics.pairwise import cosine_similarity  
  
def content_based_recommendations(train_data, item_name, top_n=10):  
    if item_name not in train_data['product_name'].values:  
        print(f"Item '{item_name}' not found in the training data.")  
        return pd.DataFrame()  
  
    tfidf_vectorizer = TfidfVectorizer(stop_words='english')  
    tfidf_matrix_content = tfidf_vectorizer.fit_transform(train_data['Tags'])  
    cosine_similarities_content = cosine_similarity(tfidf_matrix_content, □  
    ↳ tfidf_matrix_content)  
  
    item_index = train_data[train_data['product_name'] == item_name].index[0]  
    similar_items = list(enumerate(cosine_similarities_content[item_index]))  
    similar_items = sorted(similar_items, key=lambda x: x[1], reverse=True)  
    top_similar_items = similar_items[1:top_n+1]  
  
    recommended_item_indices = [x[0] for x in top_similar_items]  
    recommended_items_details = train_data.  
    ↳ iloc[recommended_item_indices][['product_name', 'rating', 'rating_count', □  
    ↳ 'img_link', 'review_content']]  
  
    return recommended_items_details
```

```
[41]: train_data['product_name'][3]
```

```
[41]: 'boat, deuce, usb, 300, 2, 1, type, c, micro, usb, stress, resistant, tangle,  
free, sturdy, cable, 3a, fast, charging, 480mbps, data, transmission, 10000,  
bends, lifespan, extended, m, red'
```

```
[42]:
```

```
item_name='boat, deuce, usb, 300, 2, 1, type, c, micro, usb, stress, resistant, u
↳tangle, free, sturdy, cable, 3a, fast, charging, 480mbps, data, u
↳transmission, 10000, bends, lifespan, extended, m, red'
```

```
content_based_rec=content_based_recommendations(train_data,item_name,10)
content_based_rec
```

[42]:

```
product_name rating rating_count \
392 boat, deuce, usb, 300, 2, 1, type, c, micro, u... 4.2 94,364
628 boat, deuce, usb, 300, 2, 1, type, c, micro, u... 4.2 94,364
92 boat, deuce, usb, 300, 2, 1, type, c, micro, u... 4.2 94,363
113 boat, type, c, a750, stress, resistant, tangle... 4.1 1,780
74 boat, type, c, a750, stress, resistant, tangle... 4.1 1,780
985 boat, type, c, a750, stress, resistant, tangle... 4.1 1,780
6 boat, micro, usb, 55, tangle, free, sturdy, mi... 4.1 15,188
418 boat, micro, usb, 55, tangle, free, sturdy, mi... 4.1 15,189
658 boat, micro, usb, 55, tangle, free, sturdy, mi... 4.1 15,189
18 boat, type, c, a325, tangle, free, sturdy, typ... 4.1 13,045

img_link \
392 https://m.media-amazon.com/images/I/41V5FtEWPk...
628 https://m.media-amazon.com/images/W/WEBP_40237...
92 https://m.media-amazon.com/images/W/WEBP_40237...
113 https://m.media-amazon.com/images/W/WEBP_40237...
74 https://m.media-amazon.com/images/I/41jlh3c7Ub...
985 https://m.media-amazon.com/images/I/41jlh3c7Ub...
6 https://m.media-amazon.com/images/W/WEBP_40237...
418 https://m.media-amazon.com/images/I/41jlwEZpa5...
658 https://m.media-amazon.com/images/W/WEBP_40237...
18 https://m.media-amazon.com/images/W/WEBP_40237...

review_content
392 Good product,long wire,Charges good,Nice,I bou...
628 Good product,long wire,Charges good,Nice,I bou...
92 Good product,long wire,Charges good,Nice,I bou...
113 Yeap this cable support 65w fast charging test...
74 Yeap this cable support 65w fast charging test...
985 Yeap this cable support 65w fast charging test...
6 Build quality is good and it is comes with 2 y...
418 Build quality is good and it is comes with 2 y...
658 Build quality is good and it is comes with 2 y...
18 Check for offera before buying,18w charging ca...
```

[43]: train_data.head(1)

```
product_id product_name category \
0 07 wayona, nylon, braided, usb, lightning, fast, ...
```

```

discounted_price actual_price discount_percentage rating rating_count \
0           399        1,099          64%      4.2      24,269

                                about_product user_id \
0  high, compatibility, compatible, iphone, 12, 1...      3

                                user_name \
0  Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jasp...

                                review_id \
0  R3HXWTOLRPONMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...

                                review_title \
0  Satisfied,Charging is really fast,Value for mo...

                                review_content \
0  Looks durable Charging is fine tooNo complains...

                                img_link \
0  https://m.media-amazon.com/images/W/WEBP_40237...

                                product_link \
0  https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...

                                Tags
0  , wayona, nylon, braided, usb, lightning, fast...

```

[44]: `user_item_matrix = train_data.pivot_table(index='user_id',
 columns='product_id', values='rating',aggfunc='mean').fillna(0).astype(int)`

[45]: `user_item_matrix`

```

[45]: product_id  0   00   002   003   004   005   006   0073   008   0083   ...   0974   0977   \
user_id
2           4   4     4     4     0     0     0     0     0     0   ...   0     0
22          0   0     0     0     0     0     0     0     0     0   ...   0     0
222         4   0     0     0     0     0     0     0     0     0   ...   0     0
226         0   0     0     0     0     0     0     0     0     0   ...   0     0
23           0   4     0     0     0     0     0     0     0     0   ...   0     0
...
757          0   0     0     0     0     0     0     0     0     0   ...   0     0
76           0   0     0     0     0     0     0     0     0     0   ...   0     0
77           0   4     0     0     0     0     0     0     0     0   ...   0     0
774          0   0     0     0     0     0     0     0     0     0   ...   0     0
775          0   0     0     0     0     0     0     0     0     0   ...   0     0

```

```

product_id 0978 098 0981 0989 099 0993 09939 0994
user_id
2          4   3   0   4   3   0   0   0
22         0   0   0   0   0   0   0   0
222        0   0   0   0   0   0   0   0
226        0   0   0   0   0   0   0   0
23         0   0   0   0   0   0   0   0
...
757        0   0   0   0   0   0   0   0
76         0   0   0   0   0   0   0   0
77         0   0   0   0   0   0   0   0
774        0   0   0   0   0   0   0   0
775        0   4   0   0   0   0   0   0

```

[88 rows x 122 columns]

```
[46]: user_similarity = cosine_similarity(user_item_matrix)
user_similarity
```

```
[46]: array([[1.          , 0.14545455, 0.18181818, ... , 0.3612361 , 0.18181818,
       0.13636364],
       [0.14545455, 1.          , 0.          , ... , 0.42385035, 0.          ,
       0.          ],
       [0.18181818, 0.          , 1.          , ... , 0.          , 0.          ,
       0.          ],
       ... ,
       [0.3612361 , 0.42385035, 0.          , ... , 1.          , 0.39735971,
       0.          ],
       [0.18181818, 0.          , 0.          , ... , 0.39735971, 1.          ,
       0.          ],
       [0.13636364, 0.          , 0.          , ... , 0.          , 0.          ,
       1.          ]])
```

```
[47]: user_item_matrix.index = user_item_matrix.index.astype(int)
```

```
[48]: target_user_id = 2
target_user_index = user_item_matrix.index.get_loc(target_user_id)
```

```
[49]: user_similarities = user_similarity[target_user_index]
```

```
similar_user_indices = user_similarities.argsort()[:-1][1:]
```

```
[50]: recommend_items = []

for user_index in similar_user_indices:
    rated_by_similar_user = user_item_matrix.iloc[user_index]
```

```

    not_rated_by_target_user = (rated_by_similar_user==0) & (user_item_matrix.
    ↪iloc[target_user_index]==0)

    recommend_items.extend(user_item_matrix.columns[not_rated_by_target_user] [:
    ↪10])

recommended_items_details = train_data[train_data['product_id']].
↪isin(recommend_items)][['product_name','rating','rating_count','img_link','review_content']]

```

[51]: recommended_items_details.head(10)

	product_name	rating	rating_count	\
8	tp, link, usb, wifi, adapter, wn725n, n150, wi...	4.2	1,79,691	
33	samsung, original, type, c, c, cable, feet, 1,...	4.3	355	
56	tp, link, wifi, dongle, 300, mbps, mini, wirel...	4.2	1,79,691	
90	gizga, essentials, usb, wifi, adapter, pc, 150...	3.6	10,134	
148	bluerigger, digital, optical, audio, toslink, ...	4.2	30,023	
189	bluerigger, digital, optical, audio, toslink, ...	4.2	30,023	
239	amazon, basics, 16, gauge, speaker, wire, 50, ...	4.4	12,091	
302	technotech, high, speed, hdmi, cable, 5, meter...	4.2	25	
431	goldmedal, curve, plus, 202042, plastic, spice...	4.5	11,339	
588	sandisk, cruzer, blade, 32, gb, usb, flash, drive	4.3	2,53,105	
	img_link \			
8	https://m.media-amazon.com/images/W/WEBP_40237...			
33	https://m.media-amazon.com/images/W/WEBP_40237...			
56	https://m.media-amazon.com/images/I/41jxZkzNcn...			
90	https://m.media-amazon.com/images/I/31mgo4D-kP...			
148	https://m.media-amazon.com/images/I/41CF6GtnpK...			
189	https://m.media-amazon.com/images/W/WEBP_40237...			
239	https://m.media-amazon.com/images/W/WEBP_40237...			
302	https://m.media-amazon.com/images/I/51UUUmio53P...			
431	https://m.media-amazon.com/images/I/31MmkM8HTi...			
588	https://m.media-amazon.com/images/I/310mw9KTJv...			
	review_content			
8	I use this to connect an old PC to internet. I...			
33	Buy it,Received in good condition and is worki...			
56	I use this to connect an old PC to internet. I...			
90	Was working fine with window 10 old computer b...			
148	Using with my Xbox and it's working perfectly,...			
189	Using with my Xbox and it's working perfectly,...			
239	Good Packaging.. Wire Quality is Good.,Very go...			
302	Ok,Quality perfect , perfect 5m, must buy,Ok,E...			
431	Product is just as described and does what it ...			
588	Very chip very good,Really happy to buy this p...			

```
[52]: def collaborative_filtering_recommendations(train_data, target_user_id, top_n):
    user_item_matrix = train_data.pivot_table(index='user_id', u
    ↵columns='product_id', values='rating', aggfunc='mean').fillna(0)

    user_similarity = cosine_similarity(user_item_matrix)
    target_user_index = user_item_matrix.index.get_loc(target_user_id)

    user_similarities = user_similarity[target_user_index]
    similar_users_indices = user_similarities.argsort()[:-1][1:]
    recommended_items = []

    for user_index in similar_users_indices:
        rated_by_similar_user = user_item_matrix.iloc[user_index]
        not_rated_by_target_user = (rated_by_similar_user == 0) & u
        ↵(user_item_matrix.iloc[target_user_index] == 0)

        recommended_items.extend(user_item_matrix.
        ↵columns[not_rated_by_target_user][:top_n])

    recommended_items_details = train_data[train_data['product_id']].
    ↵isin(recommended_items)][['product_name', 'rating', 'rating_count', u
    ↵'img_link', 'review_content']]

    return recommended_items_details.head(10)
collaborative_filtering_rec = u
    ↵collaborative_filtering_recommendations(train_data, '2', 10)
collaborative_filtering_rec
```

	product_name	rating	rating_count	\
8	tp, link, usb, wifi, adapter, wn725n, n150, wi...	4.2	1,79,691	
33	samsung, original, type, c, c, cable, feet, 1,...	4.3	355	
56	tp, link, wifi, dongle, 300, mbps, mini, wirel...	4.2	1,79,691	
90	gizga, essentials, usb, wifi, adapter, pc, 150...	3.6	10,134	
148	bluerigger, digital, optical, audio, toslink, ...	4.2	30,023	
189	bluerigger, digital, optical, audio, toslink, ...	4.2	30,023	
239	amazon, basics, 16, gauge, speaker, wire, 50, ...	4.4	12,091	
302	technotech, high, speed, hdmi, cable, 5, meter...	4.2	25	
431	goldmedal, curve, plus, 202042, plastic, spice...	4.5	11,339	
588	sandisk, cruzer, blade, 32, gb, usb, flash, drive	4.3	2,53,105	
		img_link	\	
8	https://m.media-amazon.com/images/W/WEBP_40237...			
33	https://m.media-amazon.com/images/W/WEBP_40237...			
56	https://m.media-amazon.com/images/I/41jxZkzNcn...			
90	https://m.media-amazon.com/images/I/31mgo4D-kP...			
148	https://m.media-amazon.com/images/I/41CF6GtnpK...			
189	https://m.media-amazon.com/images/W/WEBP_40237...			
239	https://m.media-amazon.com/images/W/WEBP_40237...			

```

302 https://m.media-amazon.com/images/I/51UUUmio53P...
431 https://m.media-amazon.com/images/I/31MmkM8HTi...
588 https://m.media-amazon.com/images/I/310mw9KTJv...

```

	review_content
8	I use this to connect an old PC to internet. I...
33	Buy it,Received in good condition and is worki...
56	I use this to connect an old PC to internet. I...
90	Was working fine with window 10 old computer b...
148	Using with my Xbox and it's working perfectly,...
189	Using with my Xbox and it's working perfectly,...
239	Good Packaging.. Wire Quality is Good.,Very go...
302	Ok,Quality perfect , perfect 5m, must buy,Ok,E...
431	Product is just as described and does what it ...
588	Very chip very good,Really happy to buy this p...

```
[61]: def hybrid_recommendations(train_data,target_user_id, item_name, top_n=10):

    content_based_rec = content_based_recommendations(train_data,item_name, top_n)
    target_user_id=str(target_user_id)
    collaborative_filtering_rec = collaborative_filtering_recommendations(train_data,target_user_id, top_n)

    hybrid_rec = pd.concat([collaborative_filtering_rec,content_based_rec]).drop_duplicates()
    hybrid_rec = hybrid_rec.sort_values(by='rating', ascending=False)
    return hybrid_rec.head(10)
```

```
[57]: train_data['product_name'][0]
```

```
[57]: 'wayona, nylon, braided, usb, lightning, fast, charging, data, sync, cable,
compatible, iphone, 13, x, 8, 7, 6, 5, ipad, air, pro, mini, 3, ft, pack, 1,
grey'
```

```
[62]: target_user_id=3
item_name='wayona, nylon, braided, usb, lightning, fast, charging, data, sync,
cable, compatible, iphone, 13, x, 8, 7, 6, 5, ipad, air, pro, mini, 3, ft,
pack, 1, grey'
hybrid_rec=hybrid_recommendations(train_data,target_user_id,item_name,top_n=10)
hybrid_rec
```

	product_name	rating	rating_count	\
324	redtech, usb, c, lightning, cable, ft, apple, ..	5.0	0	
615	duracell, ultra, alkaline, aa, battery, 8, pcs	4.5	28,030	
734	logitech, m235, wireless, mouse, 1000, dpi, op...	4.5	54,405	
65	amazon, basics, high, speed, hdmi, cable, 6, f...	4.4	4,26,973	

47	amazon, basics, high, speed, hdmi, cable, 6, f...	4.4	4,26,973
239	amazon, basics, 16, gauge, speaker, wire, 50, ...	4.4	12,091
721	duracell, plus, aaa, rechargeable, batteries, ...	4.3	27,201
724	logitech, b100, wired, usb, mouse, 3, yr, warr...	4.3	31,534
302	technotech, high, speed, hdmi, cable, 5, meter...	4.2	25
614	wayona, nylon, braided, usb, lightning, fast, ...	4.2	24,269

img_link \

324	https://m.media-amazon.com/images/I/31-q0xhaTA...
615	https://m.media-amazon.com/images/W/WEBP_40237...
734	https://m.media-amazon.com/images/I/31CtVvtFt+...
65	https://m.media-amazon.com/images/I/41eJqkFjCR...
47	https://m.media-amazon.com/images/I/41bCxnHksn...
239	https://m.media-amazon.com/images/W/WEBP_40237...
721	https://m.media-amazon.com/images/I/418YrbHVLC...
724	https://m.media-amazon.com/images/I/31iFF1Kbkp...
302	https://m.media-amazon.com/images/I/51UUUmio53P...
614	https://m.media-amazon.com/images/W/WEBP_40237...

review_content

324	Quick delivery.Awesome ProductPacking was good...
615	Nice one, tried on remote mouse and found work...
734	It's little small for big hands. But best avai...
65	I am using it for 14 days now. The experience ...
47	I am using it for 14 days now. The experience ...
239	Good Packaging.. Wire Quality is Good.,Very go...
721	Works good,Bought it to replace my Phillips QT...
724	Liked this Product, https://m.media-amazon.com/...
302	Ok,Quality perfect , perfect 5m, must buy,Ok,E...
614	Looks durable Charging is fine tooNo complains...

[]: