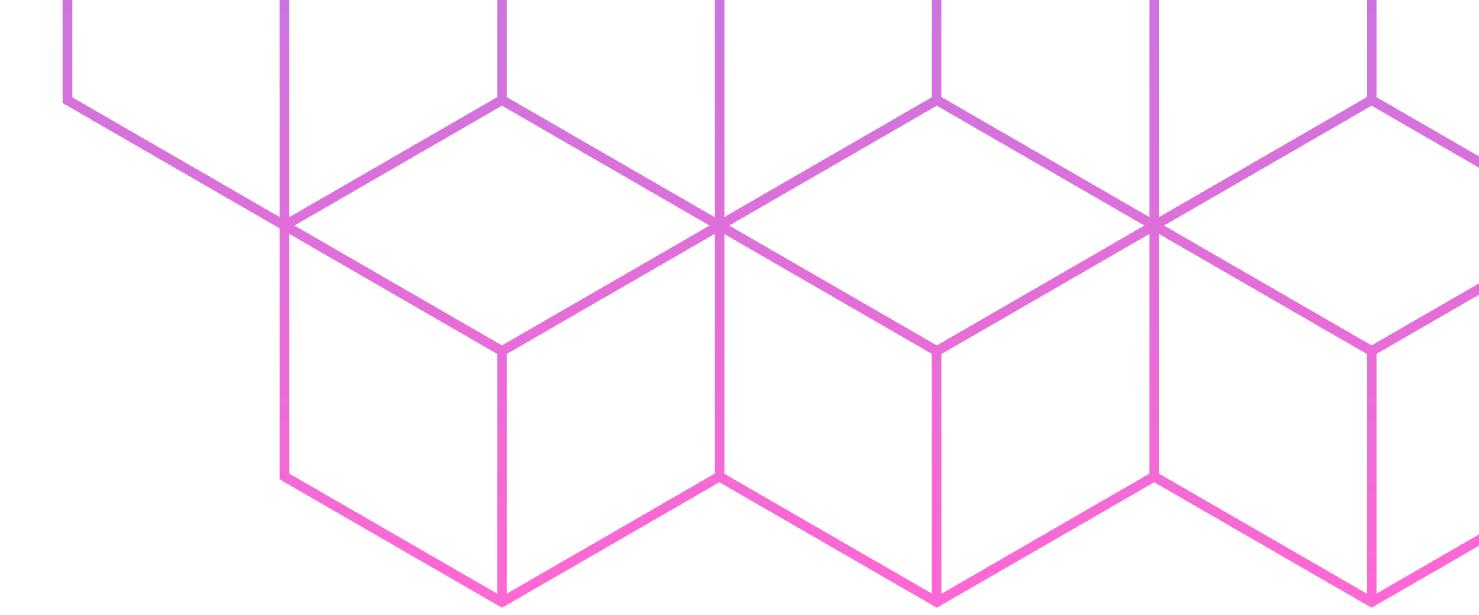


Myntra



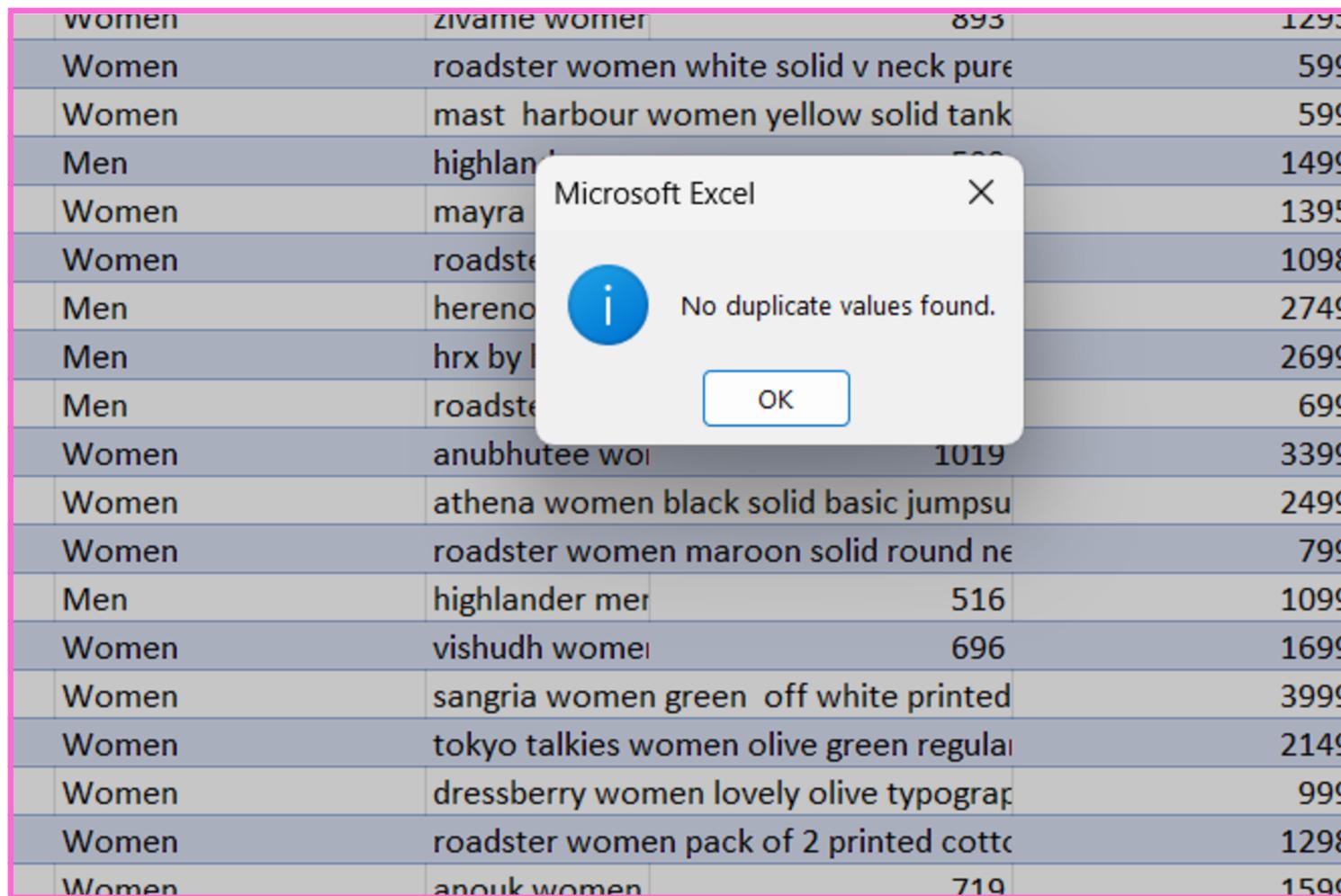
Myntra Case Study

Presented by: Nikhil Take.



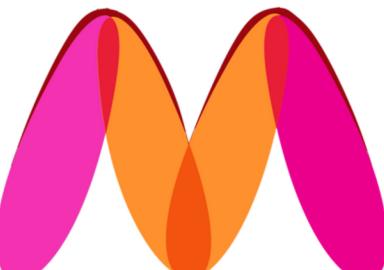
A. Data Cleaning and Preparation

1. Check for duplicate values in your dataset and remove them.



women	zivame women	895	1299
Women	roadster women white solid v neck pure	599	
Women	mast harbour women yellow solid tank	599	
Men	highlan	599	1499
Women	mayra	1395	
Women	roadste	1098	
Men	hereno	2749	
Men	hrx by l	2699	
Men	roadste	699	
Women	anubhutee wo	1019	3399
Women	athena women black solid basic jumpsu	2499	
Women	roadster women maroon solid round ne	799	
Men	highlander mer	516	1099
Women	vishudh wome	696	1699
Women	sangria women green off white printed	3999	
Women	tokyo talkies women olive green regul	2149	
Women	dressberry women lovely olive typogr	999	
Women	roadster women pack of 2 printed cottc	1298	
Women	anouk women	719	1599

- ✓ **Process to Check for Duplicates in Excel:**
- 1 Select the Dataset → Click & drag to highlight all data.
 - 2 Go to "Data" Tab → Click on "Remove Duplicates" in the toolbar.
 - 3 Select Columns → Choose relevant columns to check.
 - 4 Click "OK" → Excel will show a message if duplicates exist.
- 🔍 Result: No duplicate values were found in the dataset! ✓
- 🌟 The dataset is already clean and does not require any duplicate removal.

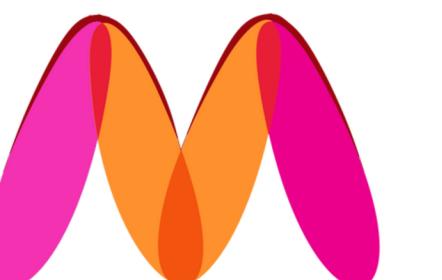


Myntra

A. Data Cleaning and Preparation

2. Standardize the "DiscountOffer" column to a single format, ensuring all values are uniform

	DiscountPrice (in Rs)	OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Discount Offer Standardize
	824	1499	45% 28, 30, 32, 34,	3.9	999		674.55
	517	1149	55% S, M, L, XL	4	999		631.95
	629	1399	55% 38, 40, 42, 44,	4.3	999		769.45
	893	1295	31% S, M, L, XL, XXL	4.2	999		401.45
women white solid v neck pure		599	35% XS, S, M, L, XL	4.2	999		209.65
women yellow solid tank		599	40% XS, S, M, L, XL	4.4	999		239.6
	599	1499	60% 30, 32, 34, 36	3.9	998		899.4
broidered a line pure cot		1395	58% S, M, L, XL	3.7	998		809.1
women pack of 2 solid t shirts		1098	XS, S, M, L, XL	4.3	997		0
blue slim fit mid rise clear		2749	28, 30, 32, 34,	3.5	996		0
	1214	2699	55% S, M, L, XL, XXL	4.4	996		1484.45
navy blue white striped p		699	XS, S, M, L, XL,	4.1	996		0
	1019	3399	70% S, M, L, XL, XXL	4.2	996		2379.3
black solid basic jumpsu		2499	50% XS, S, M, L, XL	4.3	996		1249.5
women maroon solid round ne		799	60% XS, S, M, L, XL	4	996		479.4
	516	1099	53% 39, 40, 42, 44	4.2	995		582.47
	696	1699	59% S, M, L, XL, XXL	4.2	995		1002.41
women green off white printed		3999	XS, S, M, L, XL,	4.3	995		0
women olive green regulai		2149	61% 28, 30, 32, 34	4.1	995		1310.89
women lovely olive typograp		999	75% XS, S, M, L, XL	4.3	995		749.25
women pack of 2 printed cott		1298	XS, S, M, L, XL	4.4	994		0
	719	1599	55% XS, S, M, L, XL,	3.8	993		879.45
women wired removable pad		849	S, M, L, XL	4.2	993		0
women mauve stretchable		1499	50% XS, S, M, L, XL	4.2	993		749.5
floral cinched waist top v		1299	62% XS, S, M, L, XL,	4.3	992		805.38
women maroon striped top		1299	57% S, M, L, XL	4.2	991		740.43



Myntra

I have standardized the DiscountOffer column by converting all values into Rs. format using the following formula:

```
=IF(ISNUMBER(SEARCH("Rs",[@DiscountOffer])),  
SUBSTITUTE(@DiscountOffer, "Rs. ", ""),  
[@[OriginalPrice (in Rs)]] * SUBSTITUTE(@DiscountOffer,  
"%", "") / 100)
```

Steps Followed:

- 1 Checked if the value contains "Rs". If yes, removed "Rs." and converted it into a number.
- 2 If the value was in percentage format (e.g., 50%), calculated the discount amount using the original price.
- 3 The final column now contains only discount amounts in Rs., ensuring uniformity.

Outcome:

- All discount values are now standardized into a single format (Rs.).
- This ensures consistency and makes further analysis easier. 

A. Data Cleaning and Preparation

3. Identify rows where both "DiscountPrice" and "DiscountOffer" are null and fill the "DiscountPrice" with the average discount price of the respective category.

H	I	J	K	L	M	N	O
DiscountPrice (in Rs)	OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Discount Offer Standardize	Updated DiscountPrice (in Rs)
824	1499	45% 28, 30, 32, 34,	3.9	999	674.55	824	
517	1149	55% S, M, L, XL	4	999	631.95	517	
629	1399	55% 38, 40, 42, 44,	4.3	999	769.45	629	
893	1295	31% S, M, L, XL, XXL	4.2	999	401.45	893	
	599	35% XS, S, M, L, XL	4.2	999	209.65	0	
	599	40% XS, S, M, L, XL	4.4	999	239.6	0	
599	1499	60% 30, 32, 34, 36	3.9	998	899.4	599	
	1395	58% S, M, L, XL	3.7	998	809.1	0	
	1098	XS, S, M, L, XL	4.3	997	0	1790.085106	
	2749	28, 30, 32, 34,	3.5	996	0	1351.207867	
1214	2699	55% S, M, L, XL, XXL	4.4	996	1484.45	1214	
	699	XS, S, M, L, XL,	4.1	996	0	1138.634061	
1019	3399	70% S, M, L, XL, XXL	4.2	996	2379.3	1019	
	2499	50% XS, S, M, L, XL	4.3	996	1249.5	0	
	799	60% XS, S, M, L, XL	4	996	479.4	0	
516	1099	53% 39, 40, 42, 44	4.2	995	582.47	516	
696	1699	59% S, M, L, XL, XXL	4.2	995	1002.41	696	
	3999	XS, S, M, L, XL,	4.3	995	0	923.6945977	
	2149	61% 28, 30, 32, 34	4.1	995	1310.89	0	
	999	75% XS, S, M, L, XL	4.3	995	749.25	0	
	1298	XS, S, M, L, XL	4.4	994	0	1790.085106	
719	1599	55% XS, S, M, L, XL,	3.8	993	879.45	719	
as bra sb08	849	S, M, L, XL	4.2	993	0	811.9028417	
	1499	50% XS, S, M, L, XL	4.2	993	749.5	0	
	1299	62% XS, S, M, L, XL,	4.3	992	805.38	0	
	1299	57% S, M, L, XL	4.2	991	740.43	0	

📌 **Handling Missing Discount Prices**
To ensure data consistency, missing values in the DiscountPrice column were filled using the average discount price of the respective category.

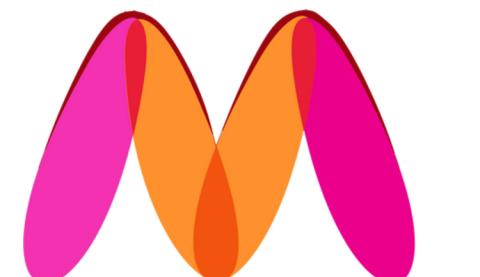
✓ **Steps Taken:**

- 1 Identified rows where both DiscountPrice and DiscountOffer were empty.
- 2 Calculated the average DiscountPrice for each Category.
- 3 Replaced missing values with the category-wise average using the formula:

```
=IF(AND(ISBLANK([@DiscountPrice (in Rs)]),  
ISBLANK([@DiscountOffer])), AVERAGEIF([Category],  
[@Category], [DiscountPrice (in Rs)]), [@DiscountPrice (in  
Rs)])
```

🎯 **Final Outcome:**

- Missing values were replaced, ensuring accurate price data.
- The dataset is now standardized for further analysis. 🚀



Myntra

A. Data Cleaning and Preparation

4. Replace all null values in the "SizeOption" column with the text "Not Available."

Product Description	Discounted Price (in Rs)	Original Price (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews
men navy blue slim fit mid rise clean look jeans	824	1499	45%	28, 30, 32, 34,	3.9	999
men black white solid slim fit track pants	517	1149	55%	S, M, L, XL	4	999
men navy white black geometric printed sustainable casual shirt	629	1399	55%	38, 40, 42, 44,	4.3	999
men black saree shapewear zi3023core0nude	893	1295	31%	S, M, L, XL, XXL	4.2	999
women white solid v neck pure cotton t shirt	599	35%	XS, S, M, L, XL	4.2	999	
women yellow solid tank top	599	40%	XS, S, M, L, XL	4.4	999	
men off white solid casual shirt	998					
men blue solid casual shirt	998					
women blue solid casual shirt	997					
women pink solid casual shirt	996					
women navy blue solid casual shirt	996					
women white solid casual shirt	996					
women black solid casual shirt	996					
women blue pink floral print a line kurta	696	2499	50%	XS, S, M, L, XL	4.3	996
women green off white printed kurta with palazzos dupatta		799	60%	XS, S, M, L, XL	4	996
women olive green regular fit solid cargos		1099	53%	39, 40, 42, 44	4.2	995
women lovely olive typography talking tee		1699	59%	S, M, L, XL, XXL	4.2	995
women pack of 2 printed cotton t shirt		3999		XS, S, M, L, XL	4.3	995
women mustard yellow gold toned printed straight kurta		2149	61%	28, 30, 32, 34	4.1	995
women black non wired removable pads high coverage medium impact sports bra sb08		999	75%	XS, S, M, L, XL	4.3	995
women mauve stretchable casual shirt		1298		XS, S, M, L, XL	4.4	994
women blue floral cinched waist top with puff sleeves		719	55%	XS, S, M, L, XL	3.8	993
women black non wired removable pads high coverage medium impact sports bra sb08		1599		S, M, L, XL	4.2	993
women mauve stretchable casual shirt		849		XS, S, M, L, XL	4.2	993
women blue floral cinched waist top with puff sleeves		1499	50%	XS, S, M, L, XL	4.2	993
women black non wired removable pads high coverage medium impact sports bra sb08		1299	62%	XS, S, M, L, XL	4.3	992

- No missing values were found in the "SizeOption" column.
- A Find & Replace operation was performed to check for null values, but no replacements were needed.
- All products already had valid size data, so no modifications were required.



Myntra

B. Data Analysis

1. Calculate the overall average original price for products with ratings greater than 4.

L	M	N	O	P
Ratings	Reviews	Discount Offer Standardize	Updated DiscountPrice (in Rs)	Avg Original Price
3.9	999	674.55	824	
4	999	631.95	517	
4.3	999	769.45	629	1966.667473
4.2	999	401.45	893	1966.667473
4.2	999	209.65	0	1966.667473
4.4	999	239.6	0	1966.667473
3.9	998	899.4	599	
3.7	998	809.1	0	
4.3	997	0	1790.085106	1966.667473
3.5	996	0	1351.207867	
4.4	996	1484.45	1214	1966.667473
4.1	996	0	1138.634061	1966.667473
4.2	996	2379.3	1019	1966.667473
4.3	996	1249.5	0	1966.667473
4	996	479.4	0	

 Calculation of Average Original Price for High-Rated Products

✓ To find the overall average original price for products with ratings greater than 4, I used the following formula:

=IF(L2>4, AVERAGEIF(L:L, ">4", I:I), "")

📌 This ensures that only products with a rating above 4 are considered, and the average original price is displayed in a separate column.



B. Data Analysis

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

Products > 50% OFF	210637
--------------------	--------

 Count of Products with Discount Offer Greater Than 50%

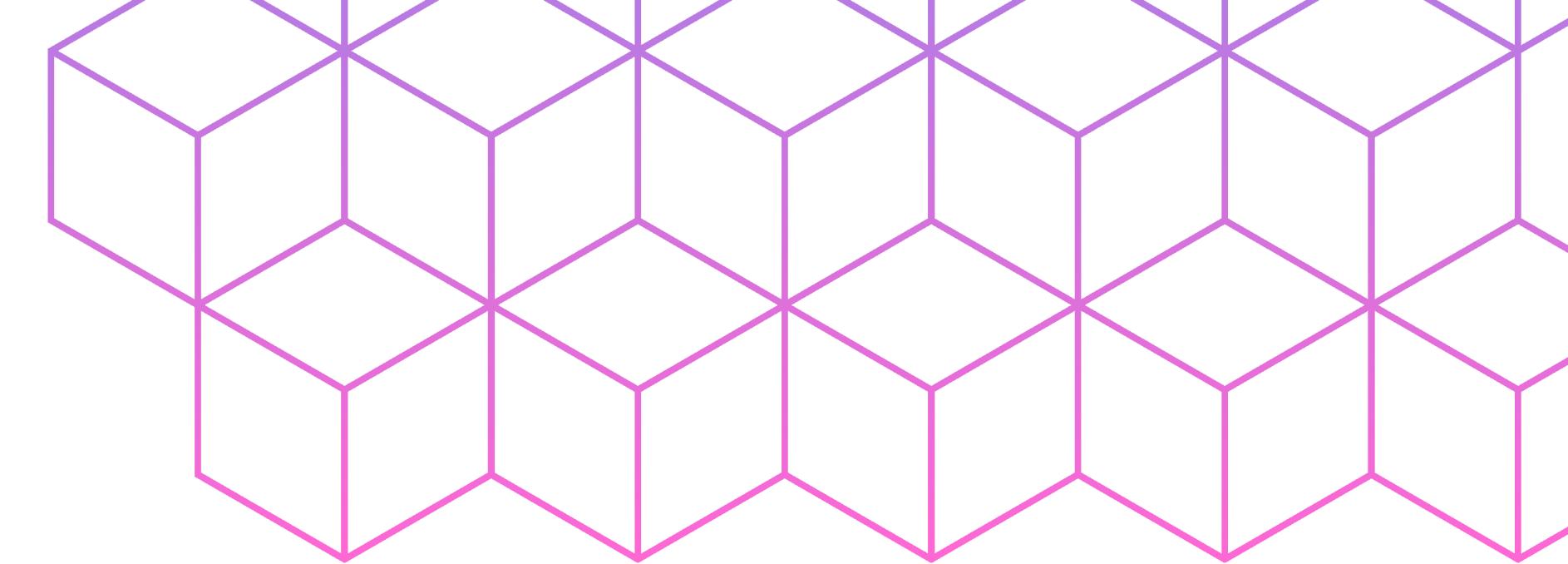
 To count the number of products where the discount offer is greater than 50%, I used the following formula:

=COUNTIF(J:J, ">50%")



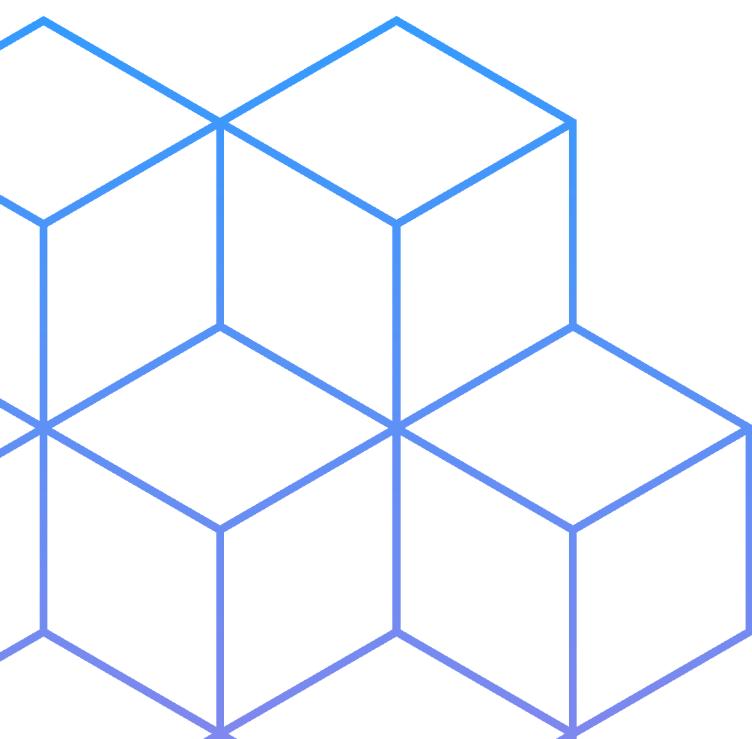
B. Data Analysis

3. Count the number of products available in size "M."



Products Available In M Size:-

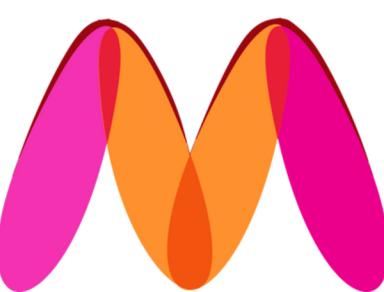
308460



To determine the number of products available in size "M", we used the formula:

Formula Used:

=COUNTIF(K:K, "*M*")

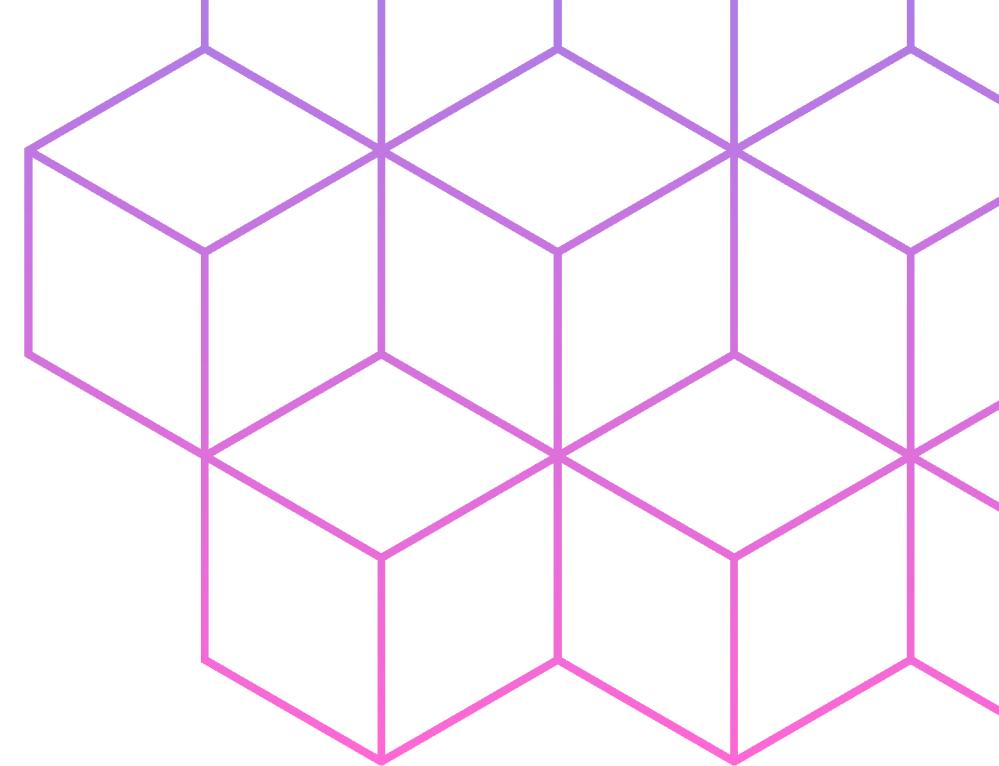


Myntra

B. Data Analysis

4. Create a new column to label the products as "High Discount" if the discount offer is greater than 50% OFF, otherwise label them as "Low Discount."

DiscountOffer	SizeOption	Ratings	Reviews	Discount Offer Standardize	Updated DiscountPrice (in Rs)	Column1
45%	28, 30, 32, 34,	3.9	999	674.55	824	Low Discount
55%	S, M, L, XL	4	999	631.95	517	High Discount
55%	38, 40, 42, 44,	4.3	999	769.45	629	High Discount
31%	S, M, L, XL, XXL	4.2	999	401.45	893	Low Discount
35%	XS, S, M, L, XL	4.2	999	209.65	0	Low Discount
40%	XS, S, M, L, XL	4.4	999	239.6	0	Low Discount
60%	30, 32, 34, 36	3.9	998	899.4	599	High Discount
58%	S, M, L, XL	3.7	998	809.1	0	High Discount
	XS, S, M, L, XL	4.3	997	0	1790.085106	Low Discount
	28, 30, 32, 34,	3.5	996	0	1351.207867	Low Discount
55%	S, M, L, XL, XXL	4.4	996	1484.45	1214	High Discount
	XS, S, M, L, XL,	4.1	996	0	1138.634061	Low Discount
70%	S, M, L, XL, XXL	4.2	996	2379.3	1019	High Discount
50%	XS, S, M, L, XL	4.3	996	1249.5	0	Low Discount
60%	XS, S, M, L, XL	4	996	479.4	0	High Discount
53%	39, 40, 42, 44	4.2	995	582.47	516	High Discount
59%	S, M, L, XL, XXL	4.2	995	1002.41	696	High Discount
	XS, S, M, L, XL,	4.3	995	0	923.6945977	Low Discount
61%	28, 30, 32, 34	4.1	995	1310.89	0	High Discount
75%	XS, S, M, L, XL	4.3	995	749.25	0	High Discount
	XS, S, M, L, XL	4.4	994	0	1790.085106	Low Discount
55%	XS, S, M, L, XL,	3.8	993	879.45	719	High Discount
	S, M, L, XL	4.2	993	0	811.9028417	Low Discount
50%	XS, S, M, L, XL	4.2	993	749.5	0	Low Discount
62%	XS, S, M, L, XL,	4.3	992	805.38	0	High Discount



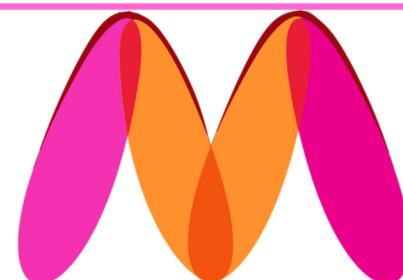
A new column 1 was created to label products as:

✓ High Discount (if discount offer is greater than 50%)

✓ Low Discount (if discount offer is 50% or less)

Formula Used:

=IF([@DiscountOffer]>50%, "High Discount", "Low Discount")



Myntra

C. Data Retrieval and Lookup

1. Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product_id "11226634".

i Have done Using XLOOKUP

Product_id "11226634"	
Using XLOOKUP:-	
Brand Name	Maniac
Price	1199
Rating	3.9

- Objective: Retrieve Brand Name, Price, and Rating for Product ID 11226634
- Formula Used:
- XLOOKUP:-

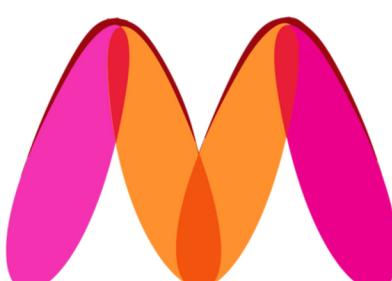
```
=XLOOKUP(11226634, B:B, C:C) ' Brand Name  
=XLOOKUP(11226634, B:B, I:I) ' Price  
=XLOOKUP(11226634, B:B, L:L) ' Rating
```

C. Data Retrieval and Lookup

2. Find the "DiscountPrice" for the product with the Product ID "6744434" using the INDEX and MATCH functions.

Disc Prize for Product ID :- 6744434	
	599 <- Disc.Price
Using Index and match Function	

- INDEX(H:H, row_number): Retrieves the value from the DiscountPrice (in Rs) column.
- MATCH(6744434, B:B, 0): Finds the row number where Product ID = 6744434.
- The combination of INDEX and MATCH efficiently finds the exact Discount Price.
- Formula Used:-
 $=INDEX(H:H, MATCH(6744434, B:B, 0))$



Myntra

C. Data Retrieval and Lookup

3.Utilize nested xlookup to find any column's detail of a product with it's product id.

Enter Product ID:-	Enter Column Name
11036420	Category
OUT Come	
Inner Wear & Sleep Wear	

Enter Product ID:-	Enter Column Name
4335679	BrandName
OUT Come	
Zivame	

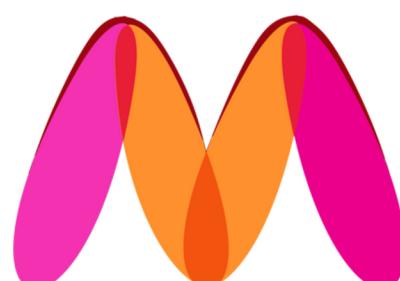
Enter Product ID:-	Enter Column Name
11036420	BrandName
OUT Come	
Jockey	

Formula Used:

=XLOOKUP(Product_ID, B:B,
XLOOKUP(Column_Name, A1:Z1, A:Z))

In the place of Product_ID & Column Name we need to give cell reference like given in below formula:-

=XLOOKUP(R3, B:B, XLOOKUP(S3, A1:Z1, A:Z))



Myntra

**Thank You for Viewing My
Presentation!**



Myntra