



Analysis of Myntra Apparel

By: Divya Sharma

Super Fast
Deliveries



www.myntra.com

About the Company



- Founded in 2007, headquartered in Bengaluru, India
- One of India's leading online fashion & lifestyle marketplaces
- Offers apparel, footwear, accessories, beauty & home products
- Acquired by Flipkart (a Walmart subsidiary) in 2014
- Known for exclusive collections & mega sales like End of Season Sale
- Uses big data to study customer behavior & preferences
- Provides personalized product recommendations
- Predicts fashion trends using advanced analytics & AI
- Optimizes pricing, inventory & supply chain management
- Enhances customer experience through data-driven insights



Learnings / What I Have Learned



01

Data Collection:
Learned how to organize datasets properly and select visible columns using "Alt + ;" for easier handling.

02

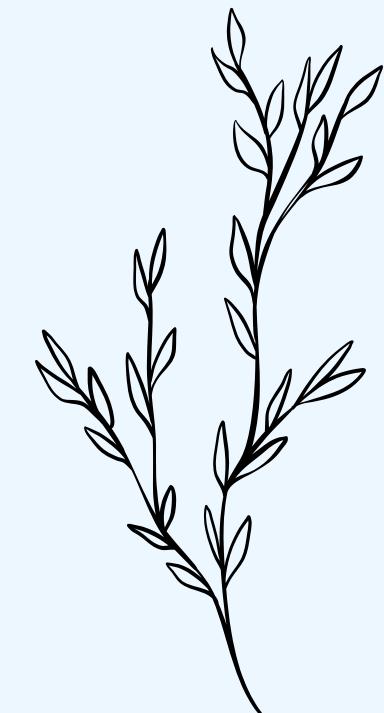
Data Cleaning:
Learned to use Ctrl + F / H for Find and Replace and ensure clean, error-free data.

03

Data Preparation:
Learned to use the Fill option to copy data left/right and Ctrl + Enter to paste values into multiple cells at once.

04

Data Analysis:
Learned how to use formulas such as IF, AVERAGEIF, COUNTIF, VLOOKUP, HLOOKUP, and XLOOKUP to analyze patterns.



Challenges Faced



Circular dependency

Learned to check formula references carefully to fix the issue.



Copying formulas

Learned to copy data and paste as values to continue working safely.



Deleting columns

Learned to remove filters before deleting columns to avoid data loss.



Cause:

A formula referred to its own cell.



Cause:

Needed to work with static values instead of formulas.

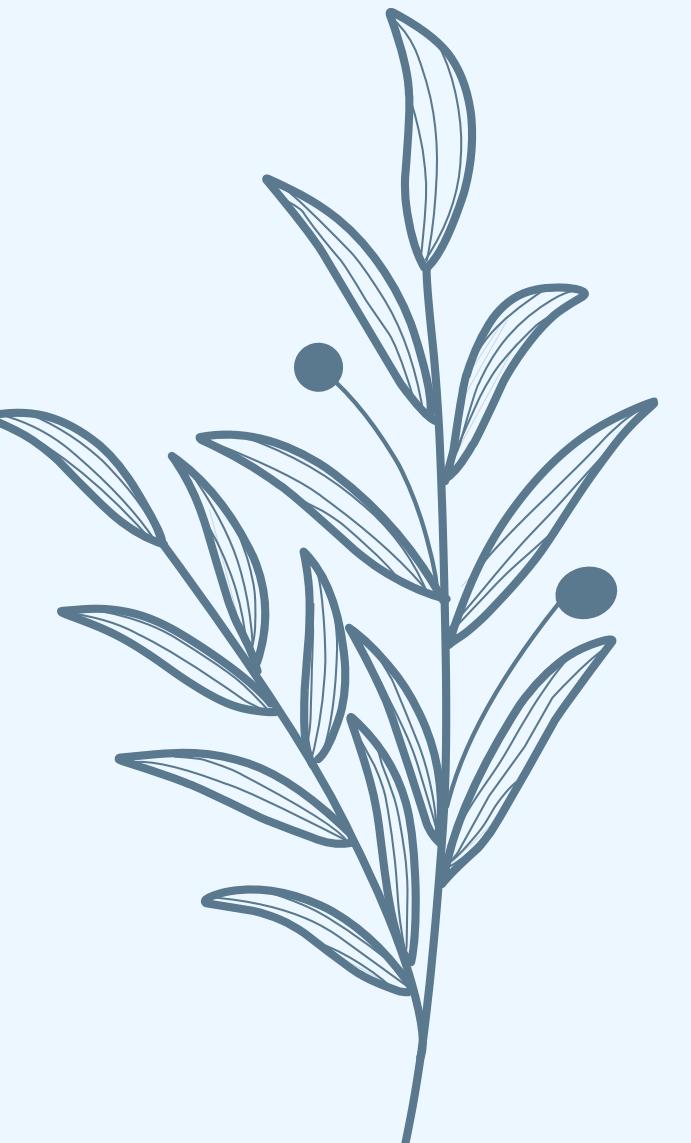


Evaluation & Scaling

Review expansion results and plan continued growth.



A. Data Cleaning and Preparation



I. Checking for Duplicate Value

Home Insert Draw Page Layout Formulas **Data** Review View Automate Help Table Design

Queries & Connections Properties Workbook Links

Transform Data Refresh All Workbook Links

Stocks Currencies

A Z Z A Z Sort Filter Clear Reapply Advanced

Text to Columns Data Tools What-If Analysis Forecast Sheet Outli

Forecast

SHIRTS MEN

Remove Duplicates ? X

To delete duplicate values, select one or more columns that contain duplicates.

Select All Unselect All My data has headers

Columns

URL
 Product_id
 BrandName
 Category
 Individual_category

OK Cancel

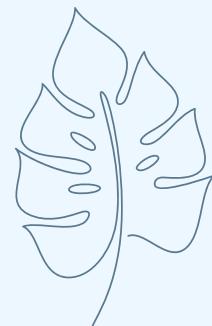
No duplicate values found.

Microsoft Excel

629 893 599

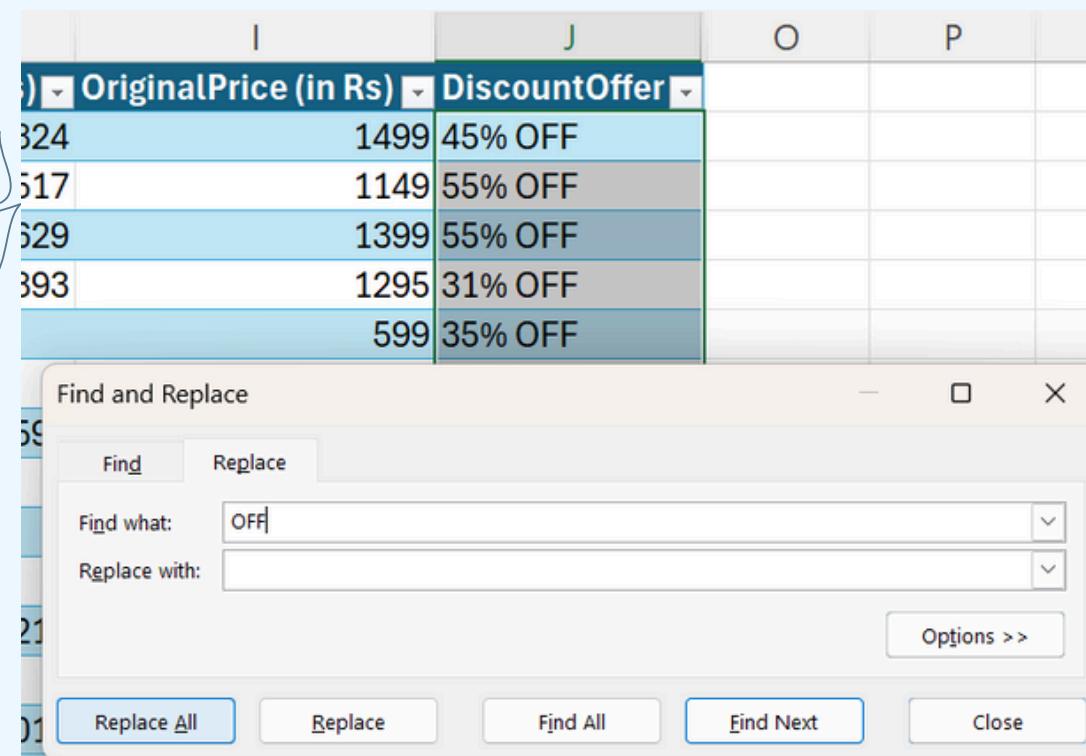
No duplicate values because each product has a unique ID.

2. Standardize the DiscountOffer values



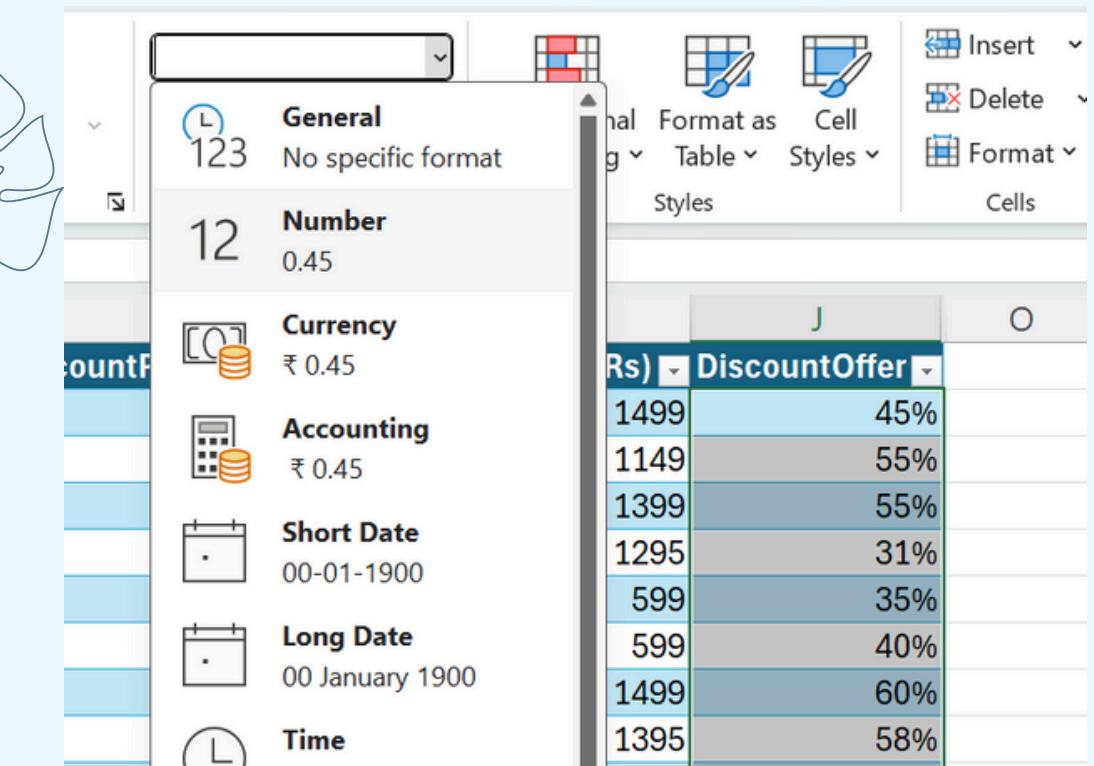
H	I	J	N	P	Q	R	S	T	U	V
DiscountPrice (in Rs)	OriginalPrice (in Rs)	DiscountOffer	DISCOUNT							
824	1499	45% OFF	=IF([@DiscountPrice (in Rs)]="", "",[@OriginalPrice (in Rs)]-[@DiscountPrice (in Rs)])							
517	1149	55% OFF								
629	1399	55% OFF								

Using formula to get values for blank DiscountOffer cells.



 Standardizing DiscountOffer is needed because values are mixed (% , flat amounts, etc.), making them comparable on a common scale.

Replacing words to get % or Flat values.

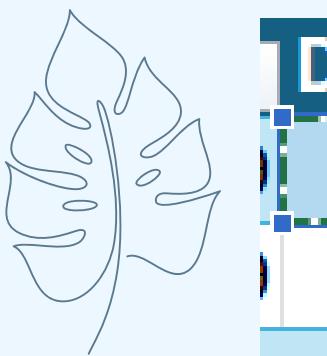


Converting to numbers to use values more/less than one.

less than 1 values are % and we want greater than 1 values because we want to convert them in %

	I	J	N
	OriginalPrice (in Rs)	DiscountOffer	Percentage disc
	999	500.00	=J184/I184
	899	400.00	
	700	300.00	

To convert all flat discounts into percentages.



DiscountOffer	Percentage disc
0.45	=J2
0.55	

Apply Filter less than 1 (percent values) & applying formula so all % values also get with flat values which was converted before



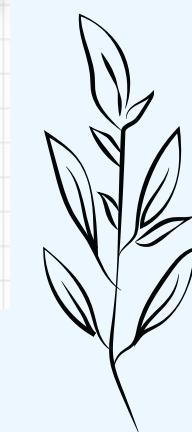
OriginalPrice (in Rs)	Percentage disc
1499	45.00%
1149	55.00%
1399	55.00%
1295	31.00%
599	35.00%
599	40.00%
1499	60.00%
1395	58.00%
1000	



"DiscountOffer values standardized"

3. Find rows where both “Discount Price” and “Discount Offer” are null, then fill “Discount Price” with the category’s average discount price.

	D	E	H	I	M	N	O	P	Q
	Category	Individual_category	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Percentage disc		=UNIQUE(Table1[Category])		
Bottom Wear	jeans		824	1499	45%				
Bottom Wear	track-pants		517	1149	55%				
Topwear	shirts		629	1399	55%				
Lingerie & Sl	shapewear		893	1295	31%				
Western	tshirts		599		35%				
Western	tops		599		40%				
Bottom Wear	trousers		599	1499	60%				
Western	tops		1395		58%				
Western	tshirts		1098						
Bottom Wear	jeans		2749						
Sports Wear	tights		1214	2699	55%				



Used UNIQUE to extract unique category names

	D	E	H	I	M	N	O	P	Q	R
	Category	Individual_category	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Percentage disc		=AVERAGEIF(Table1[Category],O2,Table1[Percentage disc])			
Bottom Wear	jeans		824	1499	45%					
Bottom Wear	track-pants		517	1149	55%					
Topwear	shirts		629	1399	55%					
Lingerie & Sl	shapewear		893	1295	31%					
Western	tshirts		599		35%					
Western	tops		599		40%					
Bottom Wear	trousers		599	1499	60%					
Western	tops		1395		58%					
Western	tshirts		1098							
Bottom Wear	jeans		2749							
Sports Wear	tights		1214	2699	55%					

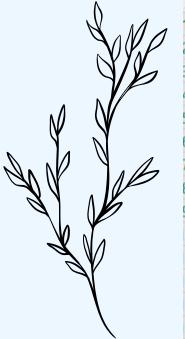
Used AVERAGEIF to calculate the average percentage of the selected category

	D	E	H	I	M	N	O	P
	Category	Individual_category	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Percentage disc			
Bottom Wear	jeans		824	1499	45%			
Bottom Wear	track-pants		517	1149	55%			
Topwear	shirts		629	1399	55%			
Lingerie & Sl	shapewear		893	1295	31%			
Western	tshirts		599		35%			
Western	tops		599		40%			
Bottom Wear	trousers		599	1499	60%			
Western	tops		1395		58%			
Western	tshirts		1098					
Bottom Wear	jeans		2749					
Sports Wear	tights		1214	2699	55%			

=XLOOKUP(D10,\$O\$2:\$O\$9,\$P\$2:\$P\$9,,0)

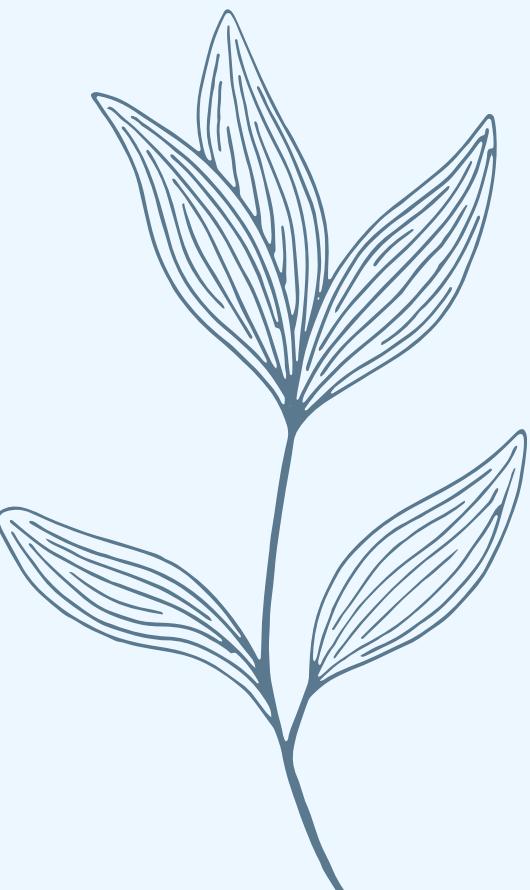
Used XLOOKUP to fill blank cells with the average % of their matching category.

4. Replace all the null values in “Size Option” column with the text “Not Available”.



A screenshot of an Excel spreadsheet titled "Myntra Fasion Clothing (1)". The "SizeOption" column is selected, and a "Text Filters" dialog box is open. The dialog lists several size categories such as XXS, XS, S, M, L, XL, XXL, and others, each with a checkbox. The "OK" button at the bottom left of the dialog is highlighted.

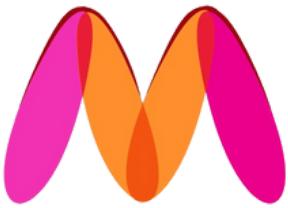
Null values, if any, in the Sizes column were to be found and replaced with ‘Not Available’. But upon querying using the FILTER button on the table, no blanks were found



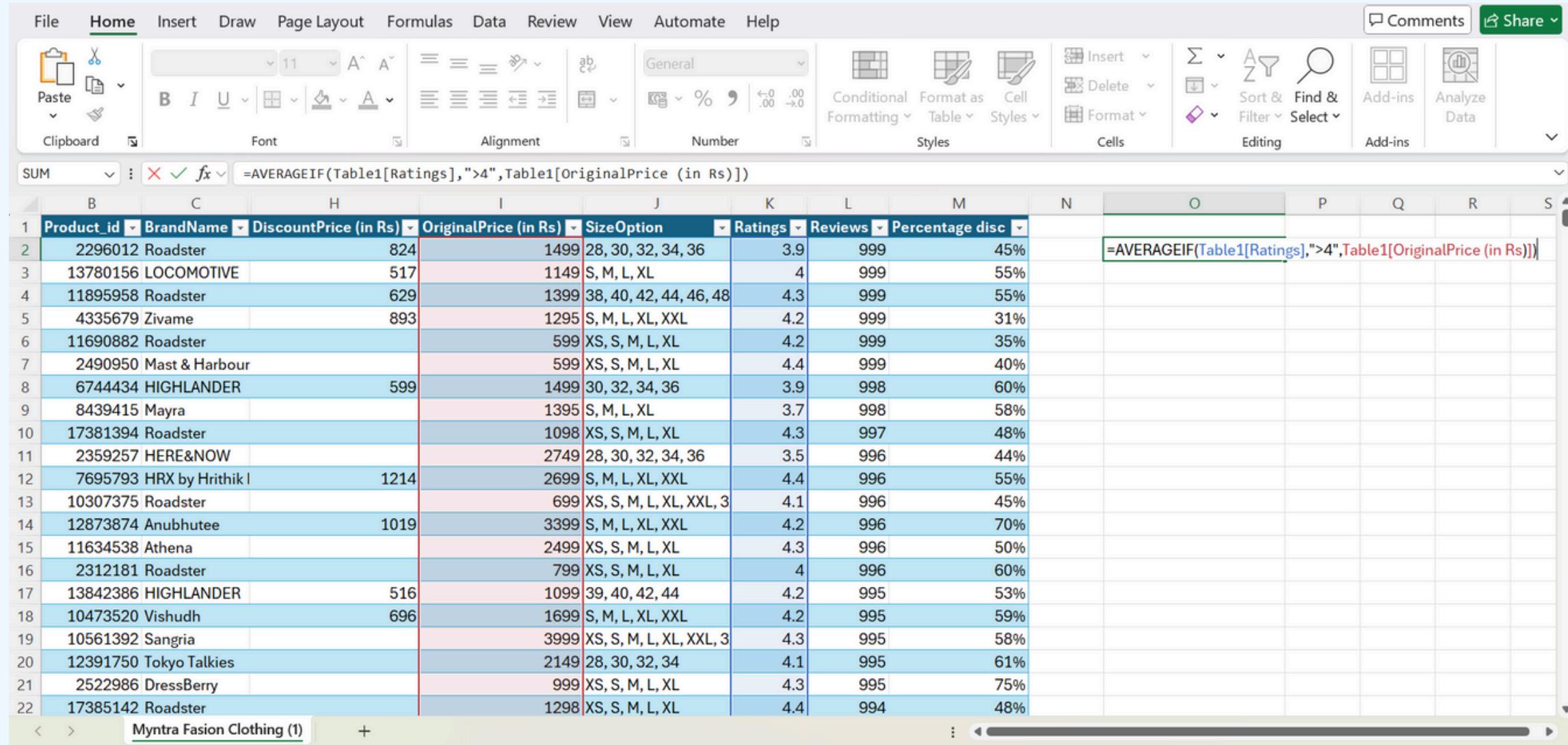
A screenshot of an Excel spreadsheet titled "Myntra Fasion Clothing (1)". The "SizeOption" column is selected, and a "Find and Replace" dialog box is open. The "Find what:" field contains "BLANK". The "Replace with:" field is empty. A message box from Microsoft Excel states: "We couldn't find anything to replace. Click Options for more ways to search." Below this, another message says: "FYI: It's possible the data you're trying to replace is in a protected sheet. Excel can't replace data in protected sheets." The "OK" button in the dialog is highlighted.

Upon checking, no null values were found to be replaced with ‘Not Available’.

B. Data Analysis



I. Calculate the overall average original price of products with ratings greater than 4.



The screenshot shows an Excel spreadsheet titled "Myntra Fasion Clothing (1)". The table has columns for Product_id, BrandName, DiscountPrice (in Rs), OriginalPrice (in Rs), SizeOption, Ratings, Reviews, and Percentage disc. The formula bar at the top displays the function =AVERAGEIF(Table1[Ratings],>4,Table1[OriginalPrice (in Rs)]). The table contains 22 rows of data, mostly for Roadster products, with various brand names like Roadster, LOCOMOTIVE, Zivame, etc., and original prices ranging from 516 to 1499.

Product_id	BrandName	DiscountPrice (in Rs)	OriginalPrice (in Rs)	SizeOption	Ratings	Reviews	Percentage disc
2296012	Roadster	824	1499	28, 30, 32, 34, 36	3.9	999	45%
13780156	LOCOMOTIVE	517	1149	S, M, L, XL	4	999	55%
11895958	Roadster	629	1399	38, 40, 42, 44, 46, 48	4.3	999	55%
4335679	Zivame	893	1295	S, M, L, XL, XXL	4.2	999	31%
11690882	Roadster		599	XS, S, M, L, XL	4.2	999	35%
2490950	Mast & Harbour		599	XS, S, M, L, XL	4.4	999	40%
6744434	HIGHLANDER	599	1499	30, 32, 34, 36	3.9	998	60%
8439415	Mayra		1395	S, M, L, XL	3.7	998	58%
17381394	Roadster		1098	XS, S, M, L, XL	4.3	997	48%
2359257	HERE&NOW		2749	28, 30, 32, 34, 36	3.5	996	44%
7695793	HRX by Hrithik I	1214	2699	S, M, L, XL, XXL	4.4	996	55%
10307375	Roadster		699	XS, S, M, L, XL, XXL, 3	4.1	996	45%
12873874	Anubhutee	1019	3399	S, M, L, XL, XXL	4.2	996	70%
11634538	Athena		2499	XS, S, M, L, XL	4.3	996	50%
2312181	Roadster		799	XS, S, M, L, XL	4	996	60%
13842386	HIGHLANDER	516	1099	39, 40, 42, 44	4.2	995	53%
10473520	Vishudh	696	1699	S, M, L, XL, XXL	4.2	995	59%
10561392	Sangria		3999	XS, S, M, L, XL, XXL, 3	4.3	995	58%
12391750	Tokyo Talkies		2149	28, 30, 32, 34	4.1	995	61%
2522986	DressBerry		999	XS, S, M, L, XL	4.3	995	75%
17385142	Roadster		1298	XS, S, M, L, XL	4.4	994	48%

We calculate the average original price only for products with ratings greater than 4 to understand the pricing of highly-rated products, rather than including low-rated ones that may skew the average.

1966.667473

it is the average of all the original prices of products whose ratings are greater than 4.

2. Count the number of products with a discount offer greater than 50% off

The screenshot shows a Microsoft Excel spreadsheet with a table of product data. The table has columns for Product_id, BrandName, DiscountPrice (in Rs), OriginalPrice (in Rs), Ratings, Reviews, and Percentage disc. A formula =COUNTIF(Table1[Percentage disc], ">0.50") is entered in cell O2, which displays the result 232123.

	B	C	H	I	K	L	M	N	O	P	Q
1	Product_id	BrandName	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Ratings	Reviews	Percentage disc		=COUNTIF(Table1[Percentage disc], ">0.50")		
2	2296012 Roadster		824	1499	3.9	999	45%				
3	13780156 LOCOMOTIVE		517	1149	4	999	55%				
4	11895958 Roadster		629	1399	4.3	999	55%				
5	4335679 Zivame		893	1295	4.2	999	31%				
6	11690882 Roadster			599	4.2	999	35%				
7	2490950 Mast & Harbour			599	4.4	999	40%				
8	6744434 HIGHLANDER		599	1499	3.9	998	60%				
9	8439415 Mayra			1395	3.7	998	58%				
10	17381394 Roadster			1098	4.3	997	48%				
11	2359257 HERE&NOW			2749	3.5	996	44%				
12	7695793 HRX by Hrithik R		1214	2699	4.4	996	55%				
13	10307375 Roadster			699	4.1	996	45%				
14	12873874 Anubhutee		1019	3399	4.2	996	70%				
15	11634538 Athena			2499	4.3	996	50%				
16	2312181 Roadster			799	4	996	60%				
17	13842386 HIGHLANDER		516	1099	4.2	995	53%				
18	10473520 Vishudh		696	1699	4.2	995	59%				
19	10561392 Sangria			3999	4.3	995	58%				
20	12391750 Tokun Talkies			2149	4.1	995	61%				

1

COUNTIF counts the number of cells in a range that meet a specific condition.

2

We want to count products with >50% discount.

3

Range contains all discount values.

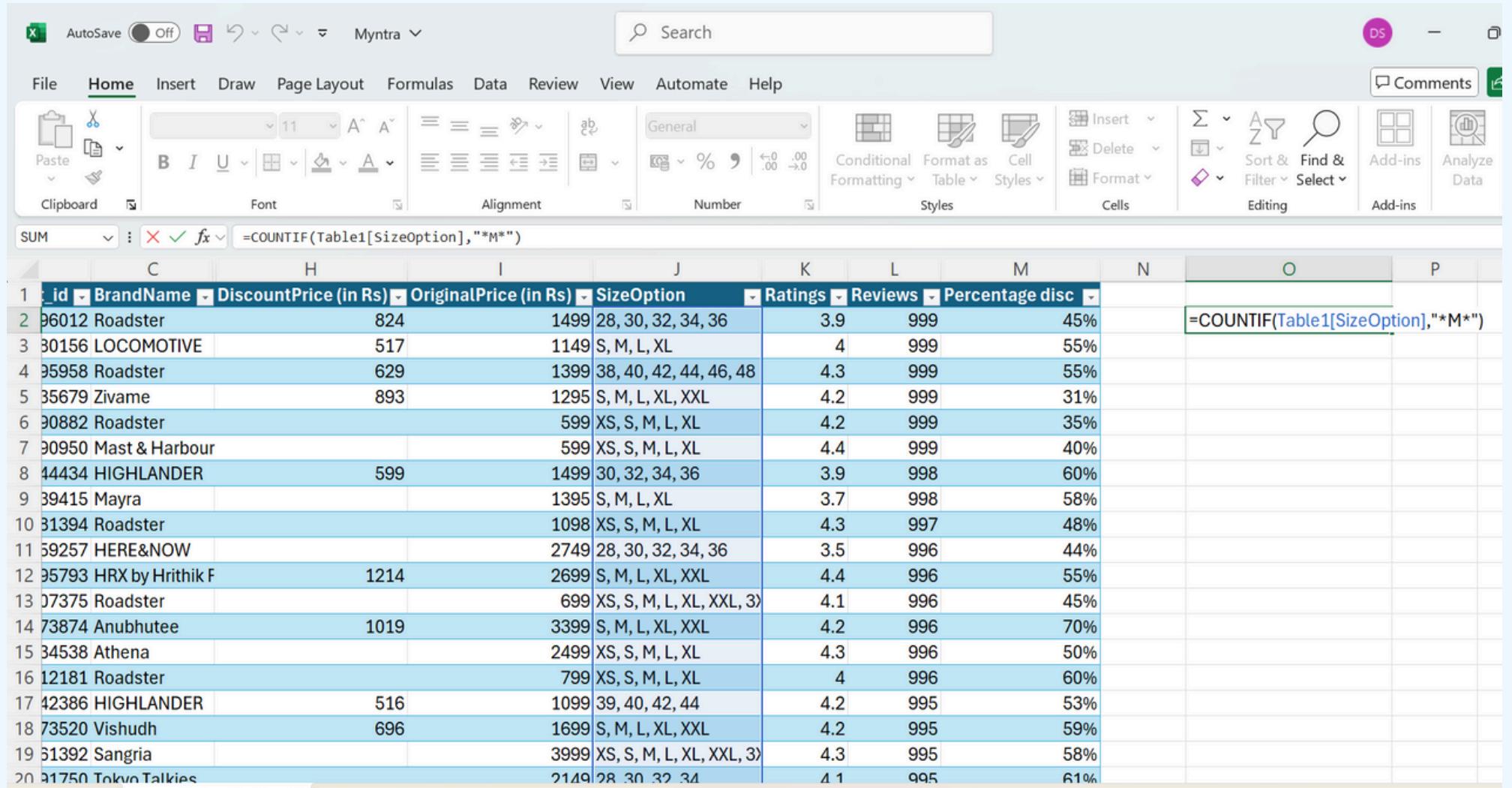
4

">0.50" checks which discounts are greater than 50%.

5

The result is the total number of products with a discount over 50%.

3. Count the number of products available in size ‘M’.



A screenshot of Microsoft Excel showing a table of product data. The table has columns for Product ID, Brand Name, Discount Price, Original Price, Size Options, Ratings, Reviews, and Percentage discount. The formula bar at the top shows the formula =COUNTIF(Table1[SizeOption],"*M*").

C	H	I	J	K	L	M	N	O	P
1	id	BrandName	DiscountPrice (in Rs)	OriginalPrice (in Rs)	SizeOption	Ratings	Reviews	Percentage disc	
2	96012	Roadster	824	1499	28, 30, 32, 34, 36	3.9	999	45%	
3	30156	LOCOMOTIVE	517	1149	S, M, L, XL	4	999	55%	
4	95958	Roadster	629	1399	38, 40, 42, 44, 46, 48	4.3	999	55%	
5	35679	Zivame	893	1295	S, M, L, XL, XXL	4.2	999	31%	
6	90882	Roadster		599	XS, S, M, L, XL	4.2	999	35%	
7	90950	Mast & Harbour		599	XS, S, M, L, XL	4.4	999	40%	
8	44434	HIGHLANDER	599	1499	30, 32, 34, 36	3.9	998	60%	
9	39415	Mayra		1395	S, M, L, XL	3.7	998	58%	
10	31394	Roadster		1098	XS, S, M, L, XL	4.3	997	48%	
11	59257	HERE&NOW		2749	28, 30, 32, 34, 36	3.5	996	44%	
12	95793	HRX by Hrithik F	1214	2699	S, M, L, XL, XXL	4.4	996	55%	
13	07375	Roadster		699	XS, S, M, L, XL, XXL, 3	4.1	996	45%	
14	73874	Anubhutee	1019	3399	S, M, L, XL, XXL	4.2	996	70%	
15	34538	Athena		2499	XS, S, M, L, XL	4.3	996	50%	
16	12181	Roadster		799	XS, S, M, L, XL	4	996	60%	
17	42386	HIGHLANDER	516	1099	39, 40, 42, 44	4.2	995	53%	
18	73520	Vishudh		1699	S, M, L, XL, XXL	4.2	995	59%	
19	61392	Sangria		3999	XS, S, M, L, XL, XXL, 3	4.3	995	58%	
20	91750	TokyoTalkies		2149	28, 30, 32, 34	4.1	995	61%	

01

The formula counts all rows where the SizeOption contains the letter “M”.

308460

02

The * symbols are wildcards that let Excel find “M” even if other text appears before or after it.

03

This ensures all sizes including “M” (like “S, M, L”) are counted.

4. Create a column labeling products as “High Discount” (Disc. Offer>50%) or “Low Discount”.

I	M	N	O	P
Price (in Rs)	Percentage disc	Disc Label		
1499	45%	=IF([Percentage disc]>50%,"High Disc","low Disc")		
1149	55%			
1399	55%			
1295	31%			
599	35%			
599	40%			
1499	60%			

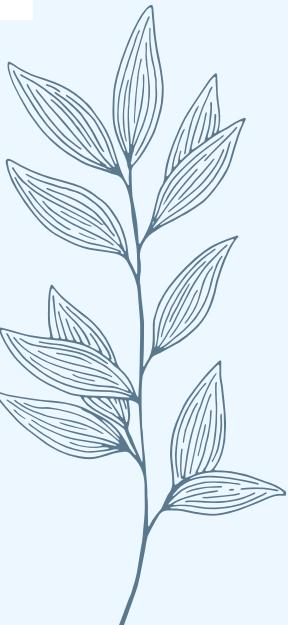
1. The IF function checks whether the Discount Offer is greater than 0.5 (i.e., 50%).

2. If true, it labels the product as “High Discount”.

3. If false, it labels it as “Low Discount”.

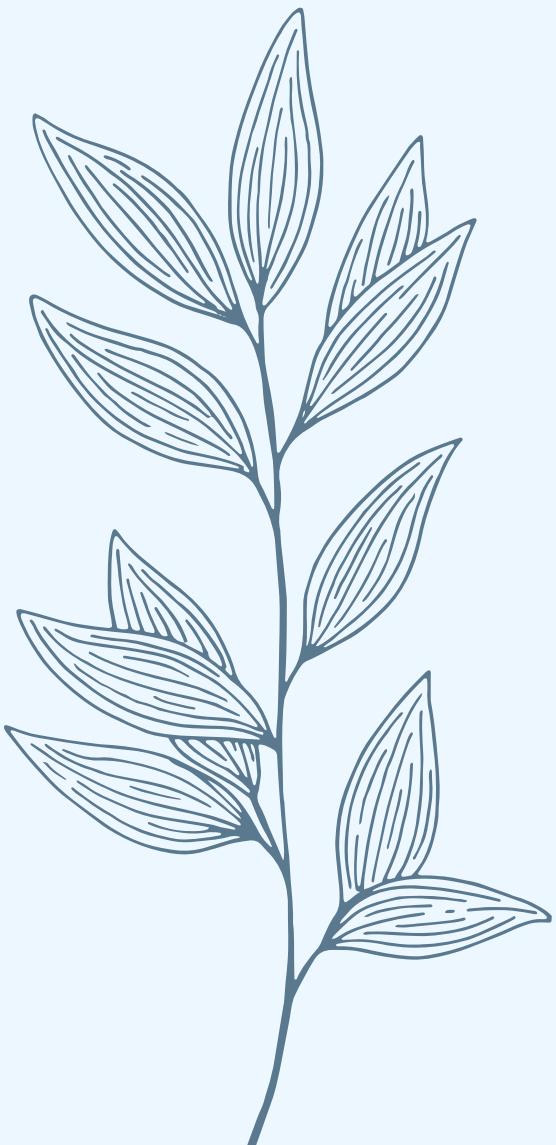
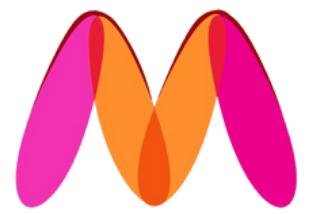


M	N
45%	low Disc
55%	High Disc
55%	High Disc
31%	low Disc
35%	low Disc
40%	low Disc
60%	High Disc
58%	High Disc
48%	low Disc
44%	low Disc
55%	High Disc
45%	low Disc
70%	High Disc
50%	low Disc
60%	High Disc

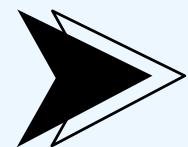




C. Data Retrieval and Lookup



1. Using XLOOKUP to find Brand, Price, and Rating for Product_ID 11226634



11226634			
Brand	Original Price	Rating	
=XLOOKUP(O2,Table1[Product_id],Table1[BrandName],,0	XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode],		

Using XLOOKUP to search a value in a column and return corresponding data from another column

Return Array
“Brand Column,,”

Return Array
“Original price Column,,”

Return Array
“Rating Column,,”

11226634			
Brand	Original Price	Rating	
Maniac	=XLOOKUP(O2,Table1[Product_id],Table1[OriginalPrice (in Rs)],,0	XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])	Searches for an Exact match, if not found return #N/A

11226634			
Brand	Original Price	Rating	
Maniac	1199	=XLOOKUP(O2,Table1[Product_id],Table1[Ratings],,0	



11226634			
Brand	Original Price	Rating	
Maniac	1199		3.9

2. Using INDEX & MATCH to find the Discount Price for Product_ID 6744434.

6744434

=INDEX(Table1[DiscountPrice (in Rs)],MATCH(O2,Table1[Product_id],0))

Table1[Discount Price] → Column from which we want the result.

MATCH Finds the row number of Product_ID 6744434.

6744434
Discount Price

INDEX returns the Discount Price at that row.

We use INDEX & MATCH to look up the Discount Price for Product_ID 6744434 flexibly, even if the lookup column isn't first.

3. Using nested XLOOKUP to find any product detail using Product_ID

A nested XLOOKUP formula is used when you need to perform multiple lookups within a single formula.



P3	B	C	D	E	H	I	M	N	O	P	Q	R	S	T
1	Product_id	BrandName	Category	Individual_category	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Percentage disc	Product ID	BrandName					
2	2296012	Roadster	Bottom Wea	jeans	824	1499	0.45	7695793	=XLOOKUP(O3,Table1[Product_id],XLOOKUP(P2,Table1[[#Headers],[BrandName]:[Individual_category]],Table1[[BrandName]:[Individual_category]],,0))					
3	13780156	LOCOMOTIVE	Bottom Wea	track-pants	517	1149	0.55							
4	11895958	Roadster	Topwear	shirts	629	1399	0.55							
5	4335679	Zivame	Lingerie & Sl	shapewear	893	1295	0.31							
6	11690882	Roadster	Western	tshirts		599	0.35							
7	2490950	Mast & Harbour	Western	tops		599	0.40							
8	6744434	HIGHLANDER	Bottom Wea	trousers	599	1499	0.60							
9	8439415	Mayra	Western	tops		1395	0.58							
10	17381394	Roadster	Western	tshirts		1098								
11	2359257	HERE&NOW	Bottom Wea	jeans		2749								
12	7695793	HRX by Hrithik	Sports Wear	tights	1214	2699	0.55							

This nested XLOOKUP first finds the column for BrandName, then looks up the ProductID to return the corresponding value from that column.



Product ID	BrandName
7695793	HRX by Hrithik Roshan

Nested XLOOKUP finds BrandName for ProductID 7695793, returning HRX by Hrithik Roshan



Thank You