

```
In [76]: 1 import numpy as np
2 import pandas as pd
3 import matplotlib.pyplot as plt
4 import seaborn as sns
```

```
In [77]: 1 df = pd.read_csv('amazon.csv')
```

```
In [78]: 1 df.head()
```

Out[78]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count	about
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Cha...	Computers&Accessories Accessories&Peripherals ...	₹399	₹1,099	64%	4.2	24,269	Con C W
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging 1.5...	Computers&Accessories Accessories&Peripherals ...	₹199	₹349	43%	4.0	43,994	C with dev
2	B096MSW6CT	Source Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	₹199	₹1,899	90%	3.9	7,928	Char Sy built
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	₹329	₹699	53%	4.2	94,363	D
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	₹154	₹399	61%	4.2	16,905	FU cc

```
In [79]: 1 df.shape
```

Out[79]: (1465, 16)

```
In [80]: 1 df.describe()
```

Out[80]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count	about
count	1465	1465		1465	1465	1465	1465	1465	1463
unique	1351	1337		211	550	449	92	28	1143
top	B07JW9H4J1	Fire-Boltt Ninja Call Pro Plus 1.83" Smart Wat...	Computers&Accessories Accessories&Peripherals ...	₹199	₹999	50%	4.1	9,378	
freq	3	5		233	53	120	56	244	9

In [81]: 1 df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1465 entries, 0 to 1464
Data columns (total 16 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   product_id       1465 non-null   object  
 1   product_name     1465 non-null   object  
 2   category         1465 non-null   object  
 3   discounted_price 1465 non-null   object  
 4   actual_price     1465 non-null   object  
 5   discount_percentage 1465 non-null   object  
 6   rating           1465 non-null   object  
 7   rating_count     1463 non-null   object  
 8   about_product    1465 non-null   object  
 9   user_id          1465 non-null   object  
 10  user_name        1465 non-null   object  
 11  review_id        1465 non-null   object  
 12  review_title     1465 non-null   object  
 13  review_content   1465 non-null   object  
 14  img_link         1465 non-null   object  
 15  product_link     1465 non-null   object  
dtypes: object(16)
memory usage: 183.2+ KB
```

In [82]: 1 df.dtypes

```
product_id      object
product_name    object
category        object
discounted_price object
actual_price    object
discount_percentage object
rating          object
rating_count    object
about_product   object
user_id         object
user_name        object
review_id        object
review_title     object
review_content   object
img_link         object
product_link    object
dtype: object
```

In [83]: 1 df.columns

```
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')
```

1. What is the average rating for each product category?

```
1 df['rating'] = pd.to_numeric(df['rating'], errors='coerce')
2 df[df['rating'].isna()]
```

Out[84]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count
1279	B08L12N5H1	Eureka Forbes car Vac 100 Watts Powerful Sucti...	Home&Kitchen Kitchen&HomeAppliances Vacuum,Cle...	₹2,099	₹2,499	16%	NaN	992

```
In [85]: 1 avg_rating_per_category = df.groupby('category')['rating'].mean().reset_index()
2 avg_rating_per_category
```

Out[85]:

	category	rating
0	Car&Motorbike CarAccessories InteriorAccessori...	3.800000
1	Computers&Accessories Accessories&Peripherals ...	4.150000
2	Computers&Accessories Accessories&Peripherals ...	3.500000
3	Computers&Accessories Accessories&Peripherals ...	3.600000
4	Computers&Accessories Accessories&Peripherals ...	4.050000
...
206	OfficeProducts OfficePaperProducts Paper Stati...	4.250000
207	OfficeProducts OfficePaperProducts Paper Stati...	4.150000
208	OfficeProducts OfficePaperProducts Paper Stati...	4.300000
209	OfficeProducts OfficePaperProducts Paper Stati...	4.133333
210	Toys&Games Arts&Crafts Drawing&PaintingSupplie...	4.300000

211 rows × 2 columns

2. What are the top rating_count products by category?

```
In [86]: 1 df.head()
```

Out[86]:

	product_id	product_name	category	discounted_price	actual_price	discount_percentage	rating	rating_count	about
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Cha...	Computers&Accessories Accessories&Peripherals ...	₹399	₹1,099	64%	4.2	24,269	Con C W
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging 1.5...	Computers&Accessories Accessories&Peripherals ...	₹199	₹349	43%	4.0	43,994	C with dev
2	B096MSW6CT	Sounce Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	₹199	₹1,899	90%	3.9	7,928	Char Sy built
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	₹329	₹699	53%	4.2	94,363	D
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	₹154	₹399	61%	4.2	16,905	[C FU cc

```
In [87]: 1 # Convert 'rating_count' column to object data type
2 df['rating_count'] = df['rating_count'].astype(str)
```

```
In [88]: 1 df_filtered = df.dropna(subset=['rating_count'])
2 top_products_by_category = df_filtered.groupby('category').apply(lambda x: x.sort_values(by='rating_count', ascending=False)
3 top_products_by_category.reset_index(drop=True, inplace=True)
4 top_products_by_category[['category', 'product_name', 'rating_count']]
5
```

Out[88]:

	category	product_name	rating_count
0	Car&Motorbike CarAccessories InteriorAccessory...	Reffair AX30 [MAX] Portable Air Purifier for C...	1,118
1	Computers&Accessories Accessories&Peripherals ...	rts [2 Pack] Mini USB C Type C Adapter Plug, T...	4,426
2	Computers&Accessories Accessories&Peripherals ...	Hp Wired On Ear Headphones With Mic With 3.5 M...	7,222
3	Computers&Accessories Accessories&Peripherals ...	HUMBLE Dynamic Lapel Collar Mic Voice Recordin...	2,804
4	Computers&Accessories Accessories&Peripherals ...	Zebronics Zeb Wonderbar 10 USB Powered 2.0 Com...	7,352
...
206	OfficeProducts OfficePaperProducts Paper Stati...	Classmate Octane Neon- Blue Gel Pens(Pack of 5...	5,792
207	OfficeProducts OfficePaperProducts Paper Stati...	Pilot Frixion Clicker Roller Pen (Blue), (9000...	6,199
208	OfficeProducts OfficePaperProducts Paper Stati...	Parker Classic Gold Gold Trim Ball Pen	5,719
209	OfficeProducts OfficePaperProducts Paper Stati...	Pentonic Multicolor Ball Point Pen, Pack of 10	4,308
210	Toys&Games Arts&Crafts Drawing&PaintingSupplie...	Faber-Castell Connector Pen Set - Pack of 25 (...	15,867

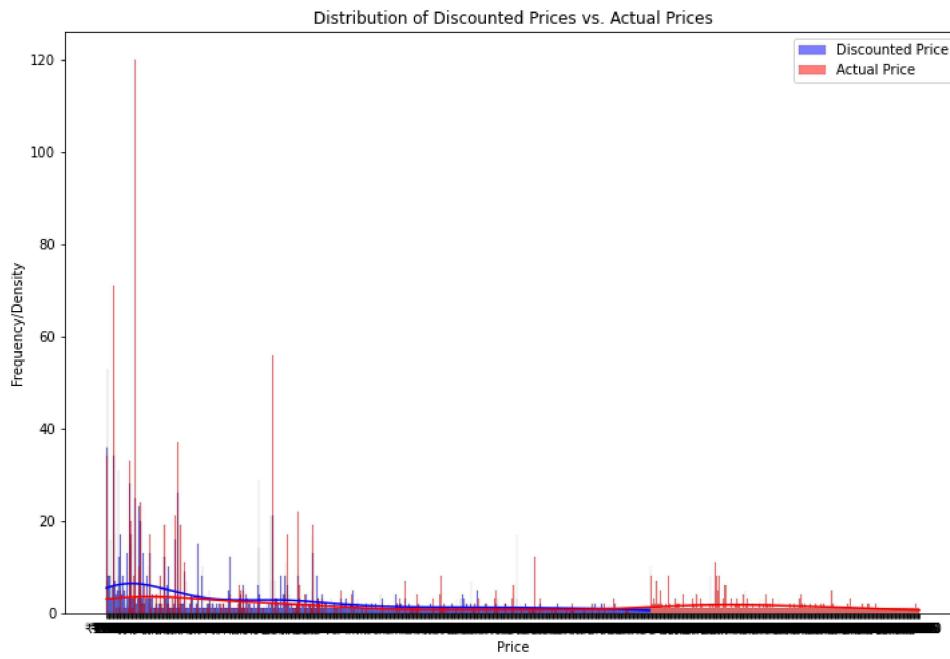
211 rows × 3 columns

```
In [89]: 1 df.dtypes
```

```
Out[89]: product_id          object
product_name        object
category           object
discounted_price   object
actual_price       object
discount_percentage    object
rating             float64
rating_count       object
about_product      object
user_id            object
user_name          object
review_id          object
review_title       object
review_content     object
img_link           object
product_link       object
dtype: object
```

3. What is the distribution of discounted prices vs. actual prices?

```
In [90]: 1 plt.figure(figsize=(12,8))
2 sns.histplot(df['discounted_price'], color='blue', kde=True, label='Discounted Price')
3 sns.histplot(df['actual_price'], color='red', kde=True, label='Actual Price')
4 plt.xlabel('Price')
5 plt.ylabel('Frequency/Density')
6 plt.title('Distribution of Discounted Prices vs. Actual Prices')
7 plt.legend()
8 plt.show()
```



```
In [98]: 1 # Convert 'discounted_price' and 'actual_price' back to object (string) data type
2 df['discounted_price'] = df['discounted_price'].astype(str)
3 df['actual_price'] = df['actual_price'].astype(str)
4
```

```
In [99]: 1 df.dtypes
```

```
Out[99]: product_id          object
product_name         object
category            object
discounted_price    object
actual_price        object
discount_percentage object
rating              float64
rating_count        object
about_product       object
user_id             object
user_name           object
review_id           object
review_title        object
review_content      object
img_link            object
product_link        object
dtype: object
```

4. How does the average discount percentage vary across categories?

In [100]: 1 df.head()

Out[100]:

	review_id	review_title	review_content	img_link	product_
	iHXWT0LRP0NMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...	Satisfied,Charging is really fast,Value for mo...	Looks durable Charging is fine tooNo complains...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Way Braided-WN3LG1-
	IQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...	A Good Braided Cable for Your Type C Device,Go...	I ordered this cable to connect my phone to An...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Ambr Unbreakable-Cf
	R3J3EQQ9T2I5ZJ,R3E7WBGK7ID0KV,RWU79XKQ6I1QF,R2...	Good speed for earlier versions,Good Product,W...	Not quite durable and sturdy,https://m.media-a...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Soul iPhone-Charging
	3EEUZKKK9J36I,R3HJVYCLYOY554,REDECAZ7AMPQC,R1...	Good product,Good one,Nice,Really nice product...	Good product,long wire,Charges good,Nice,I bou...	https://m.media-amazon.com/images/I/41V5FtEWPK...	https://www.amazon.in/Del 300-Resistant-Ta
	R1BP4L2HH9TFUP,R16PVJEXKV6QZS,R2UPDB81N66T4P,R...	As good as original,Decent,Good one for second...	Bought this instead of original apple, does th...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Portr Konnect-POR

In [102]: 1 # Remove '%' symbol from 'discount_percentage' column and convert to numeric
2 df['discount_percentage'] = df['discount_percentage'].str.rstrip('%').astype(float)

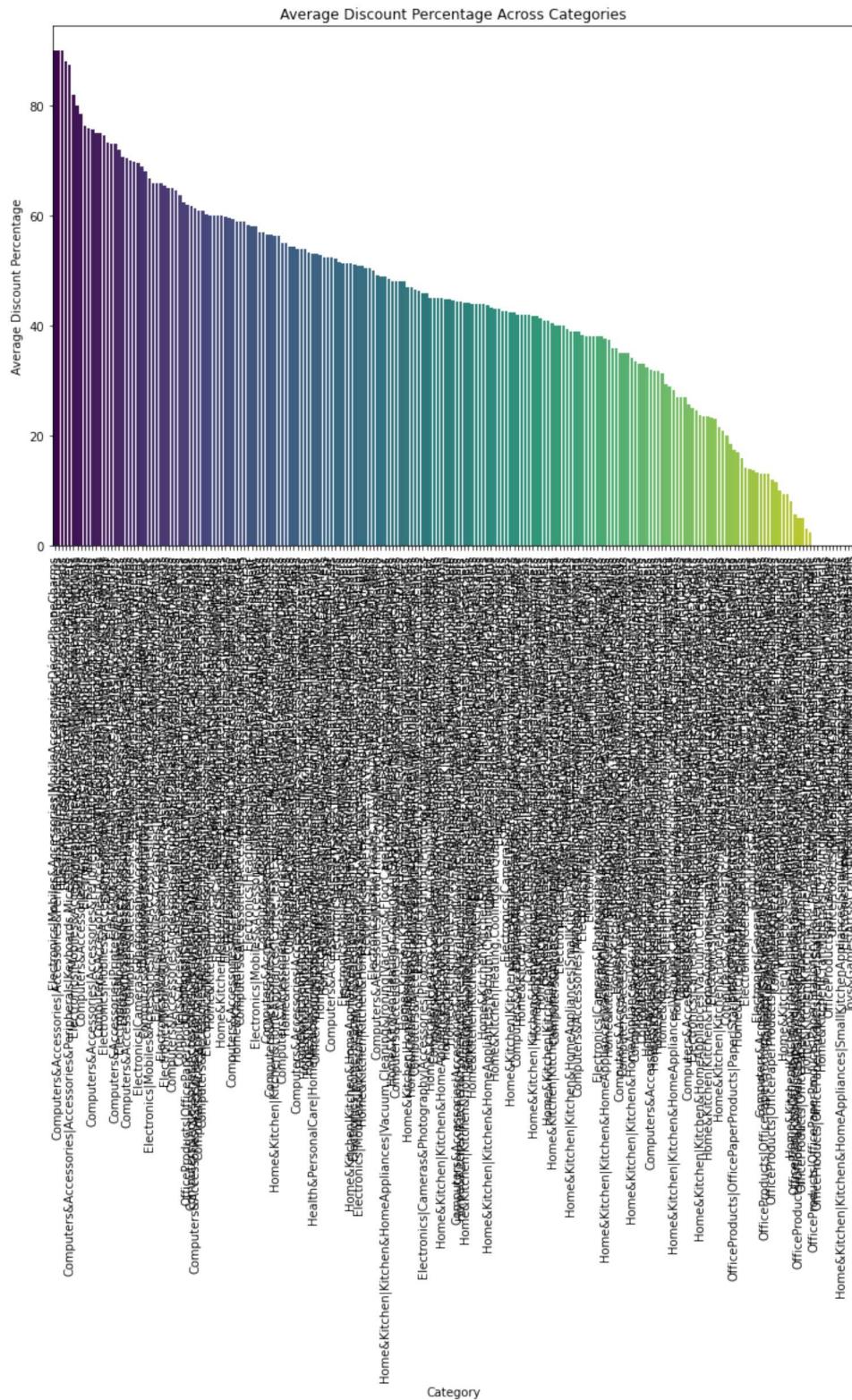
In [104]: 1 avg_discount_per_category = df.groupby('category')['discount_percentage'].mean().reset_index()
2 avg_discount_per_category = avg_discount_per_category.sort_values(by='discount_percentage', ascending=False)
3 avg_discount_per_category

Out[104]:

	category	discount_percentage
106	Electronics Mobiles&Accessories MobileAccessori...	90.0
6	Computers&Accessories Accessories&Peripherals ...	90.0
75	Electronics Headphones,Earbuds&Accessories Ear...	90.0
73	Electronics Headphones,Earbuds&Accessories Ada...	88.0
14	Computers&Accessories Accessories&Peripherals ...	87.5
...
196	OfficeProducts OfficeElectronics Calculators B...	0.0
176	Home&Kitchen Kitchen&HomeAppliances SmallKitch...	0.0
81	Electronics HomeAudio MediaStreamingDevices St...	0.0
62	Electronics Cameras&Photography Accessories Film	0.0
210	Toys&Games Arts&Crafts Drawing&PaintingSupplie...	0.0

211 rows × 2 columns

```
In [105]: 1 plt.figure(figsize=(12,8))
2 sns.barplot(data=avg_discount_per_category, x='category', y='discount_percentage', palette='viridis')
3 plt.xticks(rotation=90)
4 plt.xlabel('Category')
5 plt.ylabel('Average Discount Percentage')
6 plt.title('Average Discount Percentage Across Categories')
7 plt.show()
```



5. What are the most popular product names?

```
In [111]: 1 popular_product_name = df['product_name'].value_counts()
2
3 top_10_popular_products = popular_product_name.head(10)
4 top_10_popular_products
```

```
Out[111]: Fire-Boltt Ninja Call Pro Plus 1.83" Smart Watch with Bluetooth Calling, AI Voice Assistance, 100 Sports Modes IP67 Rating, 240 *280 Pixel High Resolution      5
Fire-Boltt Phoenix Smart Watch with Bluetooth Calling 1.3",120+ Sports Modes, 240*240 PX High Res with SpO2, Heart Rate Monitoring & IP67 Rating      4
Wayona Nylon Braided USB to Lightning Fast Charging and Data Sync Cable Compatible for iPhone 13, 12,11, X, 8, 7, 6, 5, iPad Air, Pro, Mini (3 FT Pack of 1, Grey)      3
MI Braided USB Type-C Cable for Charging Adapter (Red)
3
Amazonbasics Nylon Braided Usb-C To Lightning Cable, Fast Charging Mfi Certified Smartphone, Iphone Charger (6-Foot, Dark Grey)
3
Samsung Galaxy M13 5G (Aqua Green, 6GB, 128GB Storage) | 5000mAh Battery | Upto 12GB RAM with RAM Plus
3
Amazon Basics USB Type-C to USB-A 2.0 Male Fast Charging Cable for Laptop - 3 Feet (0.9 Meters), Black
3
boAt A400 USB Type-C to USB-A 2.0 Male Data Cable, 2 Meter (Black)
3
Duracell USB C To Lightning Apple Certified (Mfi) Braided Sync & Charge Cable For Iphone, Ipad And Ipod. Fast Charging Lightning Cable, 3.9 Feet (1.2M) - Black      3
AmazonBasics New Release Nylon USB-A to Lightning Cable Cord, Fast Charging MFi Certified Charger for Apple iPhone, iPad (6-Ft, Rose Gold)      3
Name: product_name, dtype: int64
```

6. What are the most popular product keywords?

```
In [112]: 1 df.head()
```

Out[112]:

	review_id	review_title	review_content	img_link	product_link
	T0LRP0NMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1K...	Satisfied,Charging is really fast,Value for mo...	Looks durable Charging is fine tooNo complains...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Wayona-Braided-WN3LG1-Sy...
	307R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RY...	A Good Braided Cable for Your Type C Device,Go...	I ordered this cable to connect my phone to An...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Ambrane-Unbreakable-Char...
	EQQ9TZI5ZJ,R3E7WBGK7ID0KV,RWU79XKQ6I1QF,R2...	Good speed for earlier versions,Good Product,W...	Not quite durable and sturdy,https://m.media-a...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Sounce-iPhone-Charging-C...
	JZKKK9J36I,R3HJVYCLYOY554,REDECAZ7AMPQC,R1...	Good product,Good one,Nice,Really nice product...	Good product,Long wire,Charges good,Nice,I bou...	https://m.media-amazon.com/images/I/41V5fEWPK...	https://www.amazon.in/Deuce-300-Resistant-Tang...
	4L2HH9TFUP,R16PVJEXKV6QZS,R2UPDB81N66T4P,R...	As good as original,Decent,Good one for second...	Bought this instead of original apple, does th...	https://m.media-amazon.com/images/W/WEBP_40237...	https://www.amazon.in/Portronics-Konnect-POR-1...

```
In [113]: 1 from collections import Counter
2 import re
```

```
In [119]: 1 product_keywords_counter = Counter()
2 for product_name in df['product_name']:
3     # Split product name into words using regular expression
4     words = re.findall(r'\b\w+\b', product_name.lower())
5     # Update counter with the words
6     product_keywords_counter.update(words)
7
8 # Display the top 10 most popular product keywords
9 top_10_popular_keywords = product_keywords_counter.most_common(25)
10 top_10_popular_keywords
```

```
Out[119]: [('with', 752),
('for', 674),
('black', 530),
('usb', 417),
('cable', 414),
('1', 413),
('and', 331),
('c', 289),
('3', 275),
('2', 267),
('charging', 247),
('type', 243),
('to', 219),
('fast', 212),
('in', 178),
('smart', 175),
('white', 170),
('tv', 152),
('compatible', 146),
('5', 137),
('data', 136),
('hd', 135),
('4', 133),
('led', 129),
('8', 125)]
```

7. What are the most popular product reviews?

```
In [122]: 1 popular_reviews = df['review_content'].value_counts()
2 print(popular_reviews)

repacked and requested for replacement.I checked again, and there was some green colour paste/fungus inside the micro USB connector. I cleaned with an alcoholic and starts working again.Checked the ampere of charging speed got around 1400ma-1500ma - not bad, came with braided 1.5m long cable, pretty impressive for the price.Can't blame the manufacturer.But quality issues by the distributor, they might have stored in very humid place.

7
Good product,long wire,Charges good,Nice,I bought this cable for Rs.339 worthy product for this price, i tested it in various charger adapters 33w and 18w it supports fast charging as well.,Good,Ok,I had got this at good price on sale on Amazon and product is useful with warranty but for warranty you need to go very far not practical for such a cost and mine micro to type c connector stopped working after few days only.,I like this product

7
128GB SD Card is showing 134GBDon't buy this product,It was a bit costly but product is of great quality. Could have been made little bit cheaper cost wise.,All good,Fast and smooth,Ab mujhe computer ki jarurat nahi padti,Speed is fast but not as advertised I get around read speed 70MB/s and write speed 30MB/s,Good one,Product is ok.

6

..
I generally don't write reviews on Amazon. But this product is worth every penny.Amazing Product1. It is sturdy2. Good heavy-quality metal is used yet still lightweight and portable3. Can be easily carried in 1 hand4. On Full Open, total Height reaches 5ft 9in (so can even cover if there's a stage)After buying this, I even surveyed the local market, I can now confidently say I have invested in the right product There are products available for 1,000 Rs but having another just 500Rs for such a
```

```
In [125]: 1 top_10_pop_reviews = popular_reviews.head(10).index
2
3 # Filter the DataFrame for rows containing the most popular review content
4 popular_reviews_df = df[df['review_content'].isin(top_10_pop_reviews)]
5
6
```

```
In [124]: 1 # Get the most popular review content
2 most_popular_reviews = popular_reviews.head(10).index
3
4 # Filter the DataFrame for rows containing the most popular review content
5 popular_reviews_df = df[df['review_content'].isin(most_popular_reviews)]
6
7 # Print the product name and review content for each popular review
8 for index, row in popular_reviews_df.iterrows():
9     print("Product Name:", row['product_name'])
10    print("Review Content:", row['review_content'])
11    print("*"*50)
12
```

Product Name: Ambrane Unbreakable 60W / 3A Fast Charging 1.5m Braided Type C Cable for Smartphones, Tablets, Laptops & other Type C devices, PD Technology, 480Mbps Data Sync, Quick Charge 3.0 (RCT15A, Black)

Review Content: I ordered this cable to connect my phone to Android Auto of car. The cable is really strong and the connection ports are really well made. I already has a Micro USB cable from Ambrane and it's still in good shape. I connected my phone to the car using the cable and it got connected well and no issues. I also connected it to the charging port and yes it has Fast Charging support..It quality is good at this price and the main thing is that i didn't ever thought that this cable will be so long it's good one and charging power is too good and also supports fast charging,Value for money, with extra length
Good, working fine,Product quality is good,Good,very good,Bought for my daughter's old phone.Brand new cable it was not charging, I already repacked and requested for replacement.I checked again, and there was some green colour paste/fungus inside the micro USB connector. I cleaned with an alcoholic and starts working again.Checked the ampere of charging speed got around 1400ma-1500ma - not bad, came with braided 1.5m long cable, pretty impressive for the price.Can't blame the manufacturer.But quality issues by the distributor, they might have stored in very humid place.

=====

Product Name: boAt Deuce USB 300 2 in 1 Type-C & Micro USB Stress Resistant, Tangle-Free, Sturdy Cable with 3A Fast Charging & 480mbps Data Transmission, 10000+ Bends Lifespan and Extended 1.5m Length(Martian Red)

Review Content: Good product,long wire,Charges good,Nice,I bought this cable for Rs.339 worthy product for this price, i tested it in various charger adapters 33w and 18w it supports fast charging as well.,Good,Ok,I had got this at good price on sale on Amazon and product is useful with warranty but for warranty you need to go very far not practical for such a cost and mine micro to type c connector stopped working after few days only.,I like this product

What is the correlation between discounted_price and rating?

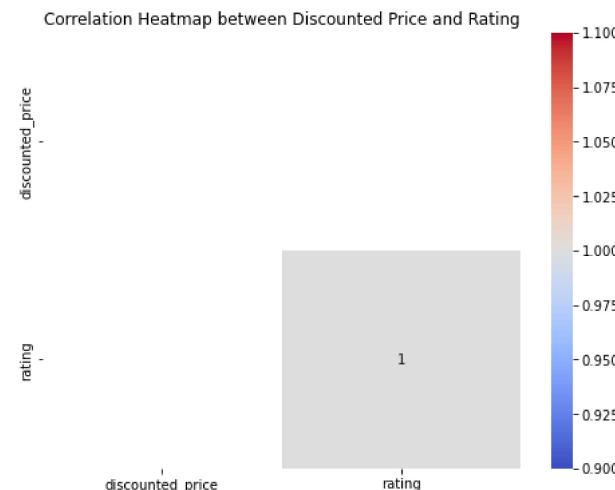
```
In [127]: 1 df['discounted_price'] = pd.to_numeric(df['discounted_price'], errors='coerce')
```

```
In [145]: 1 correlation = df[['discounted_price', 'rating']].corr()
2 correlation
```

Out[145]:

	discounted_price	rating
discounted_price	NaN	NaN
rating	NaN	1.0

```
In [146]: 1 plt.figure(figsize=(8, 6))
2 sns.heatmap(correlation, annot=True, cmap='coolwarm')
3 plt.title('Correlation Heatmap between Discounted Price and Rating')
4 plt.show()
```



```
In [129]: 1 df.dtypes
```

```
Out[129]: product_id          object
product_name        object
category           object
discounted_price   float64
actual_price       object
discount_percentage float64
rating             float64
rating_count       object
about_product      object
user_id            object
user_name          object
review_id          object
review_title       object
review_content     object
img_link           object
product_link       object
dtype: object
```

9. What are the Top 5 categories based on the highest ratings?

```
In [147]: 1 avg_rating_per_category
```

```
Out[147]:
```

	category	rating
0	Car&Motorbike CarAccessories InteriorAccessori...	3.800000
1	Computers&Accessories Accessories&Peripherals ...	4.150000
2	Computers&Accessories Accessories&Peripherals ...	3.500000
3	Computers&Accessories Accessories&Peripherals ...	3.600000
4	Computers&Accessories Accessories&Peripherals ...	4.050000
...
206	OfficeProducts OfficePaperProducts Paper Stati...	4.250000
207	OfficeProducts OfficePaperProducts Paper Stati...	4.150000
208	OfficeProducts OfficePaperProducts Paper Stati...	4.300000
209	OfficeProducts OfficePaperProducts Paper Stati...	4.133333
210	Toys&Games Arts&Crafts Drawing&PaintingSupplie...	4.300000

211 rows × 2 columns

```
In [150]: 1 top_5_category = avg_rating_per_category.sort_values(by='rating', ascending=False)
2 # print(top_5_category[['category', 'rating']])
3 top_5_category.head(5)
```

```
Out[150]:
```

	category	rating
57	Computers&Accessories Tablets	4.6
48	Computers&Accessories NetworkingDevices Networ...	4.5
62	Electronics Cameras&Photography Accessories Film	4.5
81	Electronics HomeAudio MediaStreamingDevices St...	4.5
196	OfficeProducts OfficeElectronics Calculators B...	4.5

10. Identify any potential areas for improvement or optimization based on the data analysis.

```
In [151]: 1 ...
2 1] Data Quality Assurance:
3 There are instances of missing values and incorrect data types in the dataset, such as non-numeric values in numeric columns
4 Ensuring data quality through thorough cleaning, validation, and preprocessing can improve the accuracy of analysis results.
5
6 2] Feature Engineering: Creating new features from existing ones, such as deriving additional insights from product names,
7 review content, or other textual data, can enrich the analysis and provide deeper insights into customer preferences and
8 behavior.
9
10 3] Sentiment Analysis: Analyzing the sentiment of product reviews can provide valuable insights into customer satisfaction
11 levels and identify areas for product improvement or enhancement.
12
13 4] Customer Segmentation: Segmenting customers based on their behavior, preferences, or demographics can help tailor marketing
14 strategies, product offerings, and customer service to specific customer segments, thereby improving customer satisfaction
15 and retention.
16
17 5] Price Optimization: Analyzing the relationship between pricing, discounts, and product ratings can help optimize pricing
18 strategies to maximize revenue and profitability while maintaining customer satisfaction.
19
20 6] Category-specific Analysis: Conducting more in-depth analysis within each product category can reveal category-specific
21 trends, customer preferences, and areas for improvement tailored to each category.
22
23 7] Predictive Modeling: Developing predictive models, such as forecasting sales or predicting customer churn, can help
24 anticipate future trends and make data-driven business decisions.
25
26 8] Feedback Loop Implementation: Establishing a feedback loop where insights from data analysis inform decision-making
27 processes and drive continuous improvement across all aspects of the business can lead to ongoing optimization and innovation
28 ...
```

```
Out[151]: '\n1] Data Quality Assurance: \nThere are instances of missing values and incorrect data types in the dataset, such as non-numeric values in numeric columns\nEnsuring data quality through thorough cleaning, validation, and preprocessing can improve the accuracy of analysis results.\n\n2] Feature Engineering: Creating new features from existing ones, such as deriving additional insights from product names, review content, or other textual data, can enrich the analysis and provide deeper insights into customer preferences and behavior.\n\n3] Sentiment Analysis: Analyzing the sentiment of product reviews can provide valuable insights into customer satisfaction levels and identify areas for product improvement or enhancement.\n\n4] Customer Segmentation: Segmenting customers based on their behavior, preferences, or demographics can help tailor marketing strategies, product offerings, and customer service to specific customer segments, thereby improving customer satisfaction and retention.\n\n5] Price Optimization: Analyzing the relationship between pricing, discounts, and product ratings can help optimize pricing strategies to maximize revenue and profitability while maintaining customer satisfaction.\n\n6] Category-specific Analysis: Conducting more in-depth analysis within each product category can reveal category-specific trends, customer preferences, and areas for improvement tailored to each category.\n\n7] Predictive Modeling: Developing predictive models, such as forecasting sales or predicting customer churn, can help anticipate future trends and make data-driven business decisions.\n\n8] Feedback Loop Implementation: Establishing a feedback loop where insights from data analysis inform decision-making processes and drive continuous improvement across all aspects of the business can lead to ongoing optimization and innovation.\n'
```

```
In [ ]: 1
```