

Amazon Sales 2023 Project

1- About Dataset :

- This dataset is having the data of 1K+ Amazon Product's Ratings and Reviews as per their details listed on the official website of Amazon

2- Inspiration :

- Amazon is an American Tech Multi-National Company whose business interests include E-commerce, where they buy and store the inventory, and take care of everything from shipping and pricing to customer service and returns. I've created this dataset so that people can play with this dataset and do a lot of things as mentioned below Dataset Walkthrough Understanding Dataset Hierarchy Data Preprocessing Exploratory Data Analysis Data Visualization This is a list of some of that things that you can do on this dataset. It's not definitely limited to the one that is mentioned there but a lot more other things can also be done.

3- Columns :

- **product_id** : Product ID
- **product_name** : Name of the Product
- **category** : Category of the Product
- **discounted_price** : Discounted Price of the Product
- **actual_price** : Actual Price of the Product
- **discount_percentage** : Percentage of Discount for the Product
- **rating** : Rating of the Product
- **rating_count** : Number of people who voted for the Amazon rating
- **about_product** : Description about the Product
- **user_id** : ID of the user who wrote review for the Product
- **user_name** : Name of the user who wrote review for the Product
- **review_id** : ID of the user review
- **review_title** : Short review
- **review_content** : Long review
- **img_link** : Image Link of the Product
- **product_link** : Official Website Link of the Product

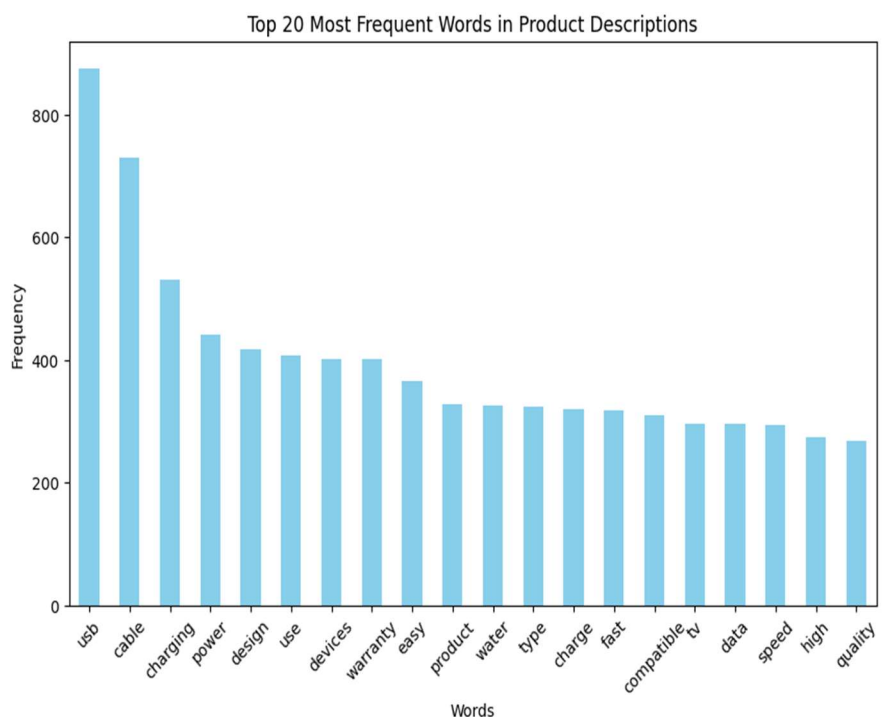
4- Business Questions :

- a) the Top 20 most common word in Product description ?
- b) what aspects of the products customers are most interested in or concerned about ?
- c) can you give overall sense of customer sentiment towards your product ?
- d) Which products have high rating ?
- e) what are the best common categories are sold ?
- f) how many products have rating more than 4 ?
- g) what are the top 5 second category ? and what are the unsold or old products ?
- h) which products have high percentage discount where percentage discount > 60% ?
- i) are there Last category ? if exists tell me the top 10 categories that have high prices ?
- j) is there any relationship between actual price and rating or rating count ?
- k) who are users that are loyalty for our company ?
- l) are the user buy products that their discount price more than 50% ?

5- Answer of Questions :

a) It show here from the visualization the customers attract to products that have these words :

- The analysis reveals a prominent presence of certain keywords among customer purchases, including "USB," "cable," "charging," "power," "use," "devices," "warranty," "easy," "product," "water," "type," "fast," "compatible TV," "data," "high," and "quality." These terms signify features and attributes highly sought after by consumers, indicating a strong appeal towards products



embodying these characteristics. Consequently, it is advisable to prioritize attention towards items incorporating these keywords either within their titles or descriptions, as they are likely to garner increased interest and demand among our target audience.

- b) Based on the analysis of customer feedback, the following aspects are the most significant in terms of customer interest and concern:

```
Topic 1: good, product, use, nice, best, quality, easy, price, money, value
Topic 2: good, product, nice, quality, money, cable, value, price, best, worth
Topic 3: product, good, nice, money, value, working, best, quality, excellent, ok
Topic 4: good, working, remote, price, product, tv, sound, buy, work, kodak
Topic 5: good, product, money, value, quality, tv, nice, price, worth, best
```

- **Topic 1: Overall Product Satisfaction**

Keywords: good, product, use, nice, best, quality, easy, price, money, value

Customers generally express satisfaction with the product, emphasizing its quality, ease of use, and value for money.

- **Topic 2: Product Quality and Value**

Keywords: good, product, nice, quality, money, cable, value, price, best, worth

Quality and value are highlighted, with specific mention of cables, suggesting that customers prioritize durable and well-made products.

- **Topic 3: Product Performance and Value**

Keywords: product, good, nice, money, value, working, best, quality, excellent, ok

Customers focus on product functionality and value for money, expecting products to perform well and provide good value.

- **Topic 4: Specific Product Features (e.g., Remote, Sound)**

Keywords: good, working, remote, price, product, tv, sound, buy, work, kodak

Customers highlight specific features like remote controls and sound quality, indicating that these features significantly impact their purchasing decisions.

- **Topic 5: Overall Product Value**

Keywords: good, product, money, value, quality, tv , nice, price, worth, best

Emphasis is placed on overall product value, encompassing various aspects such as quality, price, and perceived worthiness.

- **These topics provide insights into what customers value most in products, including quality, value for money, specific features, and overall satisfaction.**

- c) From results it show about 1031 of products the customers are satisfy about it and have a good positive

```
sentiment
Positive    1031
Negative    105
Neutral     84
Name: count, dtype: int64
```

about it , but about 105 of products that customers have a negative side of these products as it ,

So after I analyzing data I found these products that have a negative side from customers are bad quality , very less power and more I will show you some of these reviews :

Charging power is very less. It takes hours to charge even a single phone.,Ok type product. Not very sturdy,Although it is written fast charging. However it charges quite slow compared to original chargers, even when one mobile is charged at one time.,When using a regular cable, it takes about 2 hrs for a full charge. When using this cable to the same socket and charger, it showing 10 hrs.,iPhone charger last for two months only,Charging not fast,Slow charging4500amh battery18 watt Charger4 khant full charge,Slow charging speed

- The only issue is that the built quality is very poor. It broke on 2nd use. Rest the adhesion is good.,The product is not sturdy, the handle broke after a weekThe adhesive is working, but it is too diffcult to peel the tape,Product is okay but the stick is very delicate so it broke, now it very difficult to use the product.,The handle of the product broke when I used it second time so it's very poor quality would not recommend.,Its value for money,Easy to use,Product of good use. Wish the Handle of the roller was easy to use,We have a beagle dog so we have extensive use of lint rollers in the household. Have been using the 3M, but its very expensive if you consume 4to 5 rollers every month. So thought of trying this one as i though how bad can a lint roller be. The stickiness of the glue paper was good enough to remove the dog hair, but the quality of the handle is really bad. It did not even last 2 days. So ended up using a wooden stick or long pencils for the job. If the makers had spent some more on making a good handle even if the cost would have been Rs 25 higher, then it could have been a great product especially for those households which have pets as they need it more frequently. So i might go back to buying the 3M product of look for something in between.

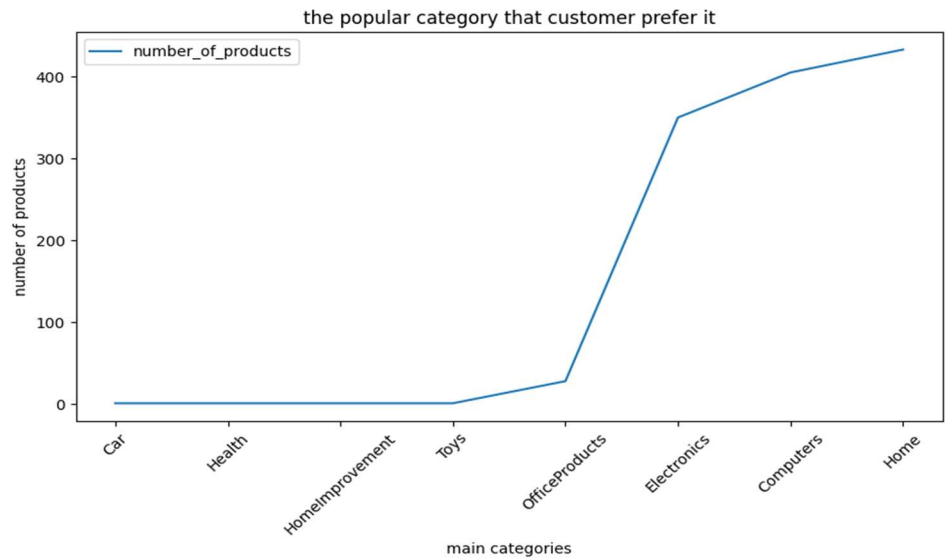
These are samples from the products that have a negative side from customers so we need to improve these products

d) it show these products have high number of rating **these mean** it have a positive side from customer and have a good reviews so we need to focus on them to increase revenue

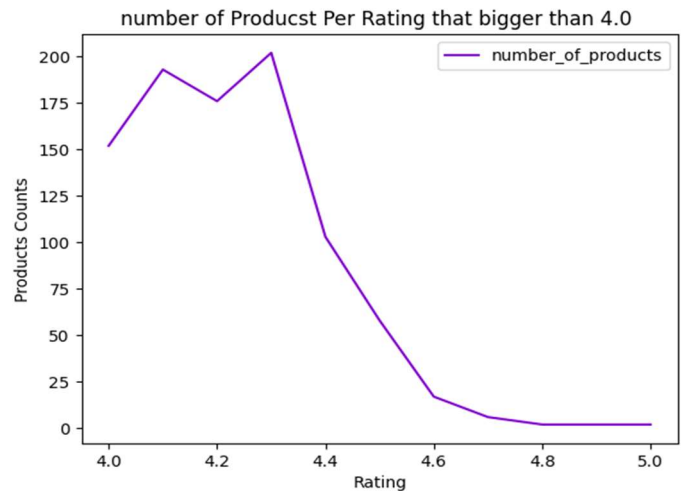
	product_name	rating
295	Duracell USB Lightning Apple Certified (Mfi) B...	13.5
589	MI Braided USB Type-C Cable for Charging Adapt...	13.2
103	Amazonbasics Nylon Braided Usb-C To Lightning ...	13.2
294	Duracell USB C To Lightning Apple Certified (M...	13.2
96	AmazonBasics New Release Nylon USB-A to Lightn...	12.9
78	Amazon Basics USB Type-C to USB-A 2.0 Male Fas...	12.9
591	MI Usb Type-C Cable Smartphone (Black)	12.9
731	Portronics Konnect L POR-1081 Fast Charging 3A...	12.9
1081	boAt A400 USB Type-C to USB-A 2.0 Male Data Ca...	12.9
1001	Wayona Nylon Braided USB to Lightning Fast Cha...	12.6

e) Here line chart show you the most popular categories that customer prefer it are Home , Computers and Electronics , also it show the lowest products that customer don't look at it are Cars , Healthy , HomeImprovement ,Toys , OfficeProducts .

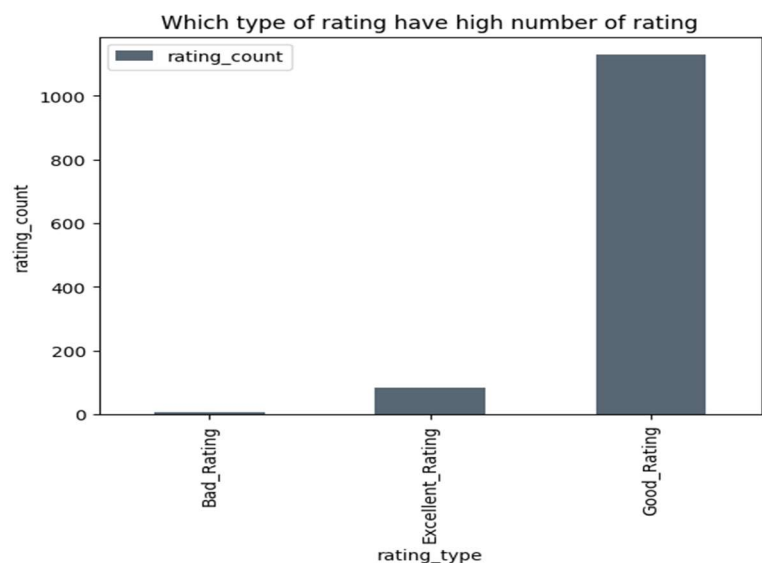
I find a reason is the most of these products have low discount price so the prices of these products are so high so customers give a low rating , also there are a bad reviews about some products cause of quality of products or bad power & cables



f) From the visualization you can see above 150 products are sold and customers have a high rating this show a lot of customers have a positive side and good reviews and good rating and this prove the loyalty of customers

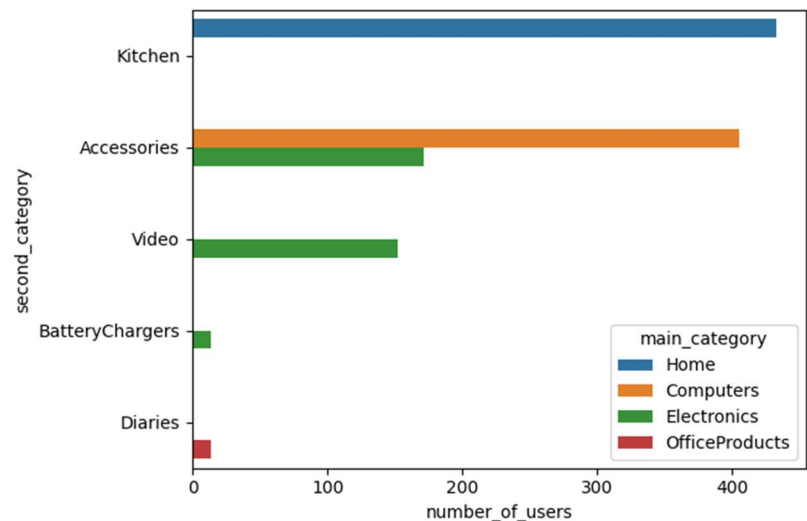


- from the another visualization it show the more than 1000 products have good rating mean have rating between (2.5 – 4)



g) - Here it show the top 5 second categories are kitchen , accessories , video , BatteryChargers , Diaries .

- Also it show kitchen are the top and popular in Home (main category) , but accessories it show it is popular in categories like : Computers (main category) and Electronics (main category) , Also it show Video are best products sold in electronics thanbatteryCharagers



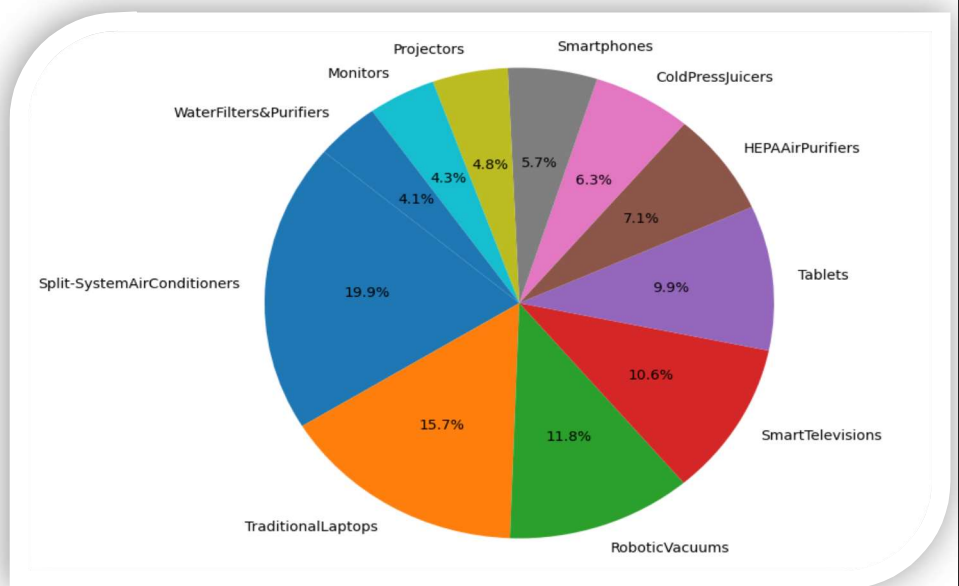
Also the Unsold products are :

- **IONIX Tap filter Multilayer | Activated Carbon** : it show people have negative side this mean this product have bad quality or not worked as well and i check it price in a other websites so here the price are so high , the discount here isso high also but not all time
- **Green Tales Heat Seal Mini Food Sealer-Impulse** : People have negative side also from this product , also his price are high compared to other websites one of the aspects that make rating of this product are low : his discount are low than 50% , the price here are so high , maybe product is bad quality
- **Personal Size Blender, Portable Blender, Batte..& MR. BRAND Portable USB Juicer Electric USB Jui** : here it show the people have postive side from these products also the prices are so high but the company give them high discount, so i think the reason that make this product are low rating : it maybethe quality not good , product not work , people not interested in this kind of products
- **Khaitan ORFin Fan heater for Home and kitchen** : have negative side from customers , it maybe cause of high price , discount lower than 50% , product not worked or quality bad
- **Eureka Forbes car Vac 100 Watts Powerful Sucti**: people are neutral his price is so good , after review his review content this product are not worked well

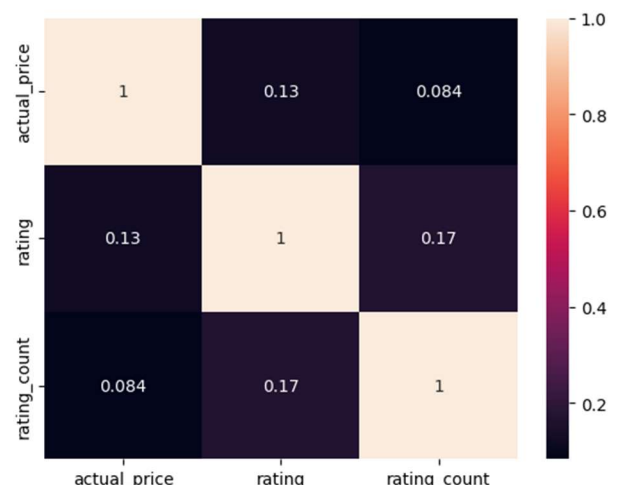
h) **From the visualization here it show these products have a lot of times good discount prices that attract customers as you see from this chart ,also customers have a give a good review on these products and it show these products are come from Electronics & computers (main categories)**

product_name	sentiment	rating_type	main_category	discount_percentage
7SEVEN® Compatible Tata Sky Remote Control Rep...	Positive	Good_Rating	Electronics	61.0
AGARO Blaze USB 3.0 to USB Type C OTG Adapter	Positive	Good_Rating	Electronics	72.0
Agaro Blaze USBA to micro +Type C 2in1 Braided...	Positive	Good_Rating	Computers	73.0
Airtel DigitalTV DTH Remote SD/HD/HD Recording...	Positive	Good_Rating	Electronics	61.0
Airtel DigitalTV DTH Television, Setup Box Rem...	Positive	Good_Rating	Electronics	78.0
Airtel DigitalTV HD Setup Box Remote	Positive	Good_Rating	Electronics	61.0
Amazon Basics Magic Slate 8.5-inch LCD Writing...	Positive	Good_Rating	Computers	71.0
Amazon Basics Multipurpose Foldable Laptop Tab...	Positive	Good_Rating	Computers	85.0
Amazon Basics USB A to Lightning PVC Molded Ny...	Positive	Good_Rating	Computers	61.0
Amazon Basics USB Type-C to USB-A 2.0 Male Fas...	Positive	Good_Rating	Computers	69.0

i) Yes these is another category are detected it show from the pie chart blow there are a lot of new categories that are popular for customers like : Monitors , Projectors , SmartPhones , ColdPressJuicers , HepaAirPurifiers , Tablets , Smart TV , Robotic Vacuums , Traditional labtops , Split-SystemAirConditiions , waterFilters&Purifiers , also from the visualization it show the most popular categories are : Split-SystemAirConditiions , TraditionalLabtop , RobotiveVacuums, smartTV , these categories show customers have a good attractive it & it will increase revenue if we take care about them .



j) **From the visualization it show there is no any relationship** between them but later I found a strong relationship between actual price & percentage price this are good it show the high price of product the customer will get a high percentage but we know correlation not imply causation so it just a probability based on the correlation.



k)

l) After some analysis I found these customers are the most loyalty for company based on high rating & high number of product sold & good review & they have a positive side & they love high discount percentage these customers are exists in the image and I think it will be more but I just get a 20 top of customers who are loyalty

user_name	rating
\$@ N TO\$ - ,Sethu madhav,Akash Thakur,Burger P...	40.0
Manav,Adarsh gupta,Sundeep,S.Sayed Ahmed,jasp...	33.6
Satheesh Kadiam,Pritom Chakraborty,Vishwa,Simr...	32.8
siddharth patnaik,Dr Sunilkumar H,Krishna,K. S...	24.6
Wraith,Krishna Engineering,Kindle Customer,Dee...	22.0
Ayush,ROHIT A.,Kedar,Haran,Santosh Ghante,KRIS...	21.5
Manoj maddheshiya,Manoj Kumar Sahoo,Saumil s,...	21.0
Actual user,Shanti lal Gurjar,Gowthami,Amrut K...	20.0
Placeholder,श्रीPKजी,Gauri salkar,aQeel,Ashish...	19.5
Binu,Sabjan,Ramkapil>User,Amazon Customer,gane...	17.2
sameer Dubey,virender,Amazon Customer,Jay,Sidd...	17.2
Pavan A H,Jayesh bagad,Shridhar,rajendran,karu...	17.2
GHOST,Amazon Customer,A Vase,Mani Jha,Anu B,Ga...	17.2
Vijayan C V,Sanjay P.,keshav,Dharemesh K,Amaz...	16.0
Jayesh,Rajesh k,Soopy,amazon customer,Aman,Sh...	15.6
Prasannavijayaraghavan G.,Preet Sikka,Theeban ...	13.5
Anonymous,Sugam Agrawal,Diganta Gogoi,Amazon C...	13.2
Amazon Customer,Govind Gujarathi,Vihari,shalem...	13.2
Birendra ku Dash,Aditya Gupta,Abdulla A N,Deep...	13.2
Tanya,Anu,Akshay,Vishal Sagara Shetty,Swatilek...	12.9

i) I test this hyposthis and I found :

- **Null Hypothesis (H0):** Customers are not attracted to products with a discounted percentage of 50% or higher
- **Alternative Hypothesis (H1):** Customers are attracted to products with a discounted percentage of 50% or higher
- **Significance Level (alpha):** 0.05
- T-statistic: 26.44703464745795
- P-value: 3.4583476712796816e-122
- **Reject the null hypothesis. There is sufficient evidence to suggest that Customers are not attracted to products with a discounted percentage of 50% or higher**

Thank you