

Analysis of Myntra Apparel

Presented By: Anchal Dayal



Introduction

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.

Problems



Project Questions

A. Data Cleaning and 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 and Lookup

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.

Data Cleaning and Preparation



Check for duplicate values in your dataset and remove them.

The screenshot shows a Microsoft Excel spreadsheet titled "Mynta Fasion Clothing - Excel". The URL in the address bar is <https://www.mynta.com/jeans/roadster/roadster-men-navy-blue-slim-fit-mid-rise-clean-look-jeans/2296012/buy>. The spreadsheet contains data from row 1 to 26, with columns A through M. A "Remove Duplicates" dialog box is open, showing the "My data has headers" checkbox is checked. The columns selected for checking duplicates are A, B, C, D, E, F, G, H, I, J, K, L, M. The dialog box also includes "Select All" and "Unselect All" buttons.

Select the dataset, go to Data > Remove Duplicates, and choose all columns to check for duplicates.

The screenshot shows the same Microsoft Excel spreadsheet after the "Remove Duplicates" operation. The message box "No duplicate values found." is displayed. The data in the spreadsheet remains the same as in the previous screenshot, with rows 1 through 26 and columns A through R.

Standardize "DiscountOffer" column to uniform format.

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

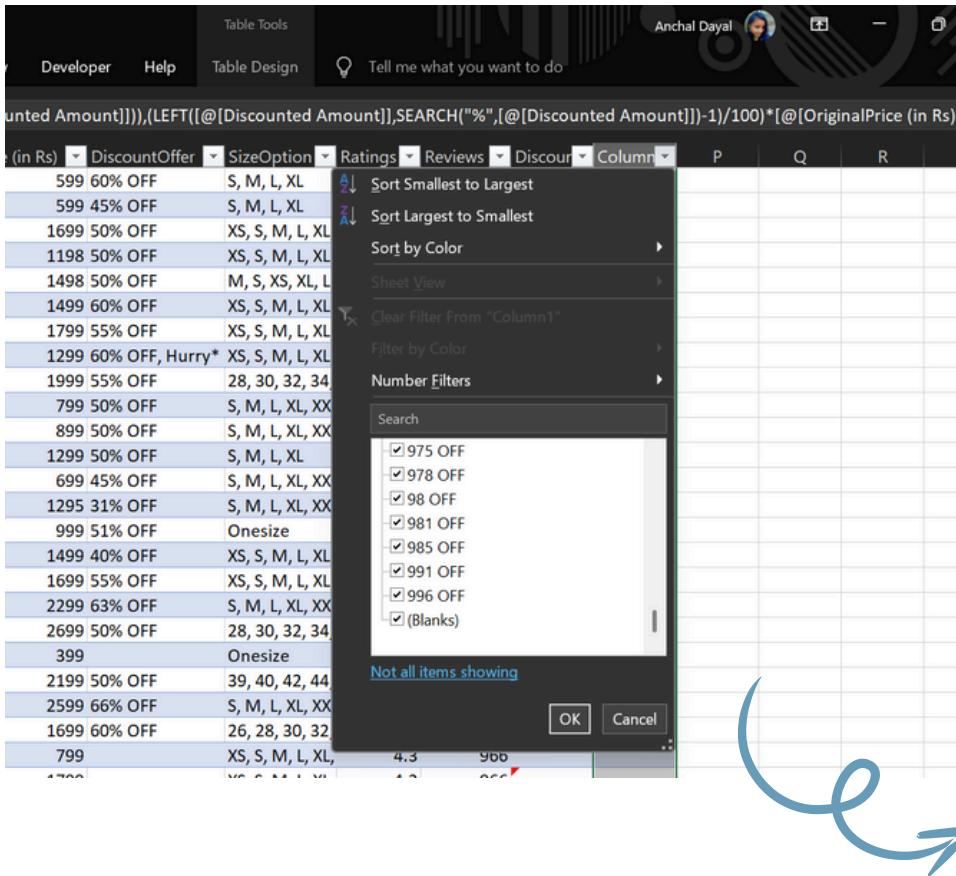
Now, the column will be standardized by converting percentage values into numbers using the formula displayed in the image

I	J	K	L	M	N	O	P	Q
Price (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews				
1499	45% OFF	28, 30, 32, 34,	3.9	999	=Trim(Substitute([@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				
3399	70% OFF	S, M, L, XL, XXL	4.2	996				
2499	50% OFF	XS. S. M. L. XL	4.3	996				

I	J	K	L	M	N	O	P	Q	R	S
OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	DiscountedAmount	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)]],[@[Discounted Amount]])				
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)]],[@[Discounted Amount]])				
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)]],[@[Discounted Amount]])				
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)]],[@[Discounted Amount]])				
599	35% OFF	XS, S, M, L, XL	4.2	999	35% OFF	209.65				
599	40% OFF	XS, S, M, L, XL	4.4	999	40% OFF	239.6				
1499	60% OFF	30, 32, 34, 36	3.9	998	60% OFF	899.4				
1395	58% OFF	S, M, L, XL	3.7	998	58% OFF	809.1				
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	1484.45				
699		XS, S, M, L, XL,	4.1	996						
3399	70% OFF	S, M, L, XL, XXL	4.2	996	70% OFF	2379.3				
2499	50% OFF	XS. S. M. L. XL	4.3	996	50% OFF	1249.5				

"DiscountPrice" fill with category average.

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



Discount Offer	Original Price (in Rs)	Discount %
599 60% OFF	599	60
599 45% OFF	599	45
1699 50% OFF	1699	50
1198 50% OFF	1198	50
1498 50% OFF	1498	50
1499 60% OFF	1499	60
1799 55% OFF	1799	55
1299 60% OFF, Hurry*	1299	60
1999 55% OFF	1999	55
28, 30, 32, 34	28, 30, 32, 34	
799 50% OFF	799	50
899 50% OFF	899	50
1299 50% OFF	1299	50
699 45% OFF	699	45
1295 31% OFF	1295	31
999 51% OFF	999	51
1499 40% OFF	1499	40
1699 55% OFF	1699	55
2299 63% OFF	2299	63
2699 50% OFF	2699	50
399 Onesize	399	
2199 50% OFF	2199	50
2599 66% OFF	2599	66
1699 60% OFF	1699	60
799 XS, S, M, L, XL	799	

Discount Offer	Original Price (in Rs)	Discount %
45% OFF	674.55	45
55% OFF	631.95	55
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
55% OFF	1484.45	55

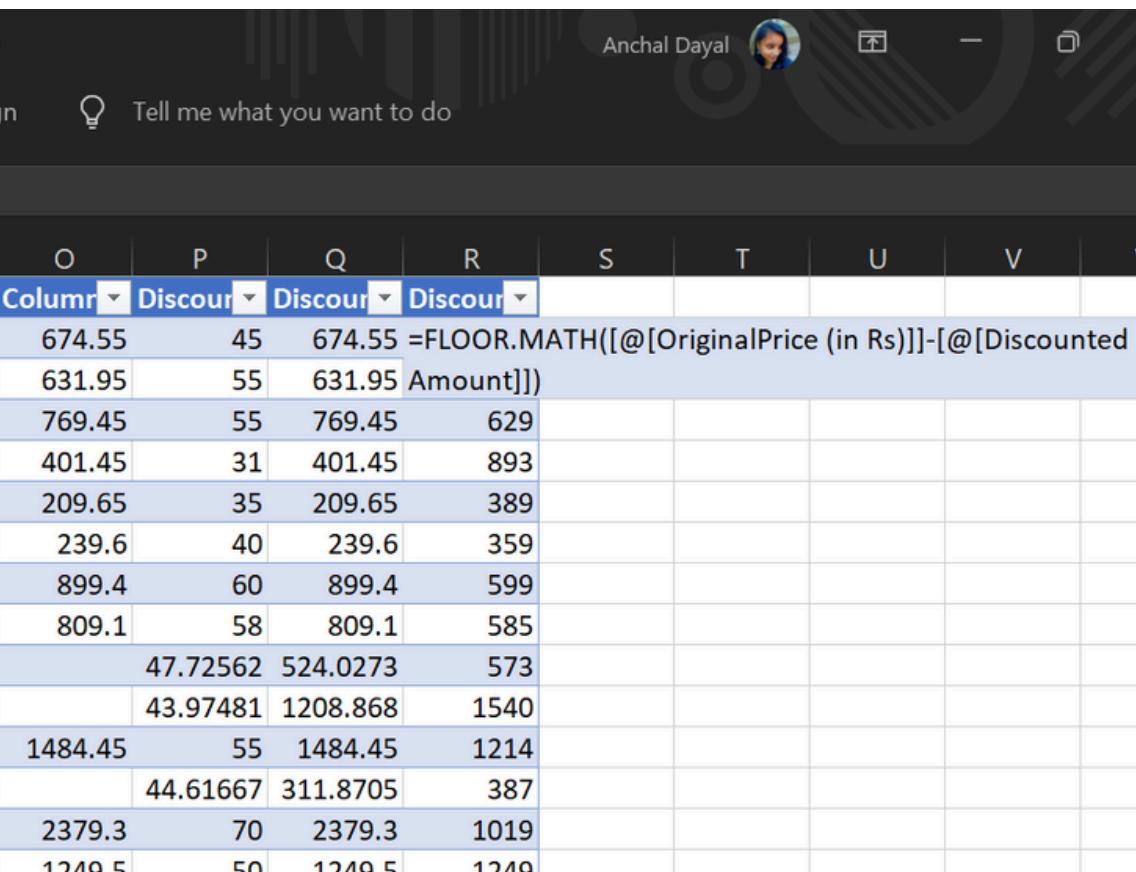
Figuring out the typical “Discount Offer” in %

Convert all numbers to percentages.



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

Table1[@[OriginalPrice (in Rs)]]									
O	P	Q	R	S	T	U	V	W	X
Column1	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			



A screenshot of the Microsoft Power BI desktop application. The ribbon at the top has "Home" selected. The main area shows a table with columns O through W. The formula bar at the bottom displays the following DAX code:

```
Column1 = FLOOR.MATH([@[OriginalPrice (in Rs)]]-[@[Discounted Amount]])
```

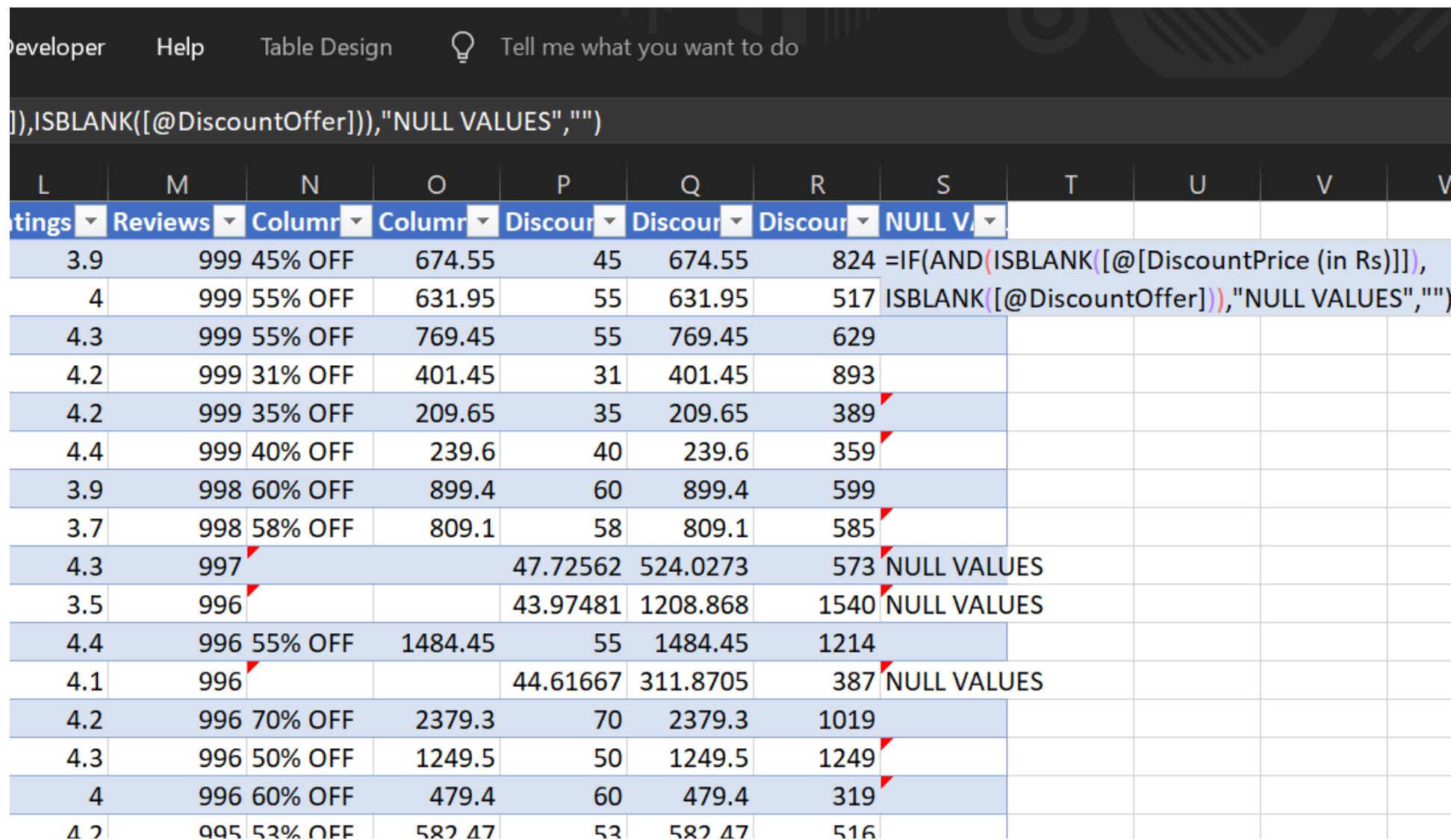
The table data includes rows for various original prices and their corresponding discounted amounts after applying the calculated discount percentage.

O	P	Q	R	S	T	U	V	W
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					
239.6	40	239.6	359					
899.4	60	899.4	599					
809.1	58	809.1	585					
47.72562	524.0273	573						
	43.97481	1208.868	1540					
1484.45	55	1484.45	1214					
	44.61667	311.8705	387					
2379.3	70	2379.3	1019					
1210.5	50	1210.5	1210					

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
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					
4.2	995	52% OFF	582.47	52	582.47	516					

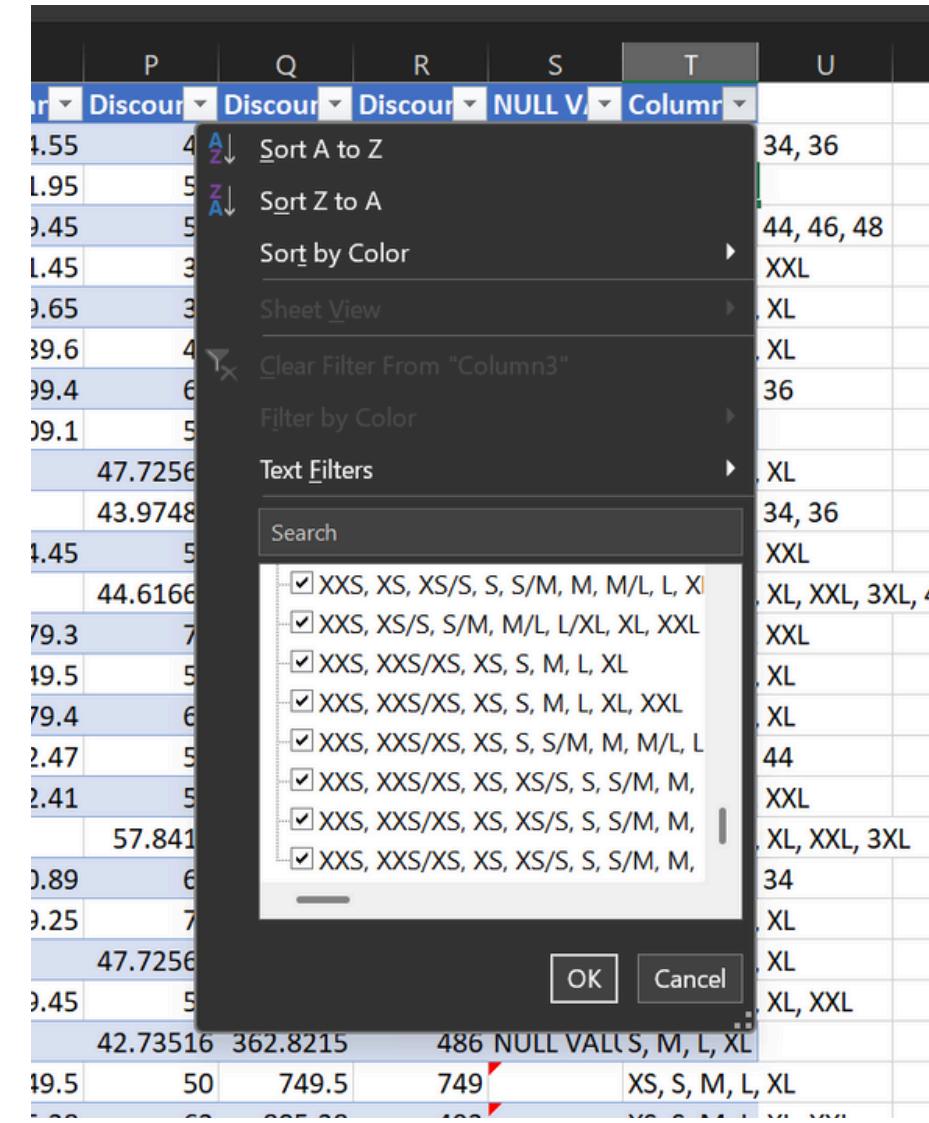
Currently, we are identifying rows containing Null values in both the "Discount Price" and "Discount Offer" columns to be displayed as "NULL VALUES" using the provided formula:

=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."

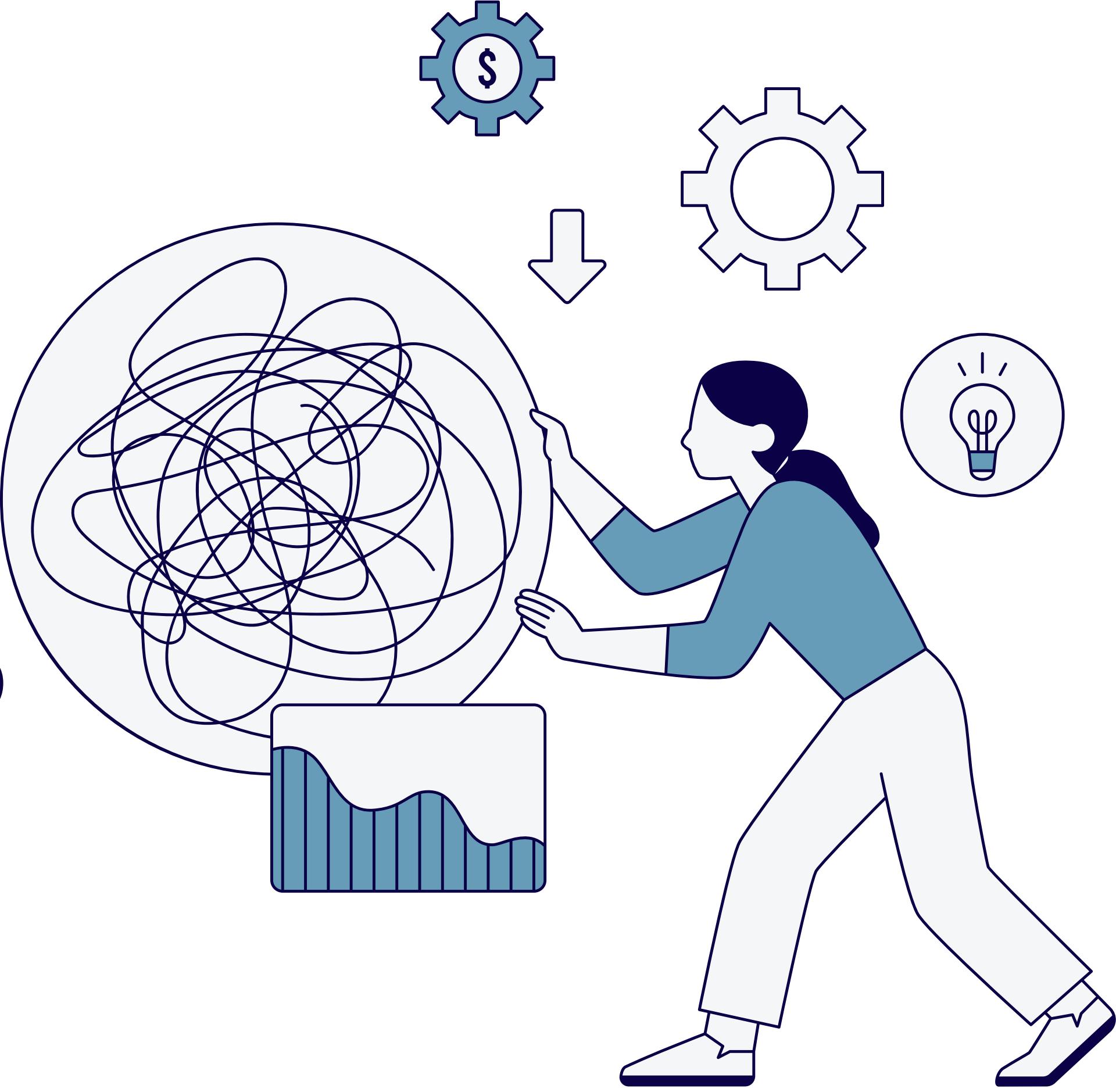
We will identify and replace null values in the "SizeOption" column with "NOT AVAILABLE".

M	N	O	P	Q	R	S	T	U
Reviews	Column	Column	Discour	Discour	Discour	NULL V		
999	45% OFF	674.55	45	674.55	824		=substitute(
999	55% OFF	631.95	55	631.95	517		[@SizeOption],"","NOT	
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			



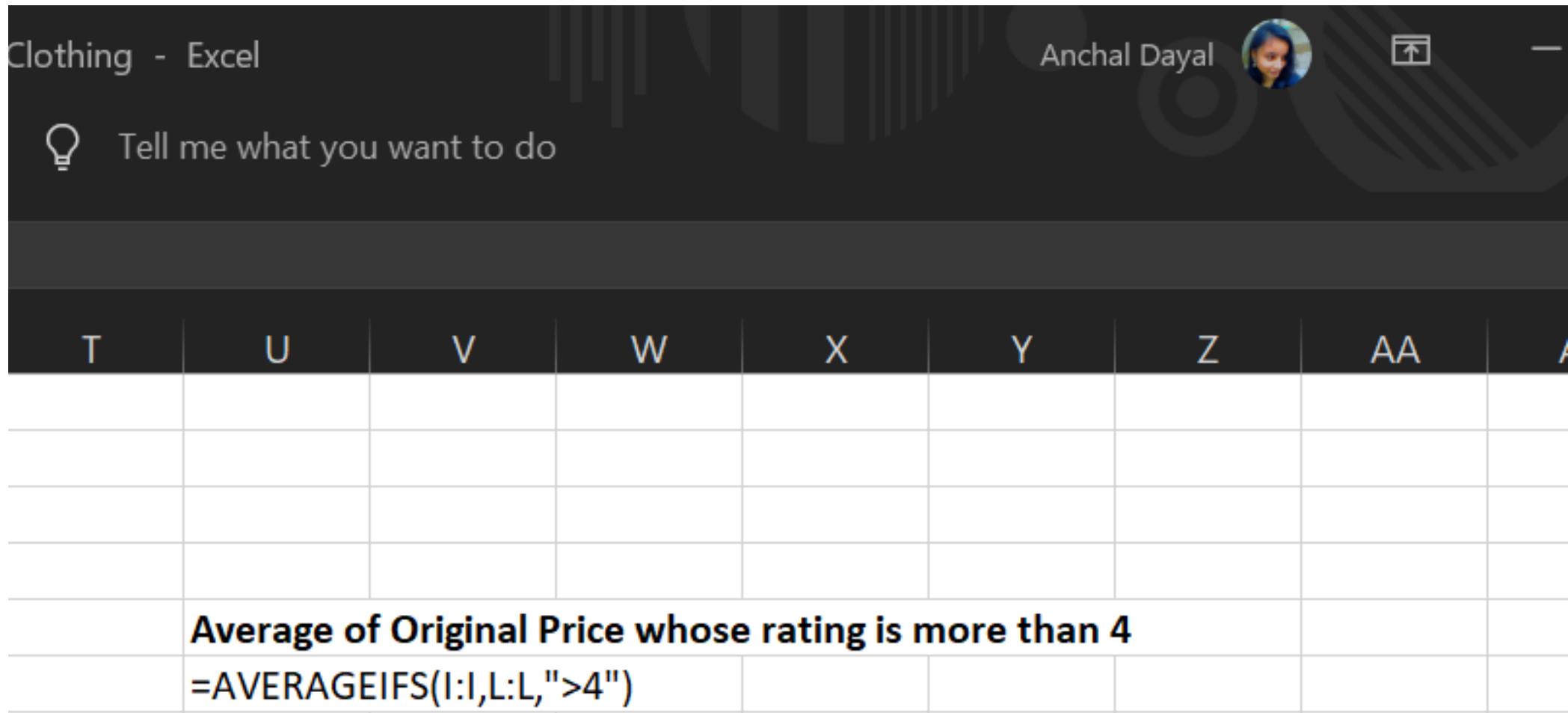
After applying the formula and checking for "NOT AVAILABLE," it was confirmed that the "SizeOption" column has no null values, leading to the decision to delete the new column.

Data Analysis





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



Clothing - Excel

Anchal Dayal

T U V W X Y Z AA A

Average of Original Price whose rating is more than 4
=AVERAGEIFS(I:I,L:L,>4")

Calculate the average original price for ratings above 4 using the formula =Averageifs(I:I,L:L,>4").

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

We will utilize the COUNTIFS function to identify products with discount offers exceeding 50%.

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



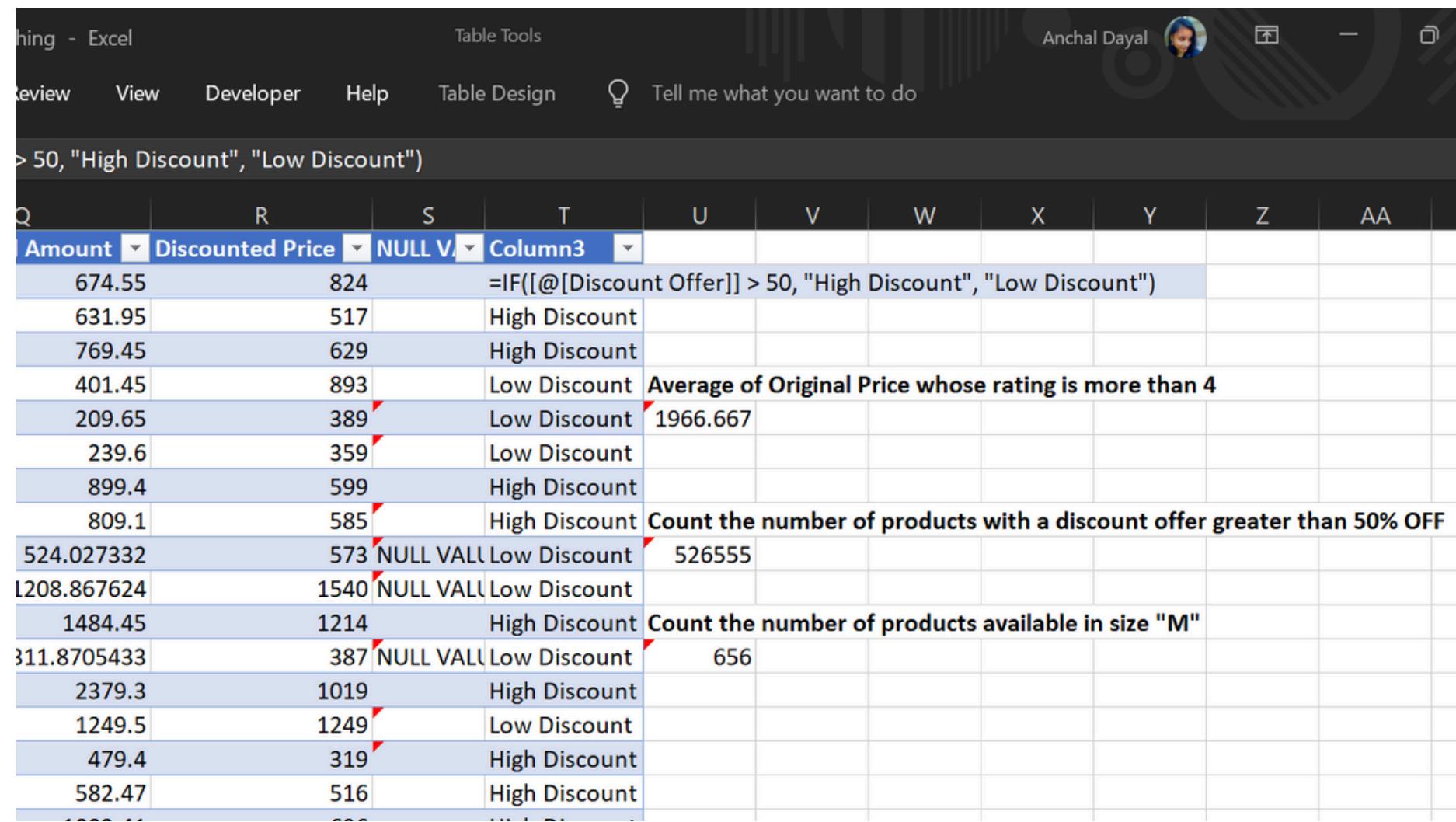
Count the number of products available in size "M."

Here we are using the COUNTIF function to count the number of products with the condition where the size is "M".

Count the number of products available in size "M"
=COUNTIF(K:K, "M")



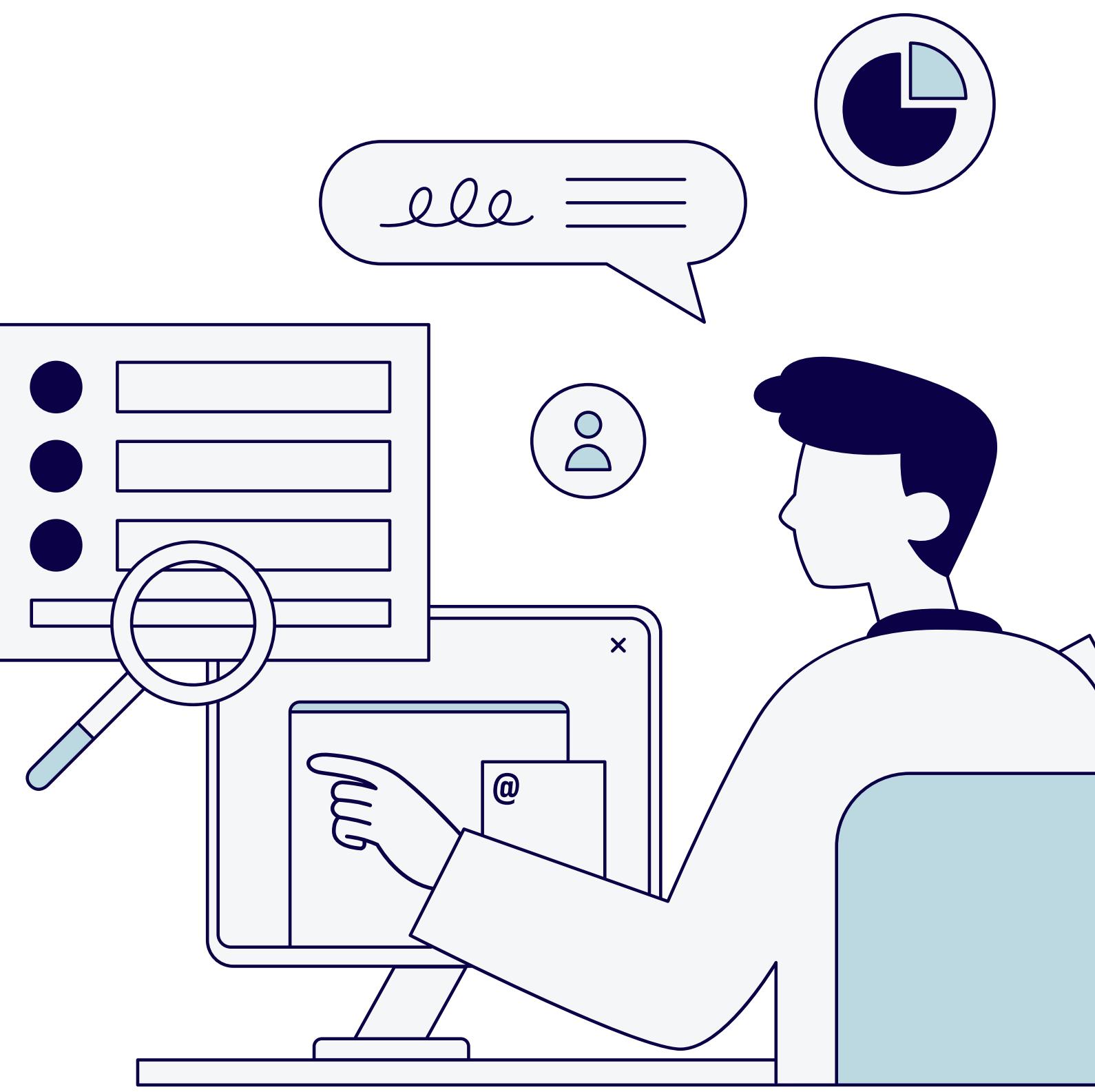
Create a new column to label the products as "High Discount" if the discount offer is greater than 50% OFF, otherwise label them as "Low Discount."



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	Count the number of products with a discount offer greater than 50% OFF
1208.867624	1540	NULL VALL Low Discount
1484.45	1214	Count the number of products available in size "M"
311.8705433	387	NULL VALL Low Discount
2379.3	1019	High Discount
1249.5	1249	Low Discount
479.4	319	High Discount
582.47	516	High Discount

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

Data Retrieval and Lookup





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

VLOOKUP	0	11226634
BRAND		
Maniac		
PRICE		
Sports Wear		
RATING		
tshirts		

BRAND
`=VLOOKUP(X15,B:C, 2, FALSE)`

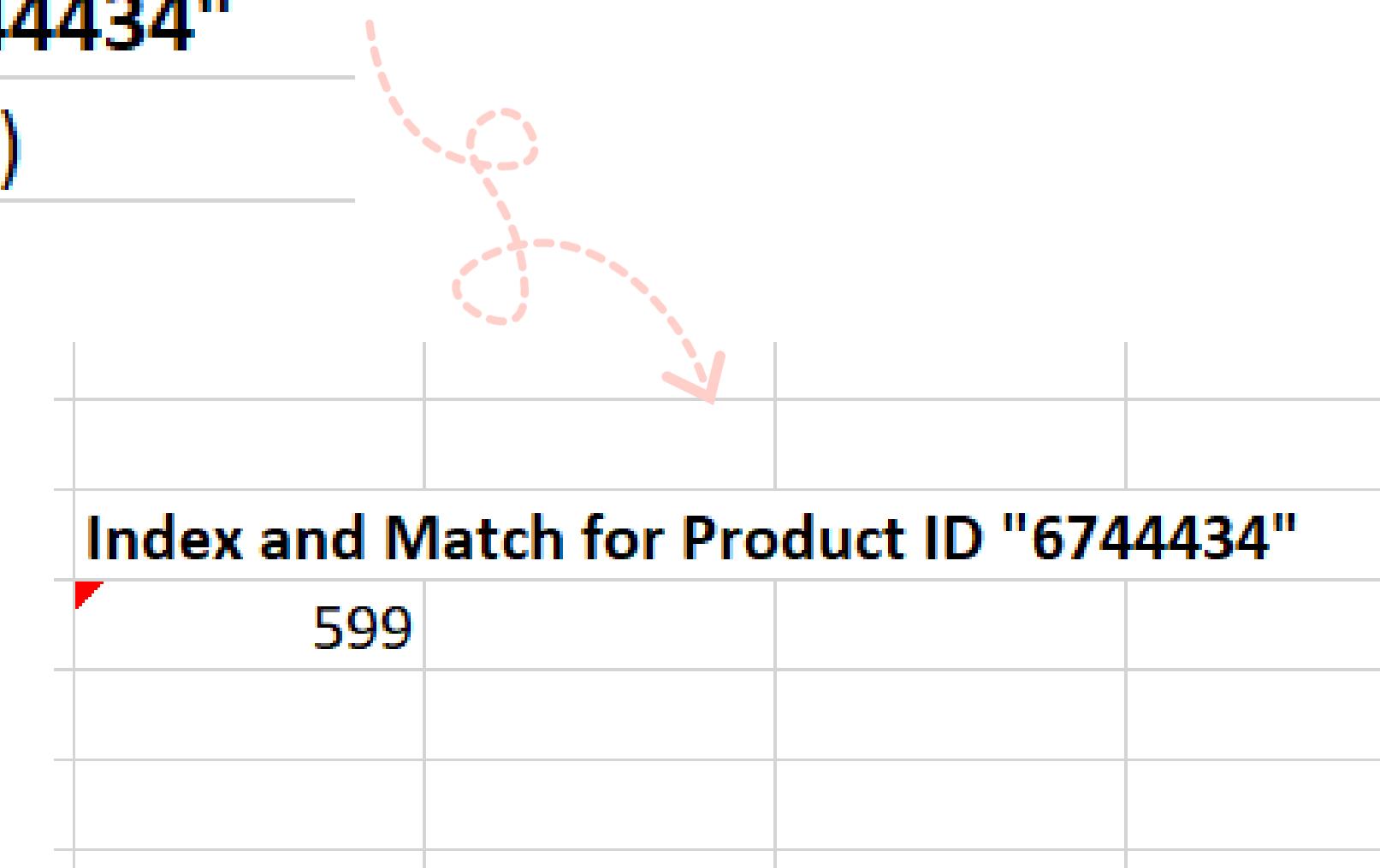
PRICE
`=VLOOKUP(X15,B:D, 3, FALSE)`

RATING
`=VLOOKUP(X15,B:E, 4, FALSE)`

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

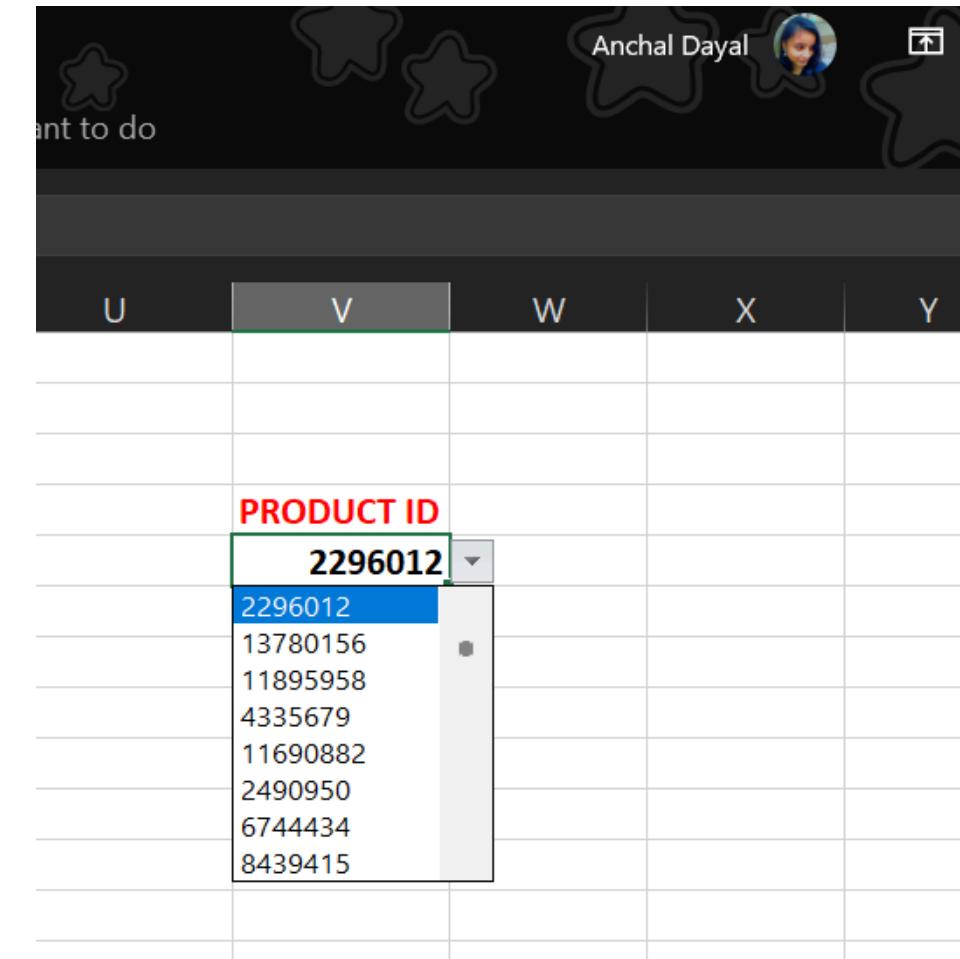
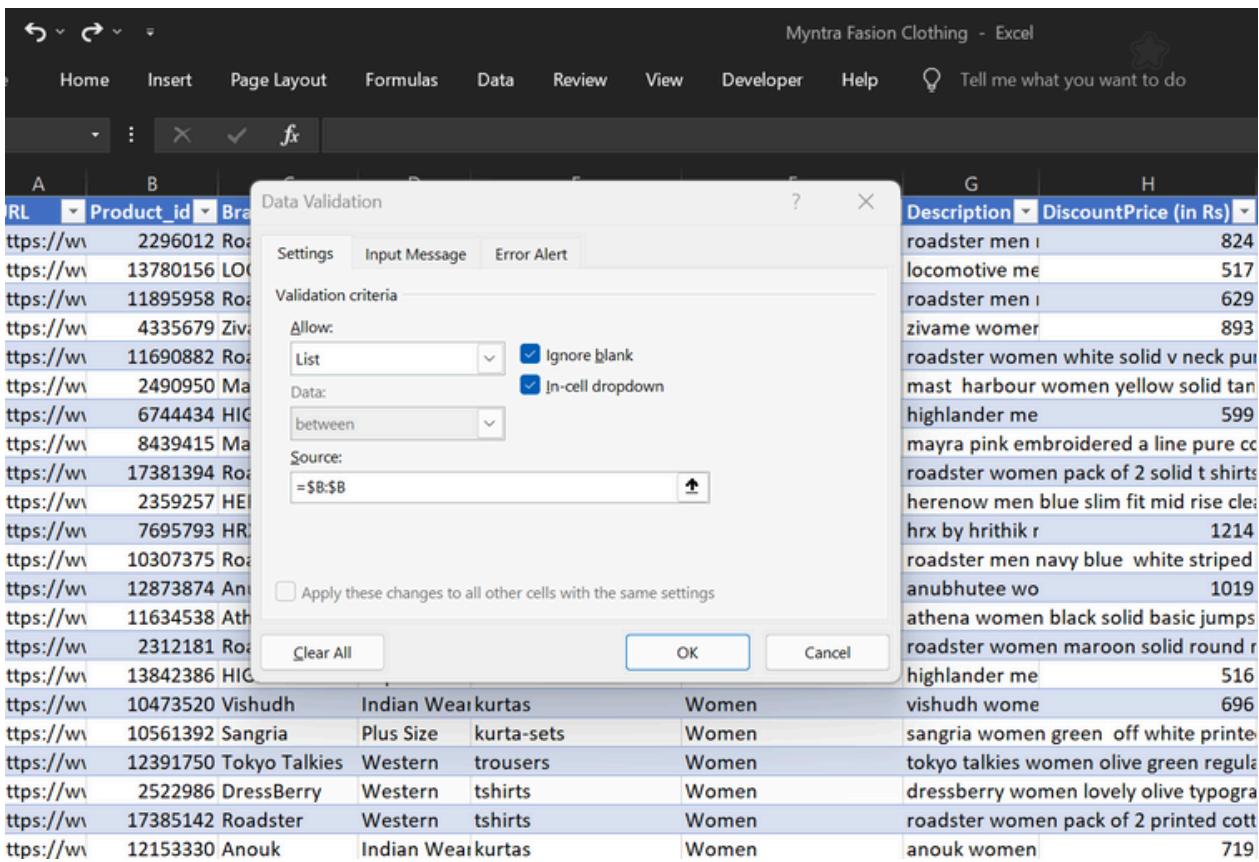
Index and Match for Product ID "6744434"

=INDEX(R:R, MATCH(6744434, B:B, 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.



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

Mynta Fasion Clothing - Excel

Review View Developer Help Tell me what you want to do

B:B, 0), 3)

Q	R	S	T	U	V	W	X	Y
Discounted Amount	Discounted Price	NULL V	Column3					
674.55	824		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	NULL V	All Low Discount					
1208.867624	1540	NULL V	All Low Discount					
1484.45	1214		High Discount					
311.8705433	387	NULL V	All Low Discount					
2379.3	1019		High Discount					
1249.5	1249		Low Discount					
479.4	319		High Discount					
582.47	516		High Discount					
1002.41	696		High Discount					

I have utilized the Index and Match functions since I cannot access the 2021 version of Excel, which would allow me to use the XLOOKUP function.

Thank You

