Analysis of Myntra Apparel

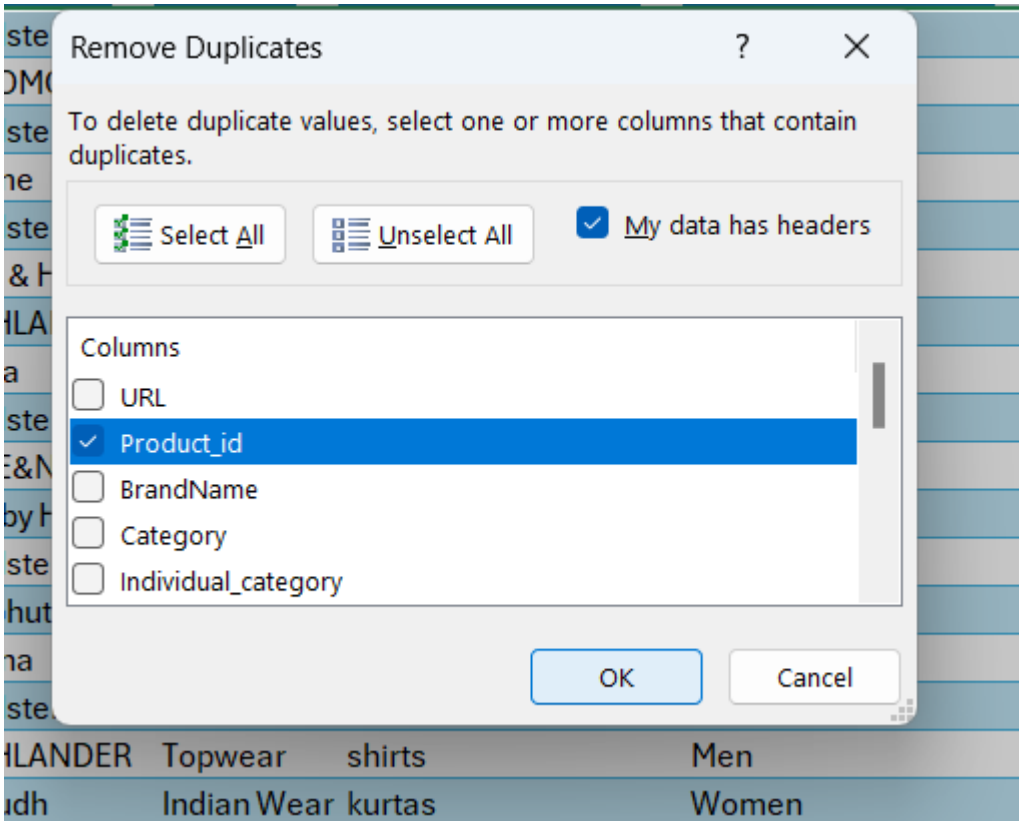
Problem Statement: You are working at Myntra, a leading online fashion retailer. The management has asked you to analyze a dataset of various apparel items to gain insights into pricing, discounts, ratings, and available sizes.

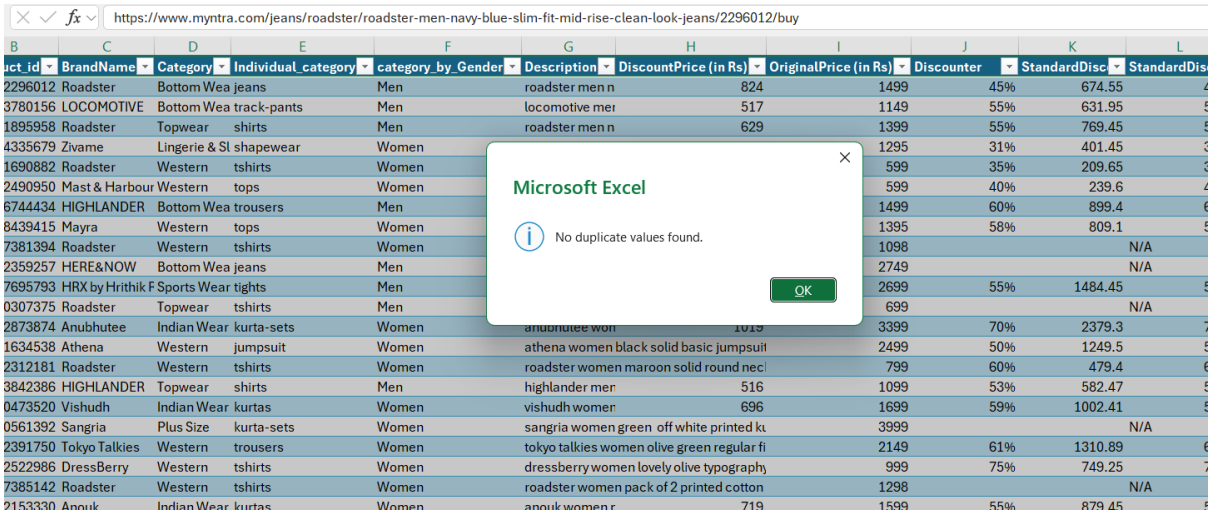
Project Questions

A. Data Cleaning and Preparation

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

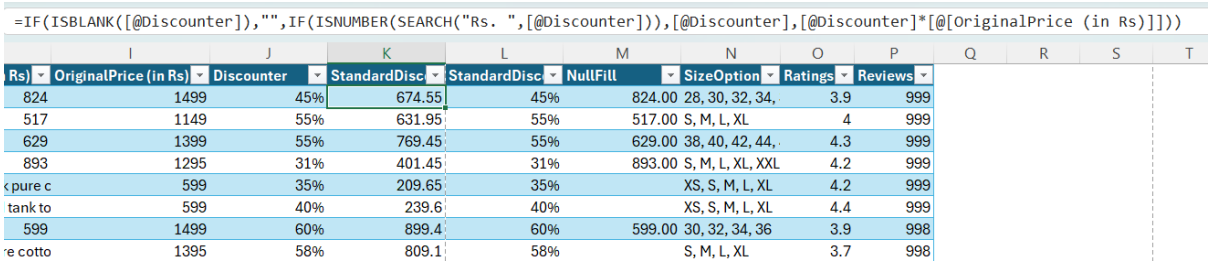
No Duplicates found using Data Tools>> Remove duplicates.





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

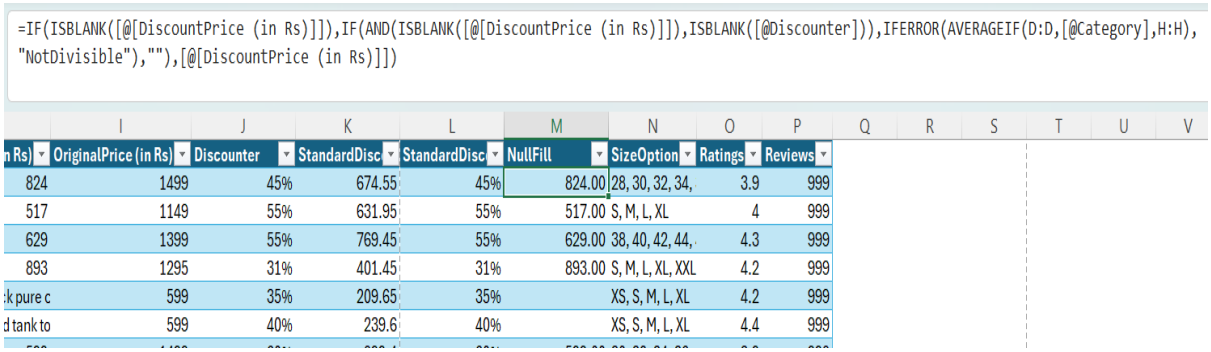
=IF(ISBLANK([@Discounter]),"",IF(ISNUMBER(SEARCH("Rs. ",[@Discounter])),[@Discounter],[@Discounter]*[@[OriginalPrice (in Rs)]]))



Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	NullFill	SizeOption	Ratings	Reviews
824	1499	45%	674.55	45%	824.00	28, 30, 32, 34,	3.9	999
517	1149	55%	631.95	55%	517.00	S, M, L, XL	4	999
629	1399	55%	769.45	55%	629.00	38, 40, 42, 44,	4.3	999
893	1295	31%	401.45	31%	893.00	S, M, L, XL, XXL	4.2	999
k pure c	599	35%	209.65	35%		XS, S, M, L, XL	4.2	999
tank to	599	40%	239.6	40%		XS, S, M, L, XL	4.4	999
599	1499	60%	899.4	60%	599.00	30, 32, 34, 36	3.9	998
e cotto	1395	58%	809.1	58%		S, M, L, XL	3.7	998

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

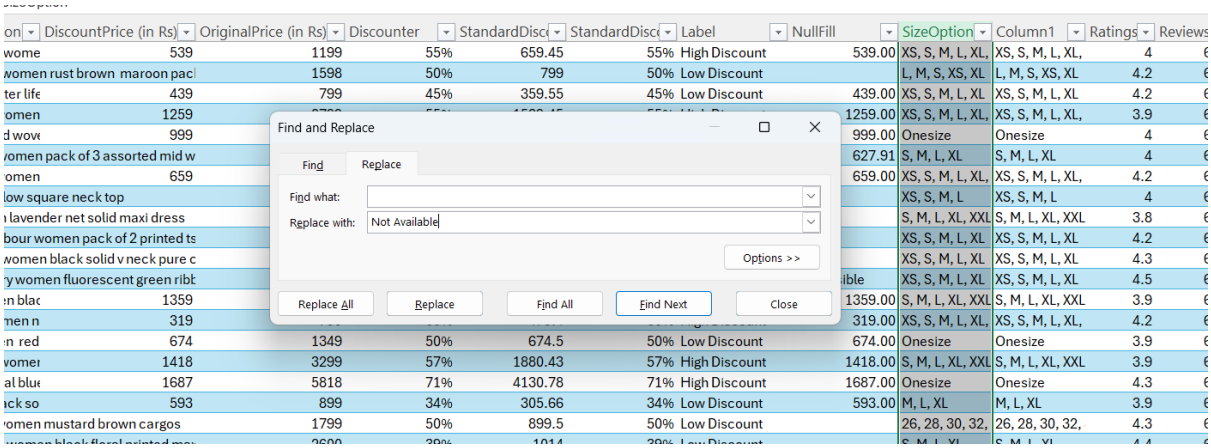
=IF(ISBLANK([@[DiscountPrice (in Rs)]]),IF(AND(ISBLANK([@[DiscountPrice (in Rs)]]),ISBLANK([@Discounter])),IFERROR(AVERAGEIF(D:D,[@Category],H:H),"NotDivisible"),""),[@[DiscountPrice (in Rs)]])



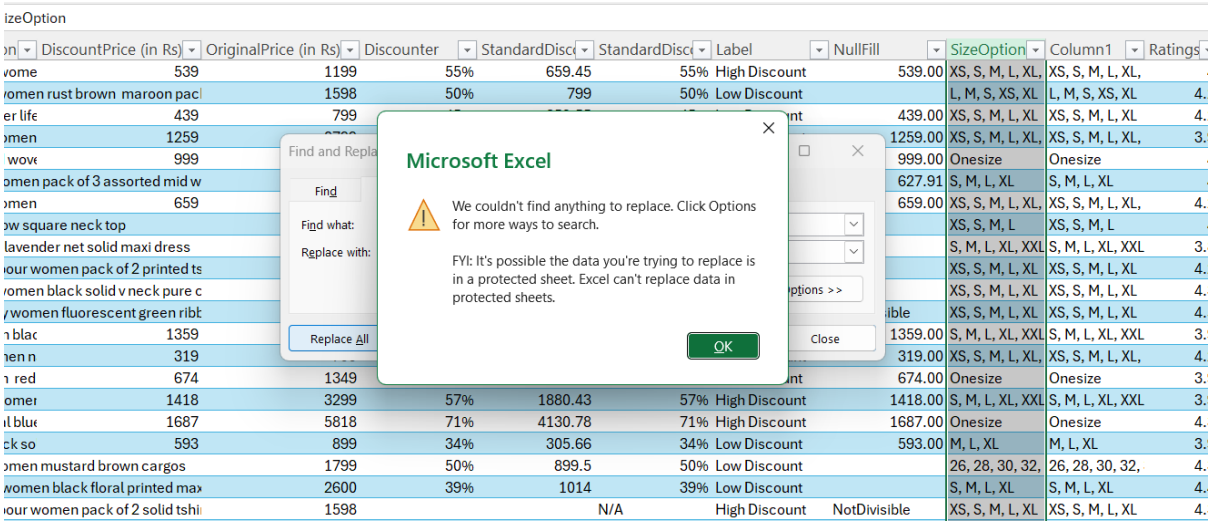
n Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	NullFill	SizeOption	Ratings	Reviews
824	1499	45%	674.55	45%	824.00	28, 30, 32, 34,	3.9	999
517	1149	55%	631.95	55%	517.00	S, M, L, XL	4	999
629	1399	55%	769.45	55%	629.00	38, 40, 42, 44,	4.3	999
893	1295	31%	401.45	31%	893.00	S, M, L, XL, XXL	4.2	999
k pure c	599	35%	209.65	35%		XS, S, M, L, XL	4.2	999
d tank to	599	40%	239.6	40%		XS, S, M, L, XL	4.4	999

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

Using Find and Replace No Null values.



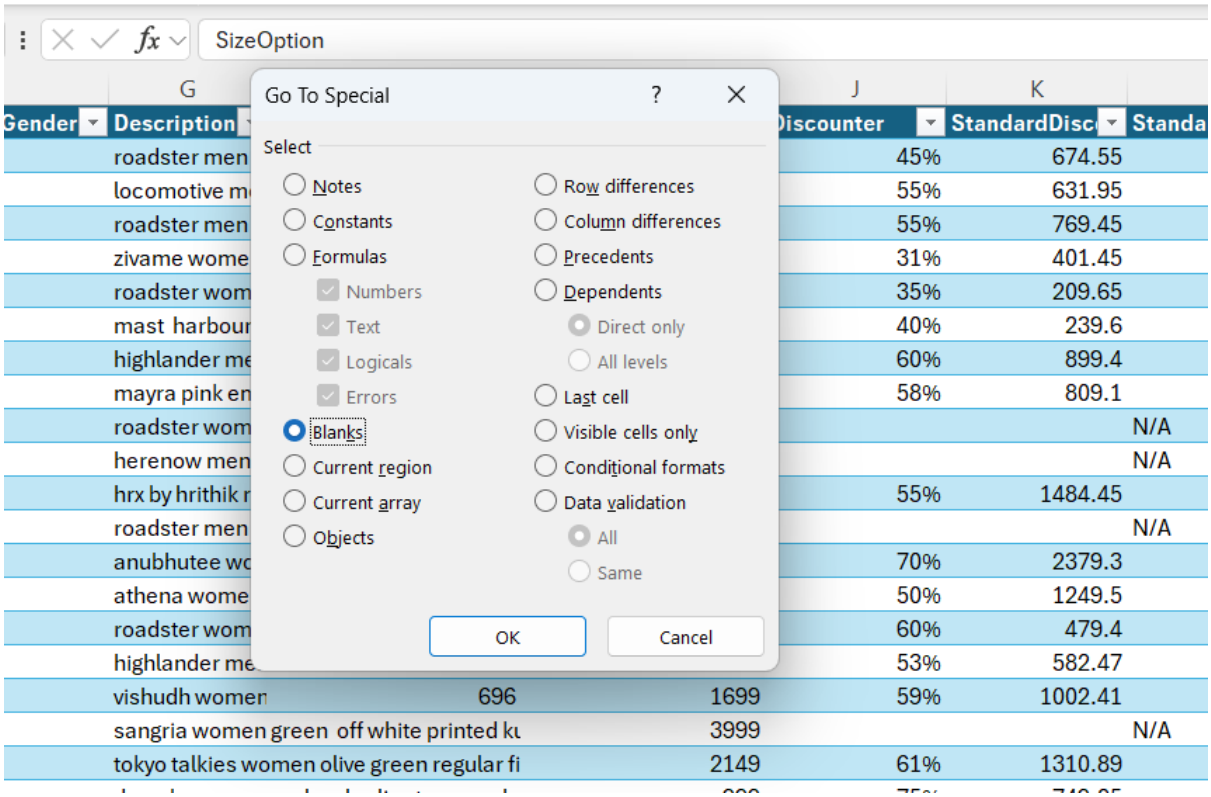
on	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	Label	NullFill	SizeOption	Column1	Ratings	Reviews
wome	539	1199	55%	659.45	55%	High Discount	539.00	XS, S, M, L, XL,	XS, S, M, L, XL,	4	6
/omen rust brown maroon pacl		1598	50%	799	50%	Low Discount		L, M, S, XS, XL	L, M, S, XS, XL	4.2	6
ter life	439	799	45%	359.55	45%	Low Discount	439.00	XS, S, M, L, XL	XS, S, M, L, XL	4.2	6
omen	1259						1259.00	XS, S, M, L, XL,	XS, S, M, L, XL,	3.9	6
d wove	999						999.00	Onesize	Onesize	4	6
/omen pack of 3 assorted mid w							627.91	S, M, L, XL	S, M, L, XL	4	6
omen	659						659.00	XS, S, M, L, XL,	XS, S, M, L, XL,	4.2	6
low square neck top								XS, S, M, L	XS, S, M, L	4	6
lavender net solid maxi dress								S, M, L, XL, XXL	S, M, L, XL, XXL	3.8	6
bour women pack of 2 printed ts								XS, S, M, L, XL	XS, S, M, L, XL	4.2	6
/omen black solid v neck pure c								XS, S, M, L, XL	XS, S, M, L, XL	4.3	6
y women fluorescent green ribb							ible	XS, S, M, L, XL	XS, S, M, L, XL	4.5	6
n blac	1359						1359.00	S, M, L, XL, XXL	S, M, L, XL, XXL	3.9	6
men n	319						319.00	XS, S, M, L, XL,	XS, S, M, L, XL,	4.2	6
n red	674	1349	50%	674.5	50%	Low Discount	674.00	Onesize	Onesize	3.9	6
/omer	1418	3299	57%	1880.43	57%	High Discount	1418.00	S, M, L, XL, XXL	S, M, L, XL, XXL	3.9	6
al blue	1687	5818	71%	4130.78	71%	High Discount	1687.00	Onesize	Onesize	4.3	6
ck so	593	899	34%	305.66	34%	Low Discount	593.00	M, L, XL	M, L, XL	3.9	6
omen mustard brown cargos		1799	50%	899.5	50%	Low Discount		26, 28, 30, 32,	26, 28, 30, 32,	4.3	6

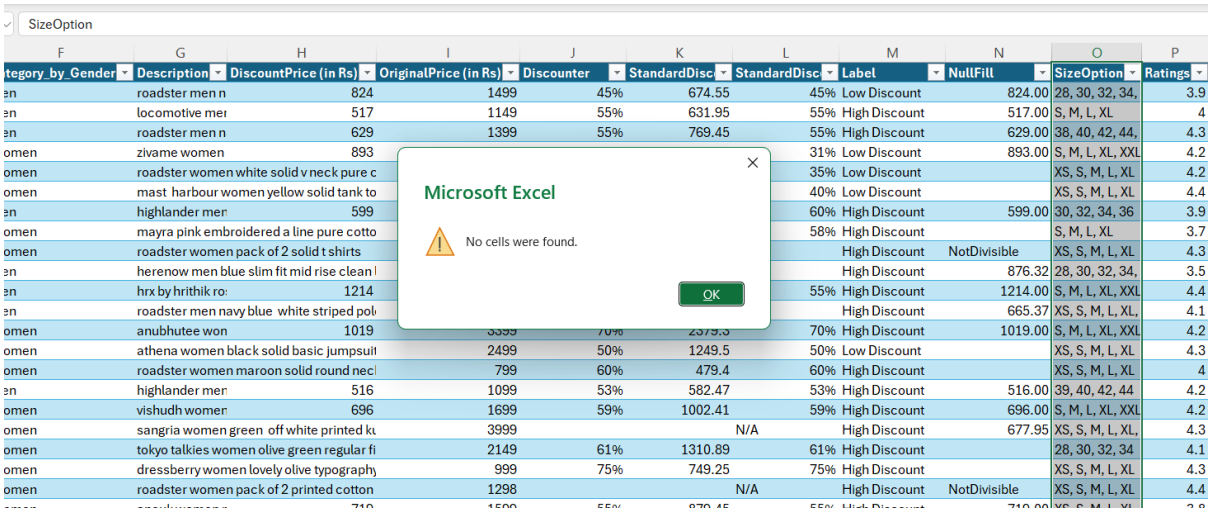


izeOption

on	DiscountPrice (in Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	Label	NullFill	SizeOption	Column1	Ratings
wome	539	1199	55%	659.45	55%	High Discount	539.00	XS, S, M, L, XL,	XS, S, M, L, XL,	
/omen rust brown maroon pacl		1598	50%	799	50%	Low Discount		L, M, S, XS, XL	L, M, S, XS, XL	4.
er life	439	799				nt	439.00	XS, S, M, L, XL	XS, S, M, L, XL	4.
omen	1259						1259.00	XS, S, M, L, XL,	XS, S, M, L, XL,	3.
wove	999						999.00	Onesize	Onesize	
omen pack of 3 assorted mid w							627.91	S, M, L, XL	S, M, L, XL	
omen	659						659.00	XS, S, M, L, XL,	XS, S, M, L, XL,	4.
ow square neck top								XS, S, M, L	XS, S, M, L	
lavender net solid maxi dress								S, M, L, XL, XXL	S, M, L, XL, XXL	3.
our women pack of 2 printed ts								XS, S, M, L, XL	XS, S, M, L, XL	4.
/omen black solid v neck pure c								XS, S, M, L, XL	XS, S, M, L, XL	4.
y women fluorescent green ribb							ible	XS, S, M, L, XL	XS, S, M, L, XL	4.
n blac	1359						1359.00	S, M, L, XL, XXL	S, M, L, XL, XXL	3.
men n	319						319.00	XS, S, M, L, XL,	XS, S, M, L, XL,	4.
n red	674	1349				nt	674.00	Onesize	Onesize	3.
omer	1418	3299	57%	1880.43	57%	High Discount	1418.00	S, M, L, XL, XXL	S, M, L, XL, XXL	3.
al blue	1687	5818	71%	4130.78	71%	High Discount	1687.00	Onesize	Onesize	4.
ck so	593	899	34%	305.66	34%	Low Discount	593.00	M, L, XL	M, L, XL	3.
omen mustard brown cargos		1799	50%	899.5	50%	Low Discount		26, 28, 30, 32,	26, 28, 30, 32,	4.
/omen black floral printed max		2600	39%	1014	39%	Low Discount		S, M, L, XL	S, M, L, XL	4.
our women pack of 2 solid tshi		1598		N/A		High Discount	NotDivisible	XS, S, M, L, XL	XS, S, M, L, XL	4.

Or we can use Find & Select, Go To Special





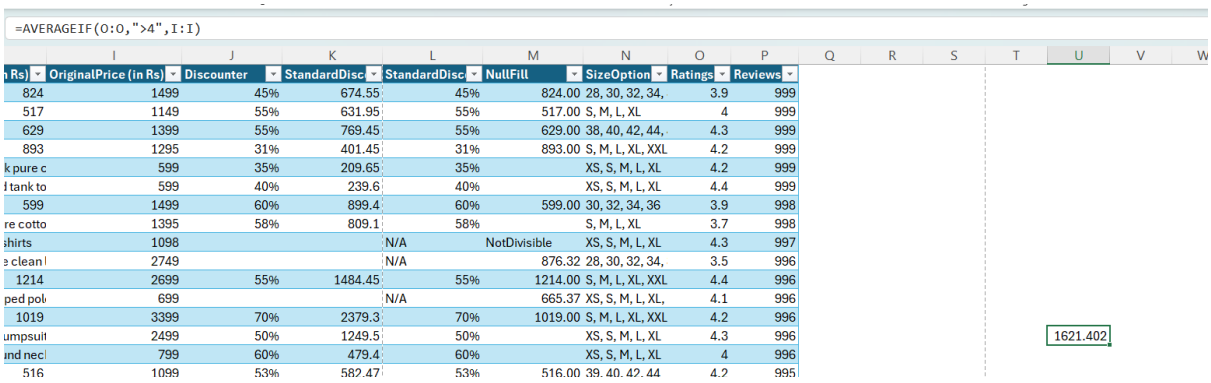
Or we can Use formula in helper column

=IF([@SizeOption]="","Not Available",[@SizeOption])

B. Data Analysis

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

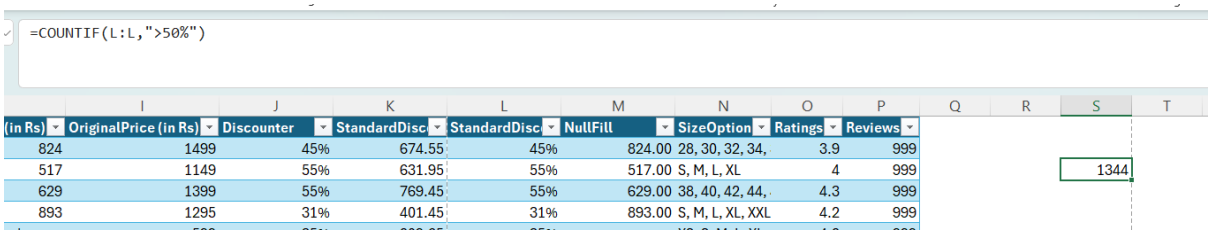
=AVERAGEIF(O:O,">4",I:I)



(in Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	NullFill	SizeOption	Ratings	Reviews
824	1499	45%	674.55	45%	824.00	28, 30, 32, 34,	3.9	999
517	1149	55%	631.95	55%	517.00	S, M, L, XL	4	999
629	1399	55%	769.45	55%	629.00	38, 40, 42, 44,	4.3	999
893	1295	31%	401.45	31%	893.00	S, M, L, XL, XXL	4.2	999
k pure c	599	35%	209.65	35%		XS, S, M, L, XL	4.2	999
l tank to	599	40%	239.6	40%		XS, S, M, L, XL	4.4	999
599	1499	60%	899.4	60%	599.00	30, 32, 34, 36	3.9	998
re cotto	1395	58%	809.1	58%		S, M, L, XL	3.7	998
shirts	1098			N/A	NotDivisible	XS, S, M, L, XL	4.3	997
e clean	2749			N/A	876.32	28, 30, 32, 34,	3.5	996
1214	2699	55%	1484.45	55%	1214.00	S, M, L, XL, XXL	4.4	996
ped pol	699			N/A	665.37	XS, S, M, L, XL,	4.1	996
1019	3399	70%	2379.3	70%	1019.00	S, M, L, XL, XXL	4.2	996
umpsuit	2499	50%	1249.5	50%		XS, S, M, L, XL	4.3	996
nd nec	799	60%	479.4	60%		XS, S, M, L, XL	4	996
516	1099	53%	582.47	53%	516.00	39, 40, 42, 44	4.2	995

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

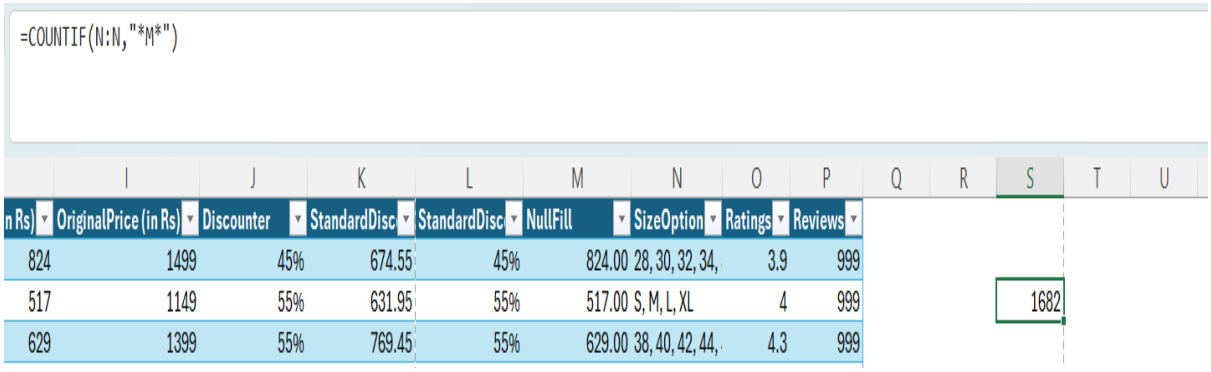
=COUNTIF(L:L,">50%")



(in Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	NullFill	SizeOption	Ratings	Reviews
824	1499	45%	674.55	45%	824.00	28, 30, 32, 34,	3.9	999
517	1149	55%	631.95	55%	517.00	S, M, L, XL	4	999
629	1399	55%	769.45	55%	629.00	38, 40, 42, 44,	4.3	999
893	1295	31%	401.45	31%	893.00	S, M, L, XL, XXL	4.2	999

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

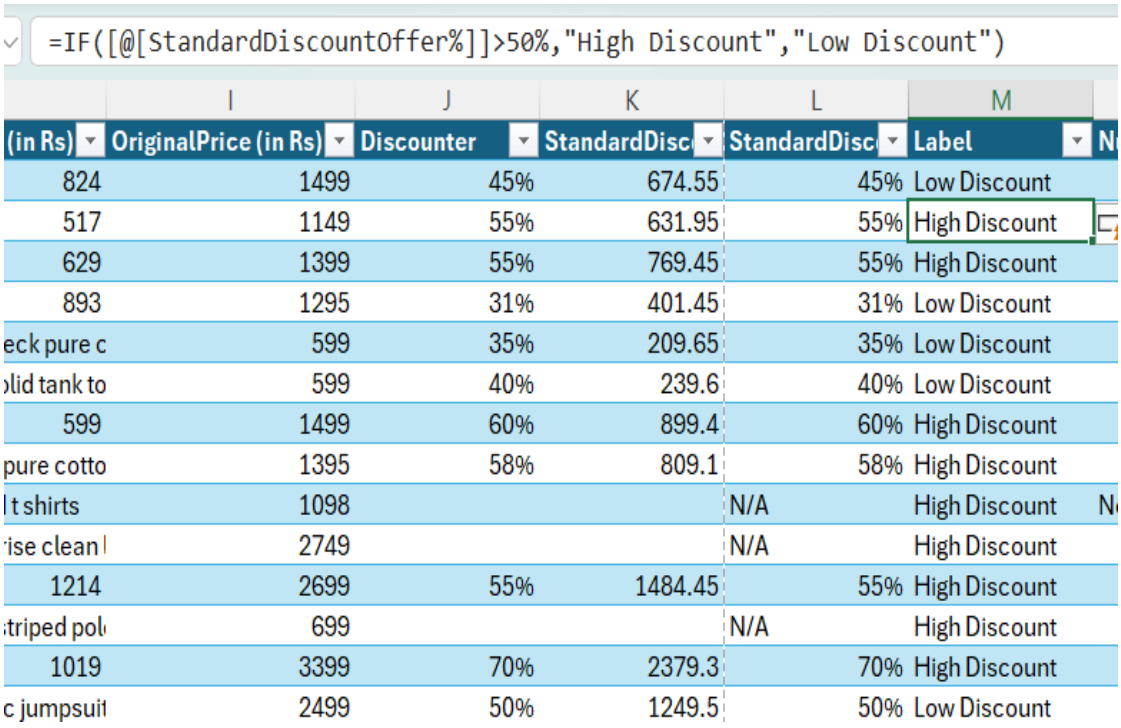
=COUNTIF(N:N,"*M*")



n Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	NullFill	SizeOption	Ratings	Reviews
824	1499	45%	674.55	45%	824.00	28, 30, 32, 34,	3.9	999
517	1149	55%	631.95	55%	517.00	S, M, L, XL	4	999
629	1399	55%	769.45	55%	629.00	38, 40, 42, 44,	4.3	999

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."

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



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

	I	J	K	L	M
(in Rs)	OriginalPrice (in Rs)	Discounter	StandardDisc	StandardDisc	Label
824	1499	45%	674.55	45%	Low Discount
517	1149	55%	631.95	55%	High Discount
629	1399	55%	769.45	55%	High Discount
893	1295	31%	401.45	31%	Low Discount
eck pure c	599	35%	209.65	35%	Low Discount
olid tank to	599	40%	239.6	40%	Low Discount
599	1499	60%	899.4	60%	High Discount
pure cotto	1395	58%	809.1	58%	High Discount
l t shirts	1098			N/A	High Discount
rise clean l	2749			N/A	High Discount
1214	2699	55%	1484.45	55%	High Discount
triped pol	699			N/A	High Discount
1019	3399	70%	2379.3	70%	High Discount
c jumpsuit	2499	50%	1249.5	50%	Low Discount

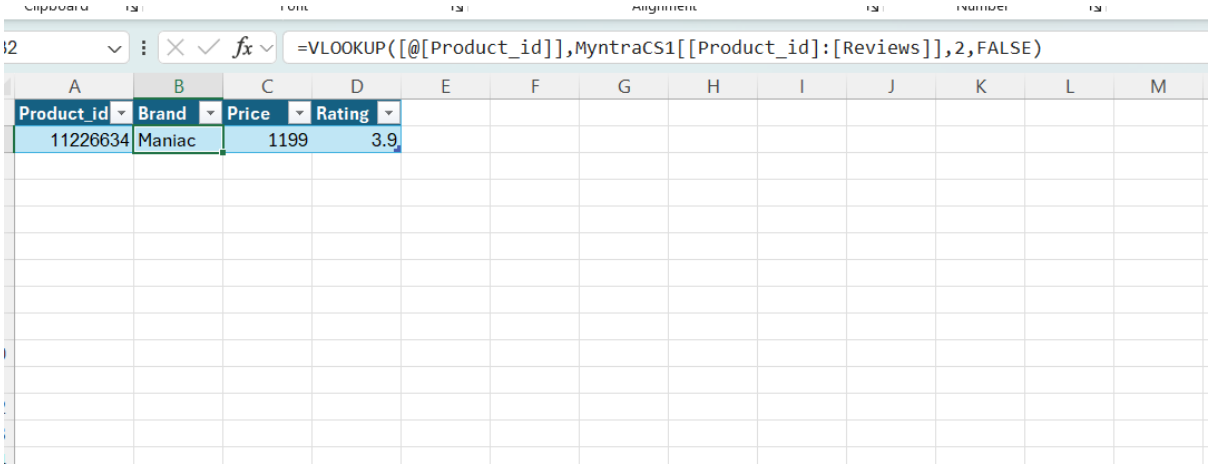
C. Data Retrieval and Lookup

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

=VLOOKUP([@[Product_id]],MyntraCS1[[Product_id]:[Reviews]],2,FALSE)

=VLOOKUP([@[Product_id]],MyntraCS1[[Product_id]:[Reviews]],8,FALSE)

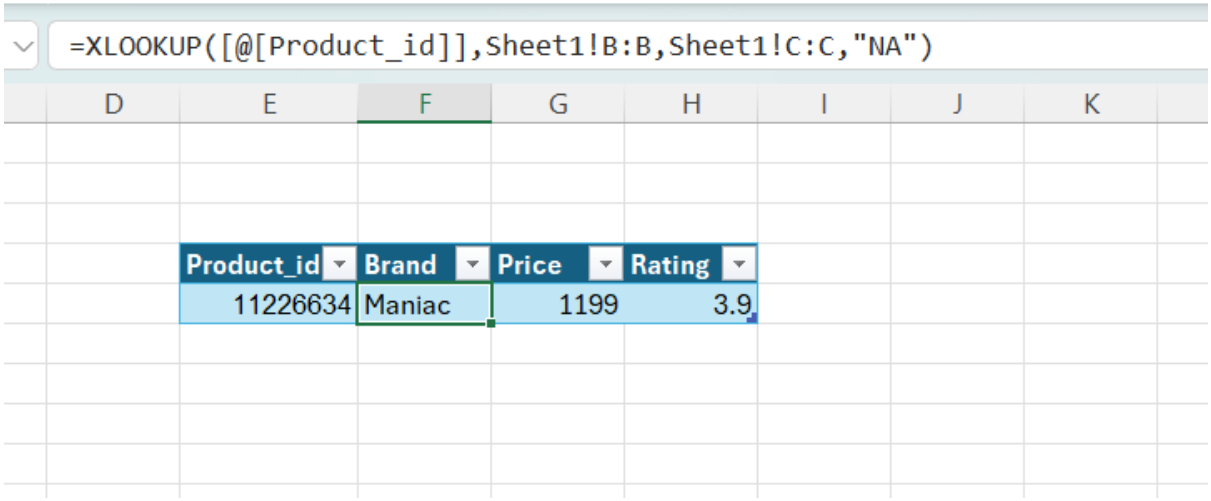
=VLOOKUP([@[Product_id]],MyntraCS1[[Product_id]:[Reviews]],15,FALSE)



=XLOOKUP([@[Product_id]],Sheet1!B:B,Sheet1!C:C,"NA")

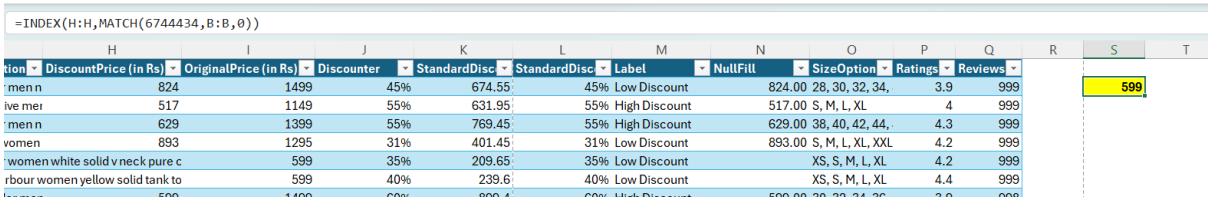
=XLOOKUP([@[Product_id]],Sheet1!B:B,Sheet1!I:I,"NA")

=XLOOKUP([@[Product_id]],Sheet1!B:B,Sheet1!P:P,"NA")



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

=INDEX(H:H,MATCH(6744434,B:B,0))



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

For this I Created a dropdown list for product id and Header.

=XLOOKUP(Table4[Product_id],Sheet1!B:B,XLOOKUP(Table4[SelectHeader],MyntraCS1[[#Headers],[Product_id]:[Reviews]],Sheet1!B:Q))