WSCUBE TECH®
Upskilling Bharat

Analysis of Myntra Apparel

Presented by
Raviraj Singh Chundawat

Company Overview

Myntra is India's leading fashion and lifestyle e-commerce marketplace, providing a comprehensive portfolio of apparel, footwear, accessories, and beauty products. Founded in 2007 and headquartered in Bengaluru, the company partners with over 6,000 brands and operates a nationwide delivery network covering more than 19,000 pin codes. With strong digital infrastructure and customer-focused policies, Myntra leverages data-driven insights to optimize operations, enhance customer experience, and drive sustainable business growth.

Objective

Having a clean and complete dataset to perform analysis and get meaningful insights.

Problems

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."

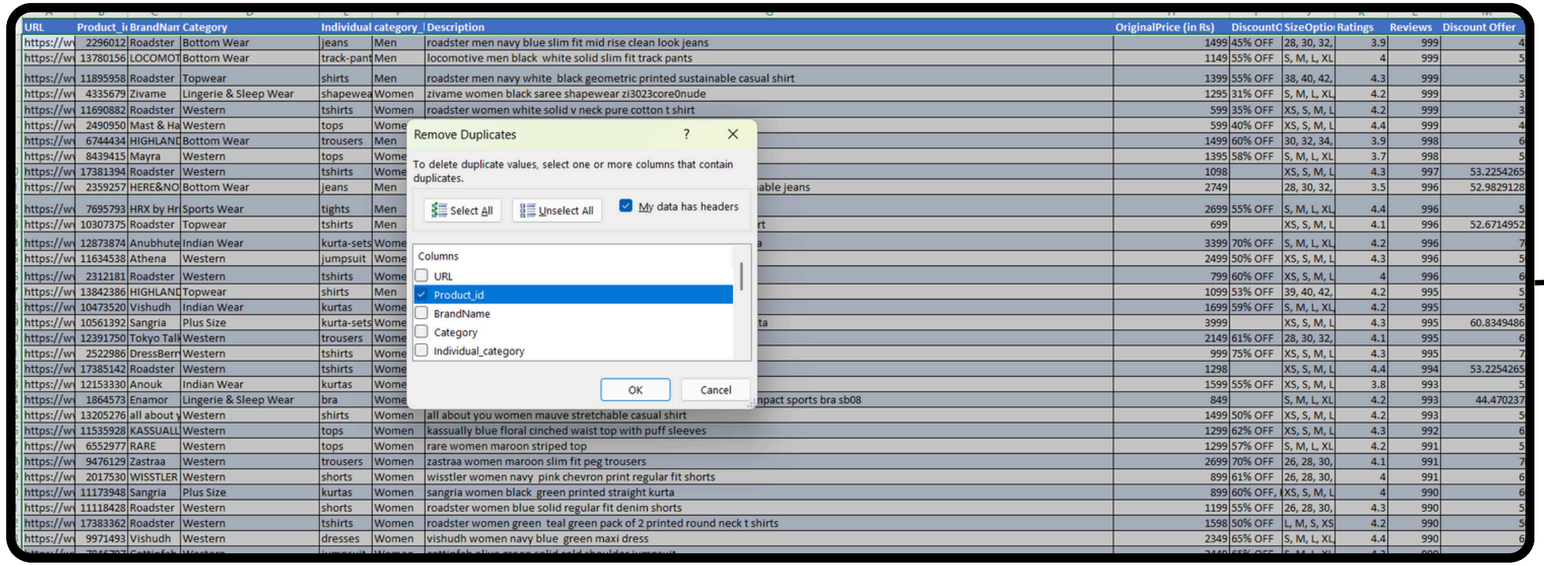
C. Data Retrieval and Lookup

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.

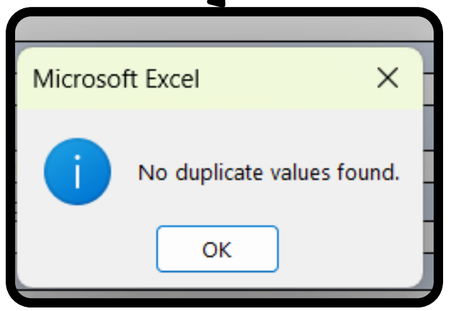
Data Cleaning and Preparation

WSCUBE TECH®
Upskilling Bharat

Check for duplicate values in your dataset and remove them.

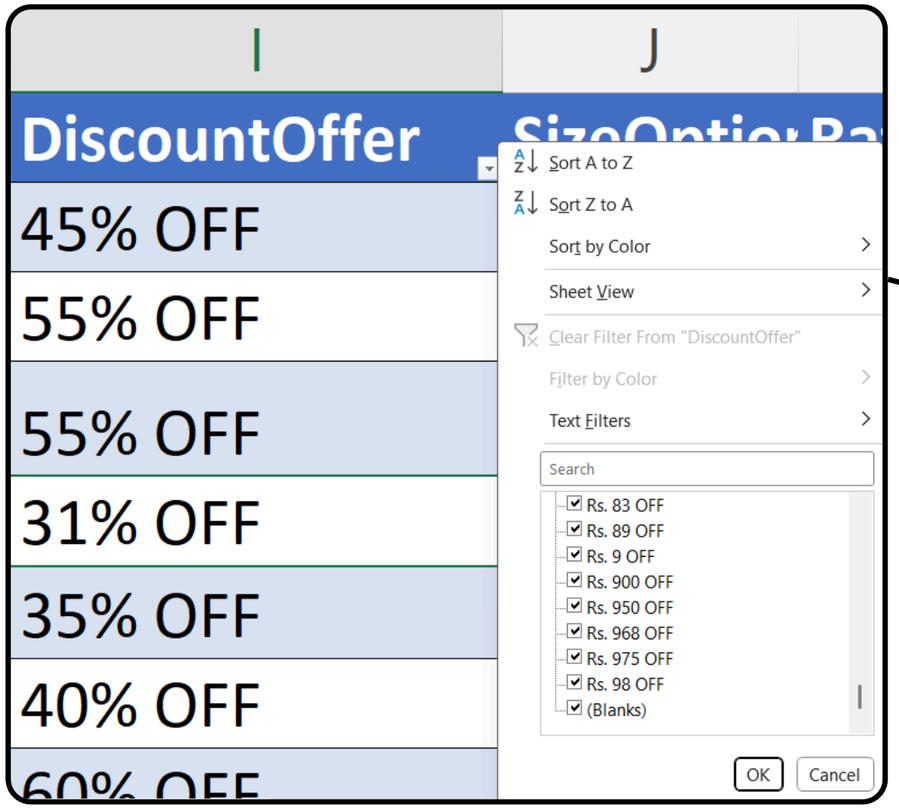


To Delete Duplicate value select entire Data, **Go to Data Tab ---> Remove Duplicate**, Click on Unselect all and then **select** Product_id and **Click Ok**



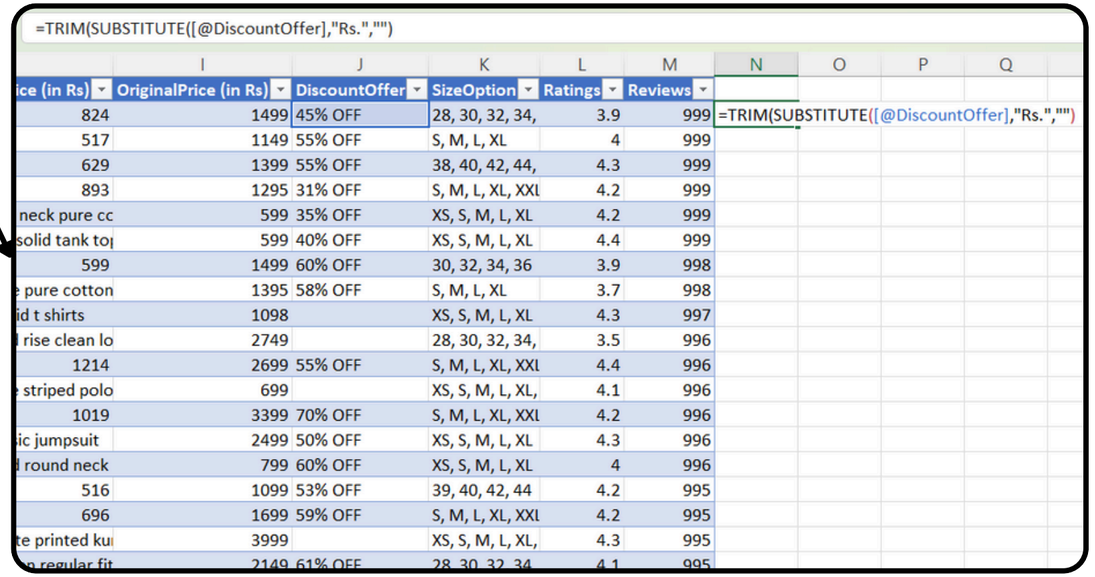
No Duplicate Values Present in our Data

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



Removing "Rs." using **SUBSTITUTE Function** and removing Whitespaces using **TRIM** function in a new coulumn

=TRIM(SUBSTITUTE([@DiscountOffer],"Rs.",""))



=TRIM(SUBSTITUTE([@DiscountOffer],"Rs.",""))

ice (in Rs)	OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	=TRIM(SUBSTITUTE([@DiscountOffer],"Rs.",""))
824	1499	45% OFF	28, 30, 32, 34,	3.9	999	
517	1149	55% OFF	S, M, L, XL	4	999	
629	1399	55% OFF	38, 40, 42, 44,	4.3	999	
893	1295	31% OFF	S, M, L, XL, XXL	4.2	999	
neck pure cc	599	35% OFF	XS, S, M, L, XL	4.2	999	
solid tank to	599	40% OFF	XS, S, M, L, XL	4.4	999	
599	1499	60% OFF	30, 32, 34, 36	3.9	998	
pure cotton	1395	58% OFF	S, M, L, XL	3.7	998	
id t shirts	1098		XS, S, M, L, XL	4.3	997	
rise clean lo	2749		28, 30, 32, 34,	3.5	996	
1214	2699	55% OFF	S, M, L, XL, XXL	4.4	996	
striped polo	699		XS, S, M, L, XL,	4.1	996	
1019	3399	70% OFF	S, M, L, XL, XXL	4.2	996	
ic jumpsuit	2499	50% OFF	XS, S, M, L, XL	4.3	996	
round neck	799	60% OFF	XS, S, M, L, XL	4	996	
516	1099	53% OFF	39, 40, 42, 44	4.2	995	
696	1699	59% OFF	S, M, L, XL, XXL	4.2	995	
te printed ku	3999		XS, S, M, L, XL,	4.3	995	
n regular fit	2149	61% OFF	28, 30, 32, 34	4.1	995	

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

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

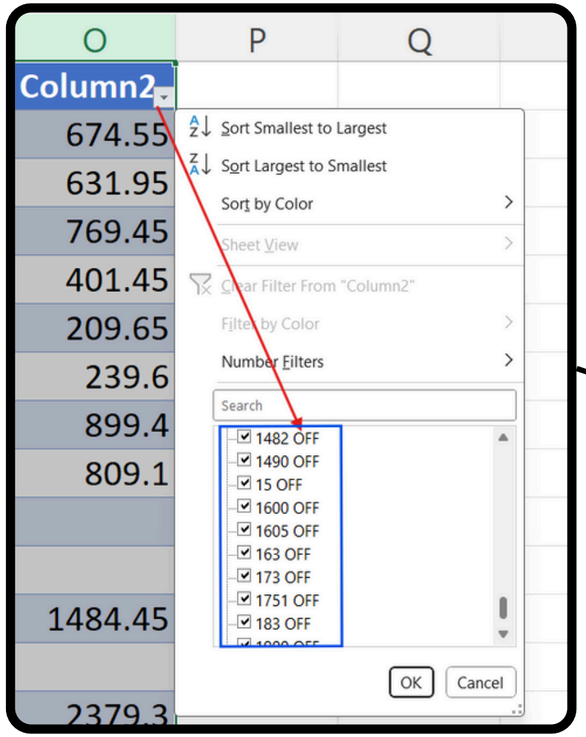
```
=IF(ISNUMBER(SEARCH("%",[@DiscountOffer2])),(LEFT([@DiscountOffer2],SEARCH("%",
[@DiscountOffer2])-1)/100)*[@[OriginalPrice (in Rs)]],[@DiscountOffer2])
```

Get & Transform Data | Queries & Connections | Sort & Filter | Data Tools | Forecast | Outline

=IF(ISNUMBER(SEARCH("%",[@Column1])),(LEFT([@Column1],SEARCH("%",[@Column1])-1)/100)*[@[OriginalPrice (in Rs)]],[@Column1])

	I	J	K	L	M	N	O	P	Q	R	S	T
1	OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Column1	Column2					
2	1499	45% OFF	28, 30, 32, 34	3.9	999	45% OFF	=IF(ISNUMBER(SEARCH("%",[@Column1])),(LEFT([@Column1],					
3	1149	55% OFF	S, M, L, XL	4	999	55% OFF	SEARCH("%",[@Column1])-1)/100)*[@[OriginalPrice (in Rs)]],					
4	1399	55% OFF	38, 40, 42, 44	4.3	999	55% OFF	[@Column1])					
5	1295	31% OFF	S, M, L, XL, XX	4.2	999	31% OFF	401.45					
6	599	35% OFF	XS, S, M, L, XL	4.2	999	35% OFF	209.65					
7	599	40% OFF	XS, S, M, L, XL	4.4	999	40% OFF	239.6					
8	1499	60% OFF	30, 32, 34, 36	3.9	998	60% OFF	899.4					
9	1395	58% OFF	S, M, L, XL	3.7	998	58% OFF	809.1					
10	1098		XS, S, M, L, XL	4.3	997							
11	2749		28, 30, 32, 34	3.5	996							
12	2699	55% OFF	S, M, L, XL, XX	4.4	996	55% OFF	1484.45					
13	699		XS, S, M, L, XL	4.1	996							
14	3399	70% OFF	S, M, L, XL, XX	4.2	996	70% OFF	2379.3					
15	2499	50% OFF	XS, S, M, L, XL	4.3	996	50% OFF	1249.5					
16	799	60% OFF	XS, S, M, L, XL	4	996	60% OFF	479.4					
17	1099	53% OFF	39, 40, 42, 44	4.2	995	53% OFF	582.47					

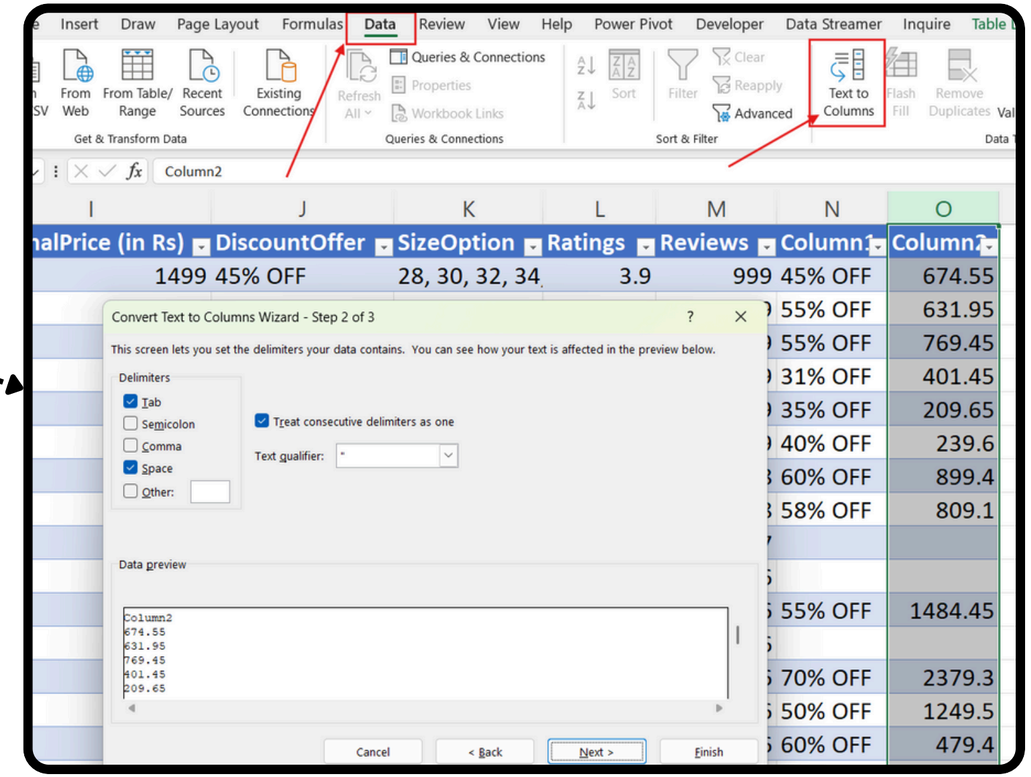
Standardize the "DiscountOffer" column to a single format, ensuring all values are uniform



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

First select the column copy and paste as values at same place.

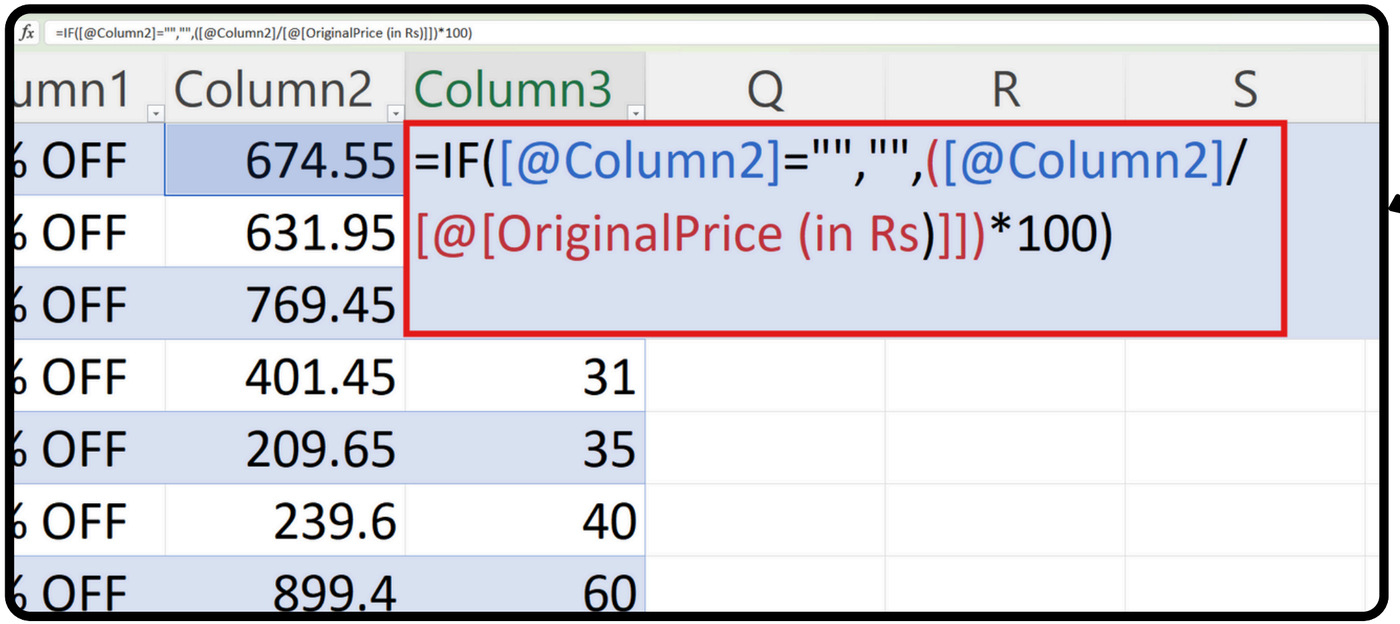
Then select the entire column, go to **Data tab** ------> **Text to columns**, check **Delimited**, click next then click on **Space**, then click **Finish**



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

Calculating percentage again using below formula

=IF([@Column2]="","",([@Column2]/[@[OriginalPrice (in Rs)]])*100)



Filling Null values with Average of respected Category by using Presented Formula

```
=IF([@[Discount %]]="",FLOOR.MATH(AVERAGEIFS([Discount %],[Category],
[@Category])),[@[Discount %]])
```

=IF([@[Discount %]]="",FLOOR.MATH(AVERAGEIFS([Discount %],[Category],[@Category])),[@[Discount %]])

I	J	K	L	M	N	O	P	Q	R	S
OriginalPrice (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Column1	Column2	Discount %	Column4		
1499	45% OFF	28, 30, 32, 34	3.9	999	45% OFF	674.55	45	=IF([@[Discount %]]="", FLOOR.MATH(AVERAGEIFS([Discount %],[Category],[@Category])),[@[Discount %]])		
1149	55% OFF	S, M, L, XL	4	999	55% OFF	631.95	55			
1399	55% OFF	38, 40, 42, 44	4.3	999	55% OFF	769.45	55			
1295	31% OFF	S, M, L, XL, XX	4.2	999	31% OFF	401.45	31			
599	35% OFF	XS, S, M, L, XL	4.2	999	35% OFF	209.65	35			
599	40% OFF	XS, S, M, L, XL	4.4	999	40% OFF	239.6	40	40		
1499	60% OFF	30, 32, 34, 36	3.9	998	60% OFF	899.4	60	60		
1395	58% OFF	S, M, L, XL	3.7	998	58% OFF	809.1	58	58		
1098		XS, S, M, L, XL	4.3	997				53		
2749		28, 30, 32, 34	3.5	996				52		
2699	55% OFF	S, M, L, XL, XX	4.4	996	55% OFF	1484.45	55	55		
699		XS, S, M, L, XL	4.1	996				52		
3399	70% OFF	S, M, L, XL, XX	4.2	996	70% OFF	2379.3	70	70		
2499	50% OFF	XS, S, M, L, XL	4.3	996	50% OFF	1249.5	50	50		
799	60% OFF	XS, S, M, L, XL	4	996	60% OFF	479.4	60	60		

Calculating Discount Amount and Discounted Price

Calculating Discount Amount using below foemula

=(([@[Discount %]]*[@[OriginalPrice (in Rs)]])/100)

=(([@[Discount %]]*[@[OriginalPrice (in Rs)]])/100)

	Q	R	T	U	V
	Discount %	Discount Amount			
45	45	=(([@[Discount %]]*[@[OriginalPrice (in Rs)]])/100)			
55	55	631.95			
55	55	769.45			
31	31	401.45			
35	35	209.65			
40	40	239.6			
60	60	899.4			
58	58	809.1			
	53	581.94			
	52	1429.48			
55	55	1484.45			
	52	363.48			
70	70	2379.3			
50	50	1249.5			
60	60	479.4			
53	53	582.47			

Calculating Discount Amount and Discounted Price

Calculating Discounted Price by usiing Floor.Math formula presented below

=FLOOR.MATH([@[OriginalPrice (in Rs)]]-[@[Discount Amount]])

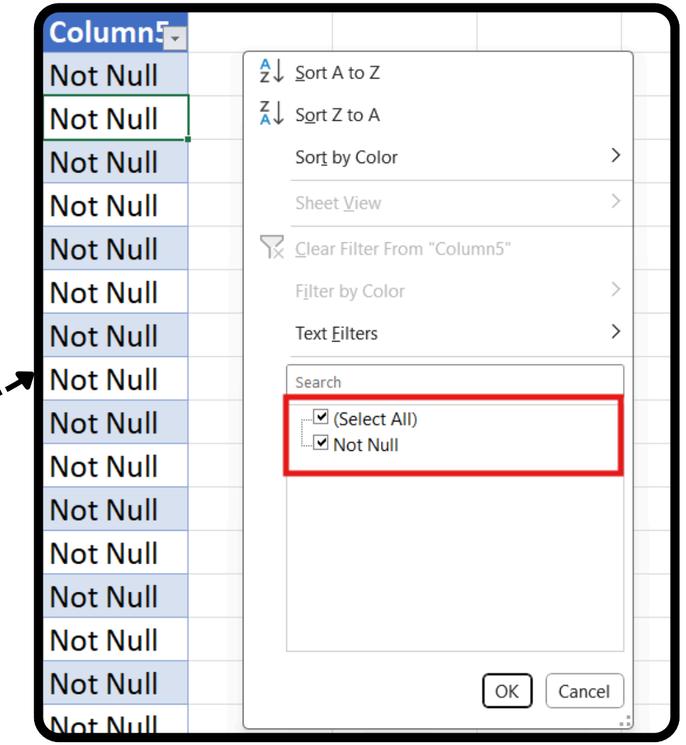
R	S	T	U	V	W	X
Discount Amount	Column					
674.55	=FLOOR.MATH([@[OriginalPrice (in Rs)]]-[@[Discount Amount]])					
631.95	517					
769.45	629					
401.45	893					
209.65	389					
239.6	359					
899.4	599					
809.1	585					
581.94	516					
1429.48	1319					
1484.45	1214					
363.48	335					
2379.3	1019					
1249.5	1249					
479.4	319					
582.47	516					

Calculating Discount Amount and Discounted Price

Checking for null values in column “Size Option” using “ISBLANK” Function presented in image

Discounted	Column5
824	=IF(ISBLANK([@SizeOption]),"Null","Not Null")
517	Not Null
629	Not Null
893	Not Null
389	Not Null
359	Not Null
599	Not Null
585	Not Null
516	Not Null
1319	Not Null
1214	Not Null
335	Not Null
1019	Not Null

No null values present in Size Option.



Data Analysis

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

Calculating the overall average original price for products with ratings greater than 4 using “AVERAGEIFS” Function.

=AVERAGEIFS(Table1[OriginalPrice (in Rs)],Table1[Ratings],">4")

=AVERAGEIFS(Table1[OriginalPrice (in Rs)],Table1[Ratings],">4")

H	I	J	K
ntPrice (in Rs)	OriginalPrice (in Rs)	SizeOptic	Ratings
824	1499	28, 30, 32,	3.9
517	1149	S, M, L, XL	4
629	1399	38, 40, 42,	4.3
893	1295	S, M, L, XL,	4.2
e solid v neck pu	599	XS, S, M, L,	4.2
n yellow solid tar	599	XS, S, M, L,	4.4
599	1499	30, 32, 34,	3.9
red a line pure c	1395	S, M, L, XL	3.7
of 2 solid t shirt	1098	XS, S, M, L,	4.3

products with ratings greater than 4.

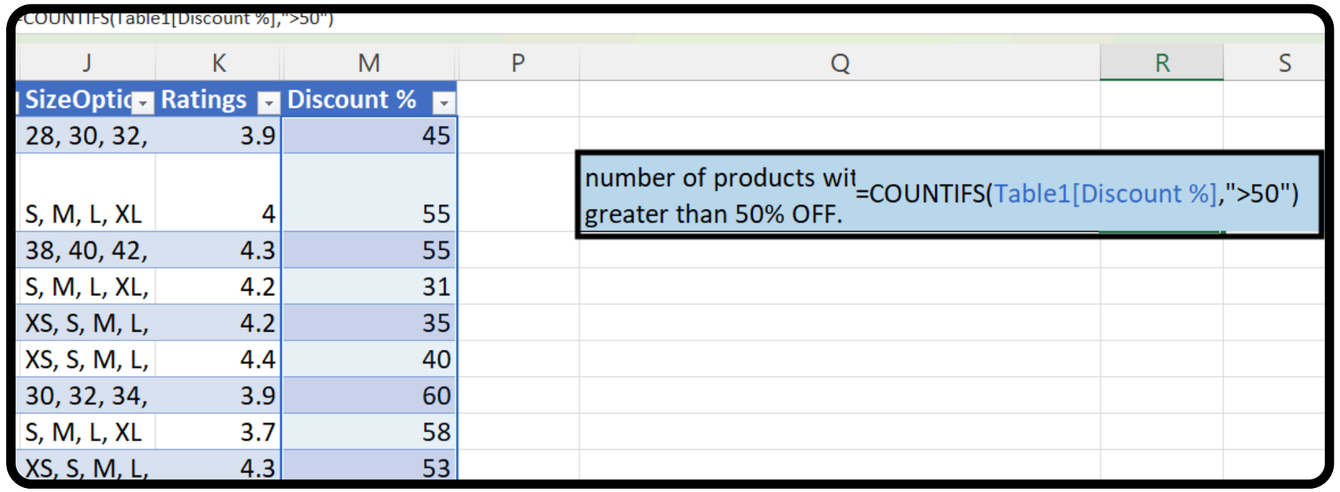
=AVERAGEIFS(Table1[OriginalPrice (in Rs)], Table1[Ratings],">4")

Overall average original price for products with ratings greater than 4.	1653.59

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

Calculating the number of products with a discount offer greater than 50% OFF Using “COUNTIFS” Function.

=COUNTIFS(Table1[Discount %],">50")



number of products with a discount offer greater than 50% OFF.	6002

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

Calculating the number of products available in size "M. Using **COUNTIFS** Function and using **Wildcard "*"**.

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

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

OriginalPrice (in Rs)	SizeOptio	Ratings	Discount %
1499	28, 30, 32,	3.9	45
1149	S, M, L, XL	4	55
1399	38, 40, 42,	4.3	55
1295	S, M, L, XL,	4.2	31
599	XS, S, M, L,	4.2	35
599	XS, S, M, L,	4.4	40
1499	30, 32, 34,	3.9	60
1395	S, M, L, XL	3.7	58
1098	XS, S, M, L,	4.3	53
2749	28, 30, 32,	3.5	52

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

number of products available in size "M."	6585

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."

Labeling products Using **"IF"** function given Below. and then applied **conditional formating** for highlighting cells.

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

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

	N	O	P	Q	R
t %	Discount Amount	Discounted Price	Discount Distribution		
45	674.55	824	=IF([@[Discount %]]>50,"High Discount","Low Discount")		
55	631.95	517	High Discount		
55	769.45	629	High Discount		
31	401.45	893	Low Discount		
35	209.65	389	Low Discount		
40	239.6	359	Low Discount		
60	899.4	599	High Discount		
58	809.1	585	High Discount		
53	581.94	516	High Discount		
52	1429.48	1319	High Discount		
55	1484.45	1214	High Discount		
52	363.48	335	High Discount		

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

Calculating the number of products available in size "M. Using “**COUNTIFS**” Function and using Wildcard “*”.

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

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

OriginalPrice (in Rs)	SizeOptio	Ratings	Discount %
1499	28, 30, 32,	3.9	45
1149	S, M, L, XL	4	55
1399	38, 40, 42,	4.3	55
1295	S, M, L, XL,	4.2	31
599	XS, S, M, L,	4.2	35
599	XS, S, M, L,	4.4	40
1499	30, 32, 34,	3.9	60
1395	S, M, L, XL	3.7	58
1098	XS, S, M, L,	4.3	53
2749	28, 30, 32,	3.5	52

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

number of products available in size "M."	6585

Data Retrieval and Lookup

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

Retrieving Data Using **Xlookup** Function displayed in image below:

Xlookup ---->	Product_id	Brand
	11226634	=XLOOKUP(S5,Table1[Product_id],Table1[BrandName], "Not Found",0)

Xlookup ---->	Product_id	Price
	11226634	=XLOOKUP(S8,Table1[Product_id], Table1[OriginalPrice (in Rs)],"Not Found",0)

Xlookup ---->	Product_id	Rating
	11226634	=XLOOKUP(S11,Table1[Product_id], Table1[Ratings],"Not Found",0)

Product_id --->	11226634
Brand --->	Maniac
Price --->	1199
Rating --->	3.9

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

Retrieving Data Using **INDEX" & "MATCH** Function displayed in image

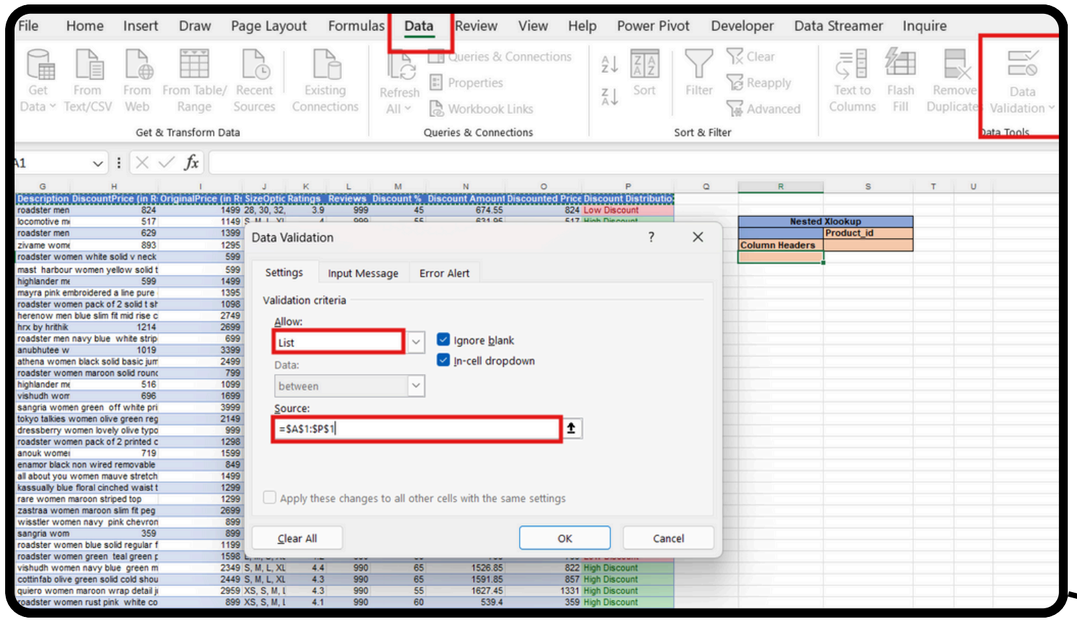
Product_id	INDEX & MATCH for finding DiscountPrice
6744434	=INDEX(Table1[Discounted Price],MATCH(R5,Table1[Product_id],0))

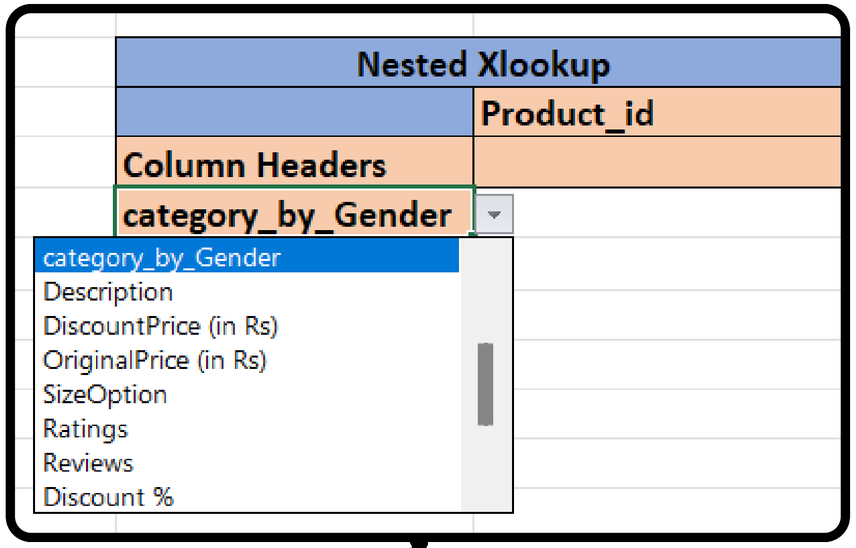
Product_id	INDEX & MATCH for finding DiscountPrice
6744434	599

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

First Using **Data Validation** to create list.

Select cell where you want to create list go to **Data tab** ---> **Data Validation**, in dropdown **select list** and select all **column headers**. do same for **product_id.**





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

Retrieving Data from any Column Using **Nested Xlookup** Function displayed in image.

Nested Xlookup	
	Product_id
Column Headers	2296012
Discounted Price	=XLOOKUP(S5,Table1[Product_id],XLOOKUP(R6, Table1[#Headers],Table1,"Nt Found"))

Nested Xlookup	
	Product_id
Column Headers	2296012
Reviews	999

URL
Product_id
BrandName
Category
Individual_category
category_by_Gender
Description

WSCUBE TECH®
Upskilling Bharat

Thank You

Conenct with me on

Linked_In

GitHub

Email