

BY JAGRUTI JADHAV

Analysis of MYNTRA FASHION APPAREL



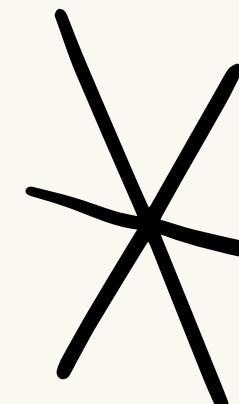
Agenda

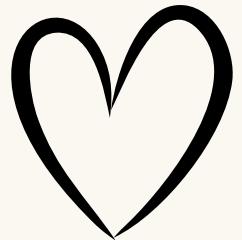
1. About the Company
2. Problem Statement
3. Our Goals
4. Raw Dataset
5. Data Cleaning and Preparation

Agenda

5. Cleaned Dataset
6. Data Analysis
7. Final Table
8. Data Retrieval and Lookup
9. Contact and Call To Action

C
C
C
C
C
C
C
C
C
C
C
C

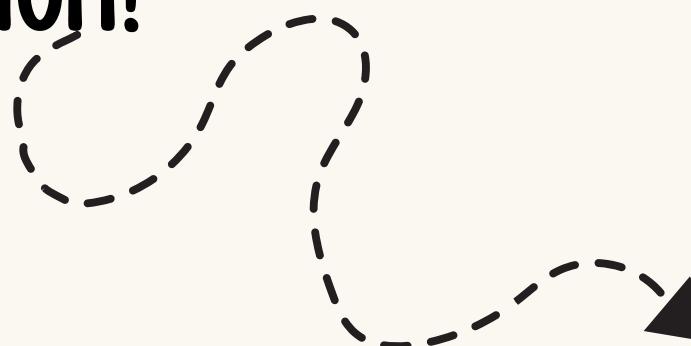
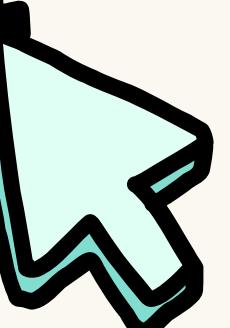




Introduction



Hi there! I'm Sarah and i work
at Myntra Fashion!



- Myntra is an Indian e-commerce company that sells fashion and lifestyle products online.
- It was founded in 2007–2008 to sell personalized gifts, but has since expanded to offer a wide range of products, including clothing, footwear, accessories, jewelry, and personal care products.
- Myntra works with over 6,000 brands, including H&M, Levis, Tommy Hilfiger, Nike, and Puma.
- The company is headquartered in Bengaluru, Karnataka, and services over 19,000 pin codes across India.

PROBLEM STATEMENTS

A. DATA CLEANING & PREPARATION

1. Check for duplicate values in your dataset and remove them.
2. Standardize the "DiscountOffer" column to a single format, ensuring all values are uniform.
3. Identify rows where both "DiscountPrice" and "DiscountOffer" are null and fill the "DiscountPrice" with the average discount price of the respective category.
4. Replace all null values in the "SizeOption" column with the text "Not Available."

B. DATA ANALYSIS

1. Calculate the overall average original price for products with ratings greater than 4.
2. Count the number of products with a discount offer greater than 50% OFF.
3. Count the number of products available in size "M."
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."

C. DATA RETRIEVAL & LOOK UP

1. Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product_id "11226634".
2. Find the "DiscountPrice" for the product with the Product ID "6744434" using the INDEX and MATCH functions.
3. Utilize nested xlookup to find any column's detail of a product with its product id.

Mind MAP





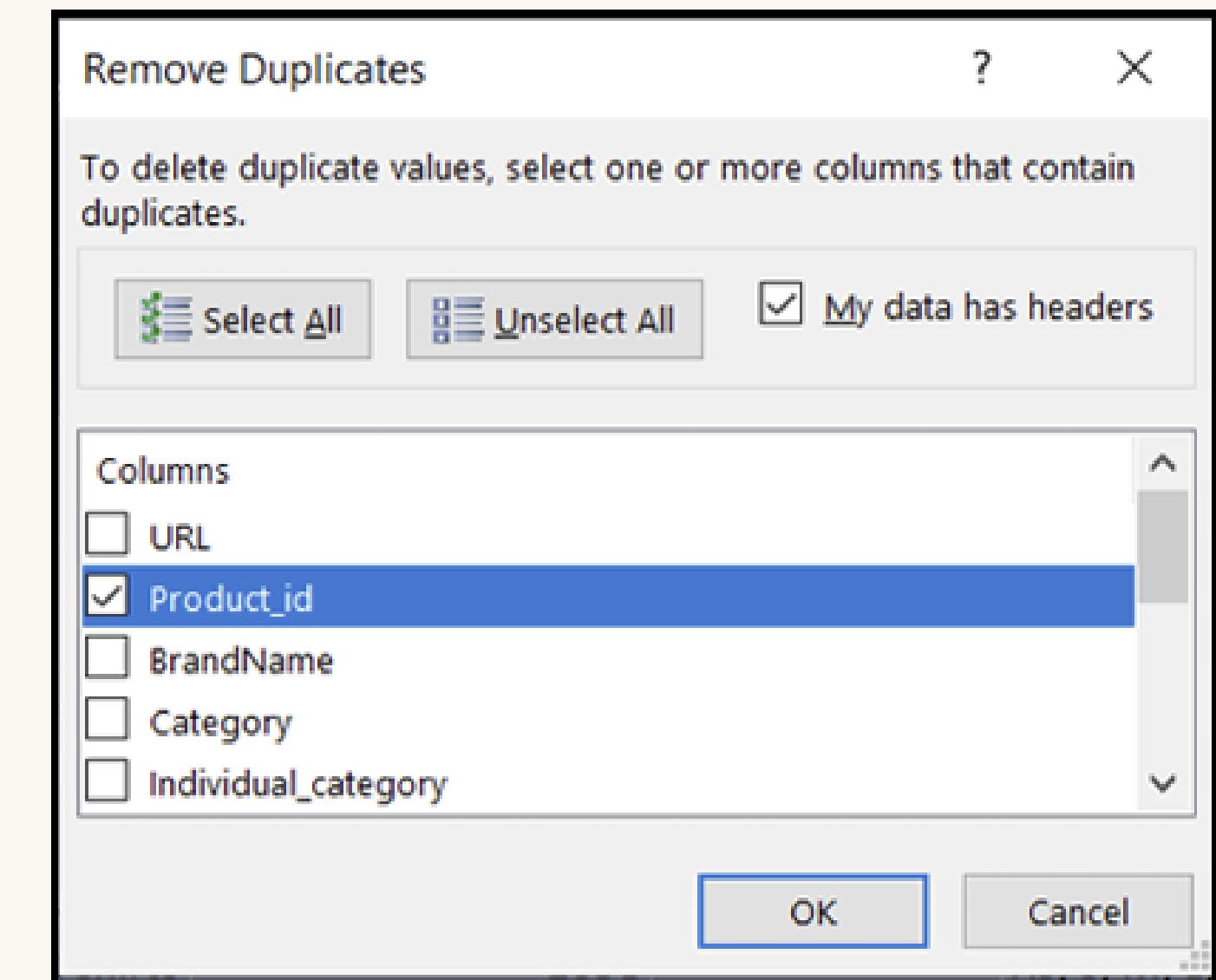
DATA CLEANING & PREPARATION

CHECK FOR DUPLICATE VALUES IN YOUR DATASET AND REMOVE THEM.

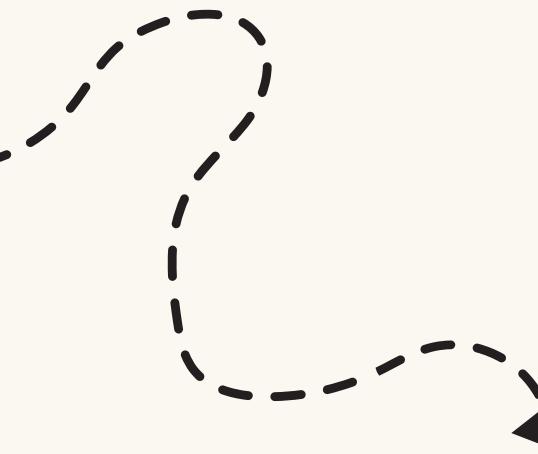
1. To remove duplicates we go to Data Menu > Remove Duplicates
2. Unselect All and then choose Product_id to remove duplicates based on Product_id.
3. We found no duplicates.

A screenshot of a Microsoft Excel spreadsheet. The data consists of three columns: Category (Men/Women), Product Name, and Price. A message box titled "Microsoft Excel" is displayed in the center, stating "No duplicate values found." with an "OK" button. The background shows several rows of data, such as "Men highlander men 599" and "Women mayra pink embroidered a line pure cot".

Category	Product Name	Price
Men	highlander men	599
Women	mayra pink embroidered a line pure cot	
Women	roadster women	
Men	herenow men	
Men	hrx by hrithik	
Men	roadster men	
Women	anubhutee women	
Women	athena women	
Women	roadster women maroon solid round neck	
Men	highlander men	516
Women	vishudh women	606



STANDARDIZE "DISCOUNTOFFER" COLUMN TO UNIFORM FORMAT.



I	J	K	L	M	N	O	P	Q
Original Price (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews				
1499 45% OFF	28, 30, 32, 34,		3.9	999	=Trim(Replace([@DiscountOffer],"Rs.",""))			
1149 55% OFF	S, M, L, XL		4	999				
1399 55% OFF	38, 40, 42, 44,		4.3	999				
1295 31% OFF	S, M, L, XL, XXL		4.2	999				
599 35% OFF	XS, S, M, L, XL		4.2	999				
599 40% OFF	XS, S, M, L, XL		4.4	999				
1499 60% OFF	30, 32, 34, 36		3.9	998				
1395 58% OFF	S, M, L, XL		3.7	998				
1098	XS, S, M, L, XL		4.3	997				
2749	28, 30, 32, 34,		3.5	996				
2699 55% OFF	S, M, L, XL, XXL		4.4	996				
699	XS, S, M, L, XL,		4.1	996				
2299 70% OFF	S, M, L, XL, XXL		4.2	996				

Now, the column will be standardized by converting percentage values into numbers using the formula displayed in the image

I	J	K	L	M	N	O	P	Q	R	S
Original Price (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Discounted Amount	Column				
1499 45% OFF	28, 30, 32, 34,		3.9	999	45% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1149 55% OFF	S, M, L, XL		4	999	55% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1399 55% OFF	38, 40, 42, 44,		4.3	999	55% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1295 31% OFF	S, M, L, XL, XXL		4.2	999	31% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
599 35% OFF	XS, S, M, L, XL		4.2	999	35% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
599 40% OFF	XS, S, M, L, XL		4.4	999	40% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1499 60% OFF	30, 32, 34, 36		3.9	998	60% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1395 58% OFF	S, M, L, XL		3.7	998	58% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
1098	XS, S, M, L, XL		4.3	997						
2749	28, 30, 32, 34,		3.5	996						
2699 55% OFF	S, M, L, XL, XXL		4.4	996	55% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					
699	XS, S, M, L, XL		4.1	996	55% OFF =IF(ISNUMBER(SEARCH("%",[@[Discounted Amount]])),LEFT([@Discounted Amount],SEARCH("%",[@[Discounted Amount]])-1)/100)*[@[OriginalPrice (in Rs)]],					

We have to remove "rupees" To remove the rupees we have used SUBSTITUTE function

"DISCOUNTPRICE" FILL WITH CATEGORY AVERAGE.

Rs	DiscountOffer	SizeOption	Ratings	Reviews	Discour	Column
599	60% OFF	S, M, L, XL				
599	45% OFF	S, M, L, XL				
1699	50% OFF	XS, S, M, L, XL				
1198	50% OFF	XS, S, M, L, XL				
1498	50% OFF	M, S, XS, XL, L				
1499	60% OFF	XS, S, M, L, XL				
1799	55% OFF	XS, S, M, L, XL				
1299	60% OFF, Hurry*	XS, S, M, L, XL				
1999	55% OFF	28, 30, 32, 34				
799	50% OFF	S, M, L, XL, XX				
899	50% OFF	S, M, L, XL, XX				
1299	50% OFF	S, M, L, XL				
699	45% OFF	S, M, L, XL, XX				
1295	31% OFF	S, M, L, XL, XX				
999	51% OFF	Onesize				
1499	40% OFF	XS, S, M, L, XL				
1699	55% OFF	XS, S, M, L, XL				
2299	63% OFF	S, M, L, XL, XX				
2699	50% OFF	28, 30, 32, 34				
399		Onesize				
2199	50% OFF	39, 40, 42, 44				
2599	66% OFF	S, M, L, XL, XX				
1699	60% OFF	26, 28, 30, 32				
799		XS, S, M, L, XL				

After converting, remove "off" from the column using "Text to Columns" and space delimiter.

Q	R	S	T	U	V	W	X	Y	Z	AA	AB
6	Column										
45	=IF([@Discounted %]="",AVERAGEIFS(Table1[[#All],[Discounted %]],Table1[[#All],[Category]],[@Category]),[@Discounted %])	55	55								
		55	55								
		31	31	48.8792							
		25	25								

Convert all numbers to percentages.

Figuring out the typical "Discount Offer" in %.

Discour	Column	Discour	
45% OFF	674.55	=IF([@Column1]:"", "",[@Column1]/	
55% OFF	631.95	[@OriginalPrice (in Rs)]*100)	
55% OFF	769.45	55	
31% OFF	401.45	31	
35% OFF	209.65	35	
40% OFF	239.6	40	
60% OFF	899.4	60	
58% OFF	809.1	58	

Calculate average discount percentage for quick “Average Discount Amount” determination.

O	P	Q	R	S	T	U	V	W
Column	Column3	Discount (%)	Discount (%)	Discount (%)	Discounted Amount			
674.55	45	45	45	45	= (Q2/100)*Table1[@[OriginalPrice (in Rs)]]			
631.95	55	55	55	55	631.95	631.95		
769.45	55	55	55	55	769.45	769.45		
401.45	31	31	31	31	401.45	401.45		

O	P	Q	R	S	T	U	V	W
Column	Discount (%)	Discount (%)	Discount (%)	Discount (%)				
674.55	45	674.55	=FLOOR.MATH([@[OriginalPrice (in Rs)]]-[@[Discounted Amount]])					
631.95	55	631.95						
769.45	55	769.45	629					
401.45	31	401.45	893					
209.65	35	209.65	389					

Now Finding the “Discount Price” using Floor Math Function

FINDING ROWS HAVING NULL VALUES BOTH IN "DISCOUNT PRICE" AND "DISCOUNT OFFER"

L	M	N	O	P	Q	R	S	T	U	V	W
tings	Reviews	Column	Column	Discou	Discou	Discou	Discou	NULL V			
3.9	999	45% OFF	674.55	45	674.55	824	=IF(AND(ISBLANK([@DiscountPrice (in Rs)]), ISBLANK([@DiscountOffer])), "NULL VALUES", "")				
4	999	55% OFF	631.95	55	631.95	517	ISBLANK([@DiscountOffer]), "NULL VALUES", "")				
4.3	999	55% OFF	769.45	55	769.45	629					
4.2	999	31% OFF	401.45	31	401.45	893					
4.2	999	35% OFF	209.65	35	209.65	389					
4.4	999	40% OFF	239.6	40	239.6	359					
3.9	998	60% OFF	899.4	60	899.4	599					
3.7	998	58% OFF	809.1	58	809.1	585					
4.3	997		47.72562	524.0273	573	NULL VALUES					
3.5	996		43.97481	1208.868	1540	NULL VALUES					
4.4	996	55% OFF	1484.45	55	1484.45	1214					
4.1	996		44.61667	311.8705	387	NULL VALUES					
4.2	996	70% OFF	2379.3	70	2379.3	1019					
4.3	996	50% OFF	1249.5	50	1249.5	1249					
4	996	60% OFF	479.4	60	479.4	319					

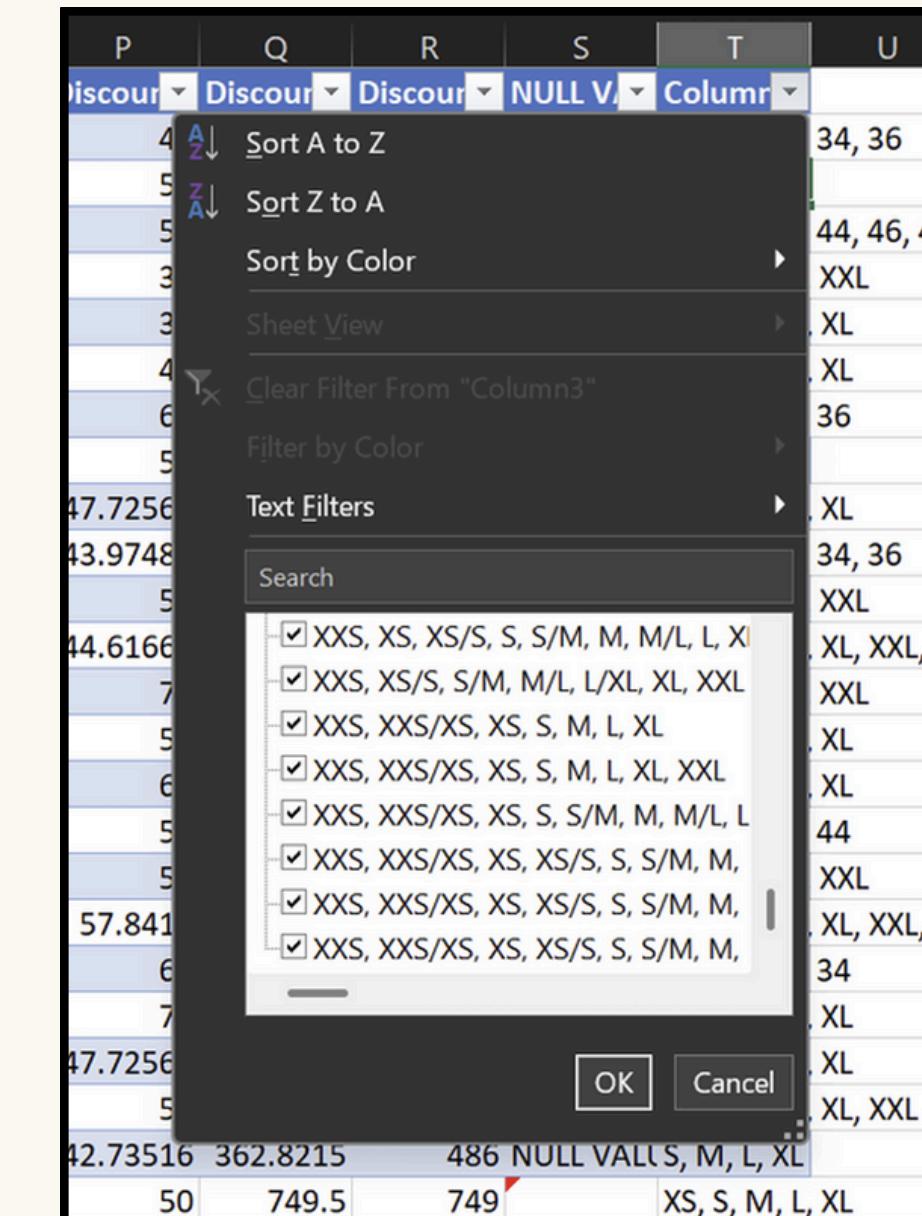
Currently, we are identifying rows containing Null values in both the "Discount Price" and "Discount Offer" columns to be displayed as "NULL VALUES" using the provided formula:

=IF(ISNULL([Discount Price]) AND ISNULL([Discount Offer]), 'NULL VALUES', 'VALID VALUES')

REPLACE ALL NULL VALUES IN THE "SIZEOPTION" COLUMN WITH THE TEXT "NOT AVAILABLE."

We will identify & replace null values in the "SizeOption" column with "NOT AVAILABLE".

I	N	O	P	Q	R	S	T	U
999	45% OFF	674.55	45	674.55	824		=substitute(
999	55% OFF	631.95	55	631.95	517		[@SizeOption],"","N	
999	55% OFF	769.45	55	769.45	629		AVAILABLE")	
999	31% OFF	401.45	31	401.45	893			
999	35% OFF	209.65	35	209.65	389			
999	40% OFF	239.6	40	239.6	359			
998	60% OFF	899.4	60	899.4	599			
998	58% OFF	809.1	58	809.1	585			
997		47.72562	524.0273		573	NULL VALUES		
996		43.97481	1208.868		1540	NULL VALUES		
996	55% OFF	1484.45	55	1484.45	1214			



After applying the formula and checking for "NOT AVAILABLE," it was confirmed that the "SizeOption" column has no null values, leading to the decision to delete the new column.

DATA ANALYSIS

CALCULATE THE OVERALL AVERAGE ORIGINAL PRICE FOR PRODUCTS WITH RATINGS GREATER THAN 4.

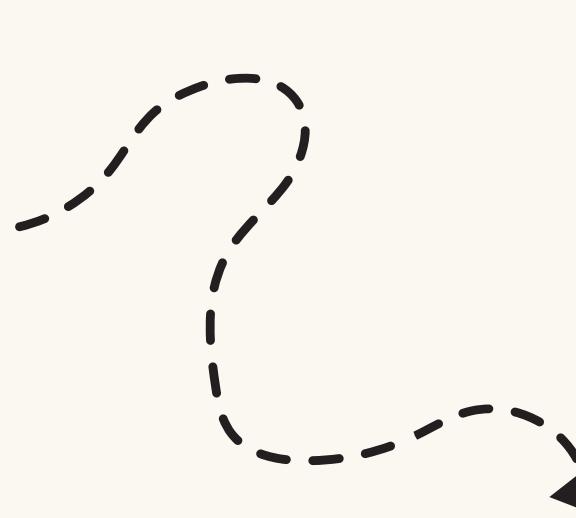
Average of Original Price whose rating is more than 4

```
=AVERAGEIFS(Table2[OriginalPrice (in Rs)],Table2[Ratings],">4")
```

CALCULATE THE OVERALL AVERAGE ORIGINAL PRICE FOR PRODUCTS WITH RATINGS GREATER THAN 4.

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

```
=COUNTIF(Table3[DiscountOffer],">50")
```



COUNT THE NUMBER OF PRODUCTS AVAILABLE IN SIZE "M."

Count the number of products available in size "M"

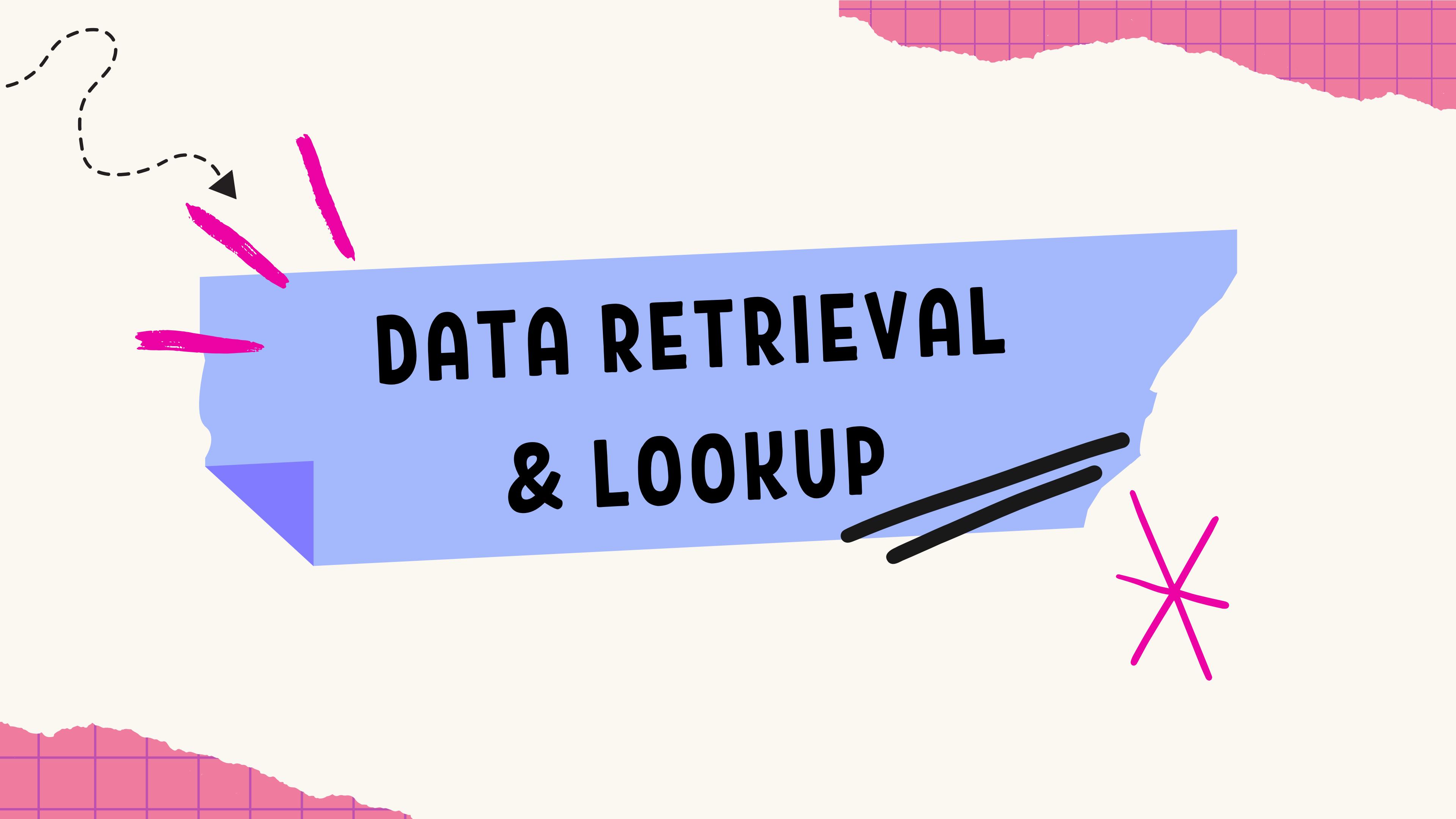
=COUNTIFS(Table1[SizeOption],"*M*")

Count: 6584

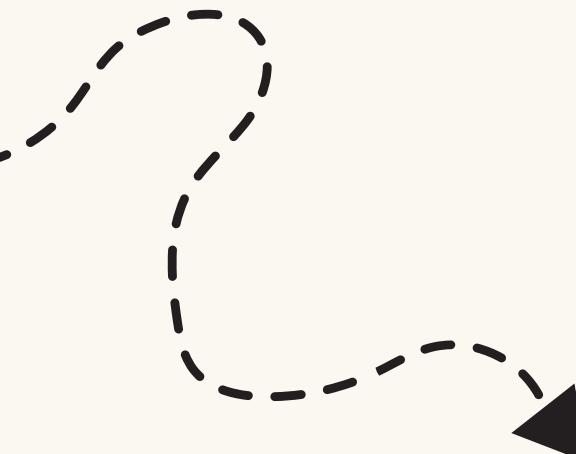
CREATE A NEW COLUMN TO LABEL THE PRODUCTS AS "HIGH DISCOUNT" IF THE DISCOUNT OFFER IS GREATER THAN 50% OFF, OTHERWISE LABEL THEM AS "LOW DISCOUNT."

Create a column labeling offers as "High Discount" for over 50% and "Low Discount" for under 50% using the "If" function.

Amount	Discounted Price	Column3
674.55	824	=IF([@Discount Offer] > 50, "High Discount", "Low Discount")
631.95	517	High Discount
769.45	629	High Discount
401.45	893	Low Discount
209.65	389	Low Discount
239.6	359	Low Discount
899.4	599	High Discount
809.1	585	High Discount
524.027332	573	Average of Original Price whose rating is more than 4
208.867624	1540	Count the number of products with a discount offer greater than 50% OFF
1484.45	1214	Low Discount
11.8705433	387	Count the number of products available in size "M"
2379.3	1019	High Discount
1249.5	1249	Low Discount
479.4	319	High Discount
582.47	516	High Discount



DATA RETRIEVAL & LOOKUP



PROJECT QUESTIONS

- 1) Use VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product_id "11226634".
- 2) Find the "DiscountPrice" for the product with the Product ID "6744434" using the INDEX and MATCH functions.
- 3) Utilize nested xlookup to find any column's detail of a product with it's product id.

USE VLOOKUP/XLOOKUP TO FIND THE PRODUCT BRAND, PRICE, AND RATING OF THE PRODUCT WITH PRODUCT_ID "11226634"

To find the product brand, price, and rating of the product with Product_id “11226634”

- We make use of the following formulas

product brand | =VLOOKUP(S15,Table1[[Product_id]:[High/Low Discount]],2,0)

price | =VLOOKUP(S15,Table1[[Product_id]:[High/Low Discount]],7,0)

rating | =VLOOKUP(S15,Table1[[Product_id]:[High/Low Discount]],12,0)

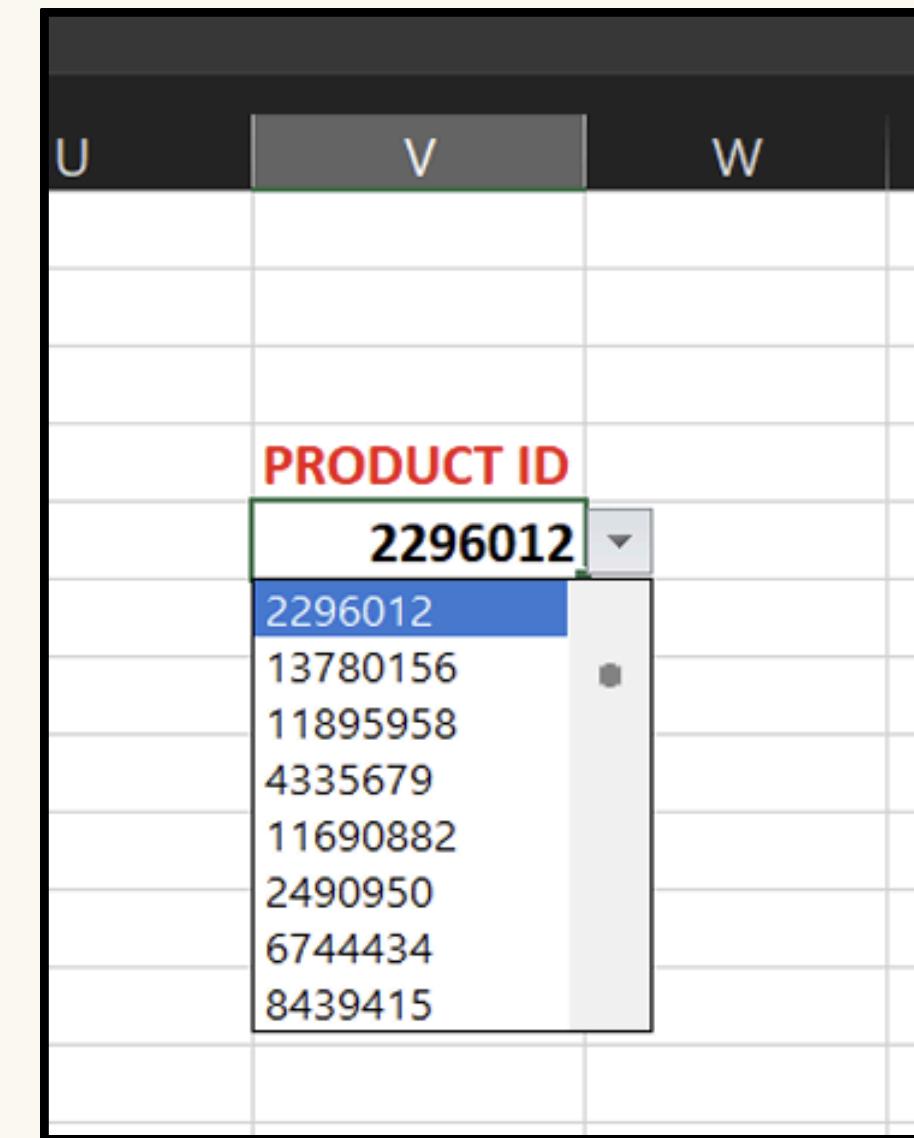
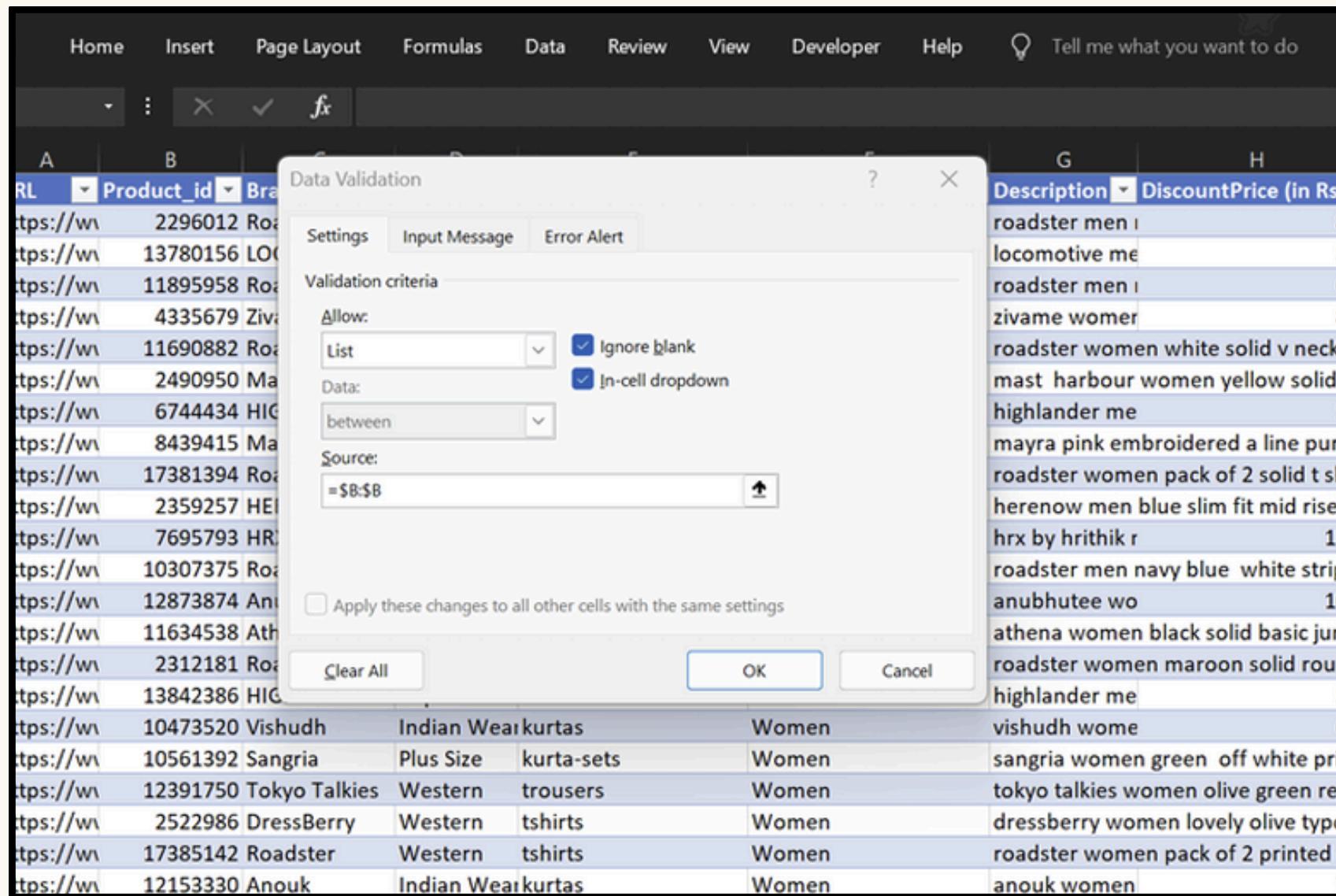
FIND THE "DISCOUNTPRICE" FOR THE PRODUCT WITH THE PRODUCT ID "6744434" USING THE INDEX AND MATCH FUNCTIONS.

To find the "DiscountPrice" for the product with the Product ID "6744434" using the INDEX and MATCH functions

DiscountPrice | =INDEX(Table1[DiscountAmount],MATCH(S25,Table1[Product_id],0))

UTILIZE NESTED XLOOKUP TO FIND ANY COLUMN'S DETAIL OF A PRODUCT WITH IT'S PRODUCT ID.

DATA>DATA VALIDATION> SELECT LIST>FILL SOURCE>CLICK OK



The product list has been curated utilizing data validation techniques.

The team



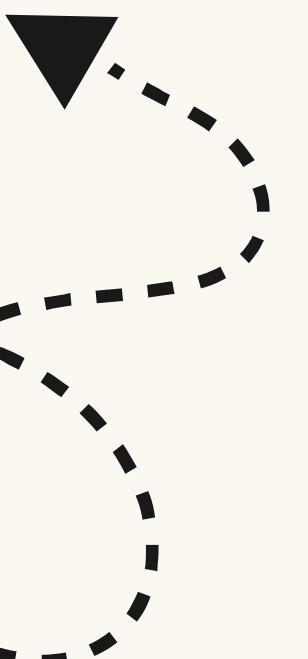
Olivia Wilson
Project manager



Howard Ong
Marketing



Brigitte Schwartz
Project manager



Call To action!

1

As a data enthusiast, I am passionate about turning raw data into actionable insights that drive business growth. My focus lies in identifying trends, solving problems, and making data-backed decisions.

2

As a proactive learner with a keen eye for detail, I am eager to apply my analytical skills to real-world challenges while continuously expanding my knowledge to add value to any organization.

CONTACT ME



8928382552



<https://github.com/jagruti217>



<https://www.linkedin.com/in/jagruti-jadhav-0275192a8>



jagruti.jadhav21@gmail.com

THANK YOU!

