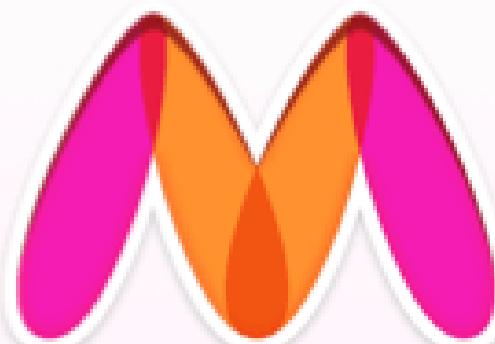


# **ANALYSIS OF MYNTRA APPAREL**

**Analyzed By Pankaj Kumar**





# ABOUT PROJECT

**Objective:** To clean, analyze, and retrieve data from a dataset using Excel.

**Data Cleaning:**

- Removed duplicates.
- Standardized "DiscountOffer" values.
- Filled null "DiscountPrice" with category averages.
- Replaced null "SizeOption" with "Not Available."

**Data Analysis:**

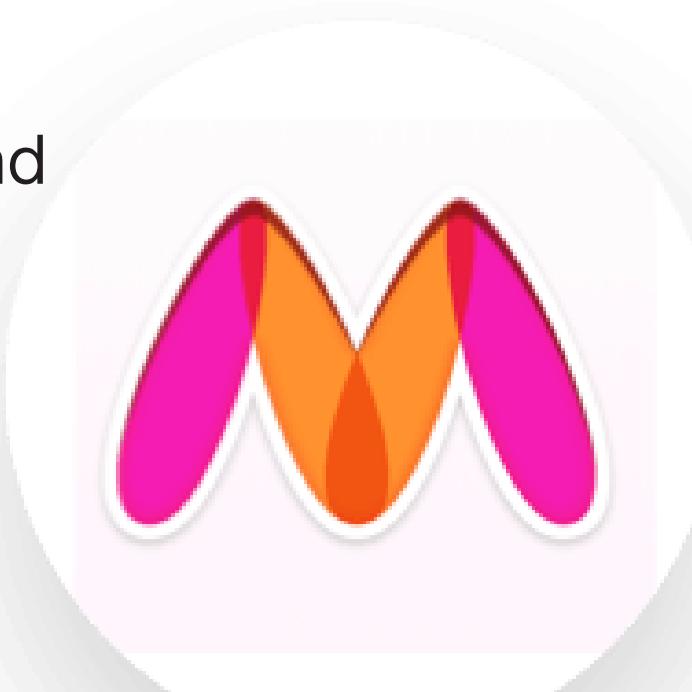
- Calculated average original price for high-rated products.
- Counted products with discounts > 50%.
- Counted products in size "M."
- Labeled products as "High Discount" or "Low Discount."

**Data Retrieval:**

- Used VLOOKUP/XLOOKUP to find details for Product\_id "11226634."
- Applied INDEX/MATCH to find "DiscountPrice" for Product ID "6744434."

**Tools Used:** Excel

**Outcome:** Achieved improved data accuracy and insights into product pricing and discount patterns.



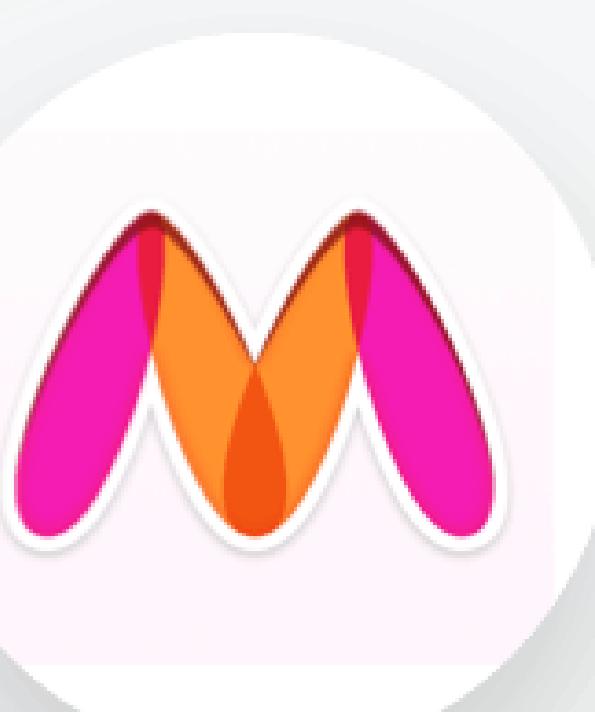
# DATA CLEANING AND PREPARATION



Removed  
Duplicate values

I have checked Data set  
and removed duplicate  
value based on product  
Id.

The screenshot shows a data cleaning interface with a table of product descriptions in the background. A modal dialog box titled "Remove Duplicates" is open in the foreground. The dialog includes instructions: "To delete duplicate values, select one or more columns that contain duplicates." It features "Select All" and "Unselect All" buttons, and a checked checkbox for "My data has headers". A list of columns is provided, with "Product\_id" selected (indicated by a blue highlight). At the bottom are "OK" and "Cancel" buttons. The background table lists various product descriptions such as "roadster men navy blue slim fit mid rise clean look jeans", "locomotive men black", and "vishudh women blue pink floral print a line kurta".



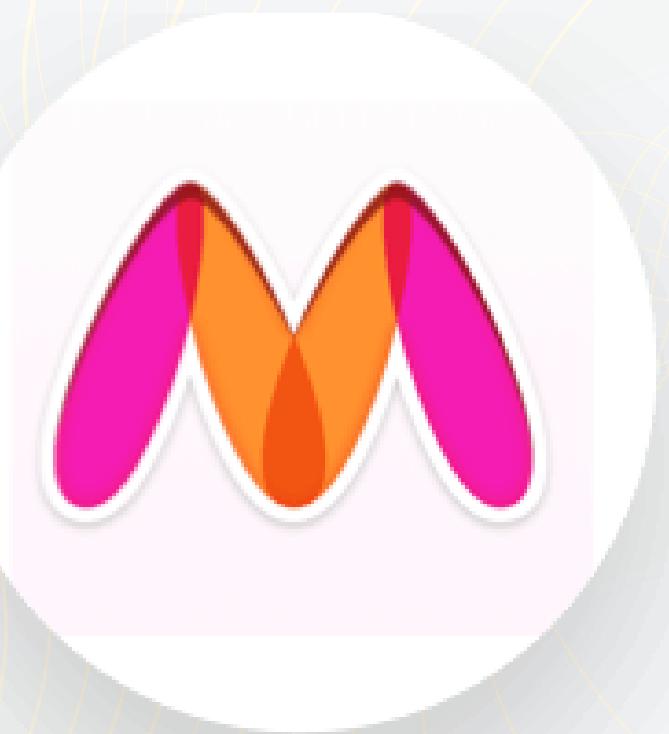
# DATA CLEANING AND PREPARATION

Replaced Rupees  
with blank

- I have Replaced rupees value with blank value to Standardize the Discount Offer Column

M	N	O	P	Q	R
Reviews	Column1				
3.9	999 =TRIM(SUBSTITUTE([@DiscountOffer],"Rs.",""))				
4	999 5 TRIM(text)				
4.3	999 55% OFF				
4.2	999 31% OFF				
4.2	999 35% OFF				
4.4	999 40% OFF				

# DATA CLEANING AND PREPARATION



# Extracting Numeric amount

1. Extracted Numeric amount from discount offer,  
Converted it into  
Discounted amount by  
below formula

# DATA CLEANING AND PREPARATION

Converting Discount amount To percentage again

1. After this i have converted numeric amount to percentage discount again by using below formula

	M	N	O	P	Q	R	S	T	U	V
	Reviews	Discount	Column1	Column2						
9	999	674.55	=Table2[@Discount]/Table2[@[OriginalPrice (in Rs)]]*100							
4	999	631.95		55	55					
3	999	769.45		55	55					
2	999	401.45		31	31					
2	999	209.65		35	35					

# DATA CLEANING AND PREPARATION

Replaced zero with blank values

1. Replaced Zero with blank values using below formula

M	N	O	P	Q	R	S	T	U
Reviews	Discount	Column1	Column2					
999	674.55	45	=IF([@Column1]=0,"",[@Column1])					
999	631.95	55	IF(logical_test, [value_if_true], [value_if_false])					
999	769.45	55						
999	401.45	31						
999	209.65	35						

# DATA CLEANING AND PREPARATION

Filled Blank value based on criteria

1. I had filled blank value using averageifs formula based on category criteria.

K	L	M	N	O	P	Q	R	S	T	U
Option	Ratings	Reviews	Discount A	Discount Percentage	Column1	Column2				
30, 32, 34,	=IF([@Discount Percentage]=""",AVERAGEIFS([Discount Percentage],Table2[Category],Table2[@Category]),[@Discount Percentage]))									
I, L, XL	4	999	631.	AVERAGEIFS(average_range, criteria_range1, criteria1, [criteria_range2, criteria2], ...)						
40, 42, 44,	4.3	999	769.45		55	55	55			
I, L, XL, XXL	4.2	999	401.45		31	31	31			
S, M, L, XL	4.2	999	209.65		35	35	35			
S, M, L, XL	4.4	999	239.6		40	40	40			

# DATA CLEANING AND PREPARATION

## Discount Offer Calculation

1. Done a basic discount offer calculation based on original price

M	N	O	P
Discount Offer	Discount Amount	Discount Price	
9	=([@Discount Offer])/100)*Table2[@[OriginalPrice (in Rs)]]		
9	55	631.95	517.05

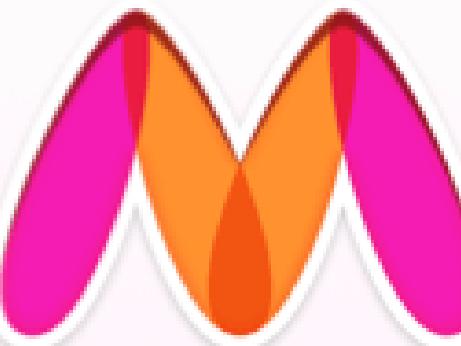
# DATA CLEANING AND PREPARATION

## Discount amount Calculation

1. Done discount amount calculation using below approach
2. Now we had standardized all column.

fer	N	O	P	Q	R	S
er	Discount Amount	Discount Price				
	=Table2[@[OriginalPrice (in Rs)]]-[@[Discount Amount]]					
	631.95	517.05				
	769.45	629.55				
	401.45	893.55				

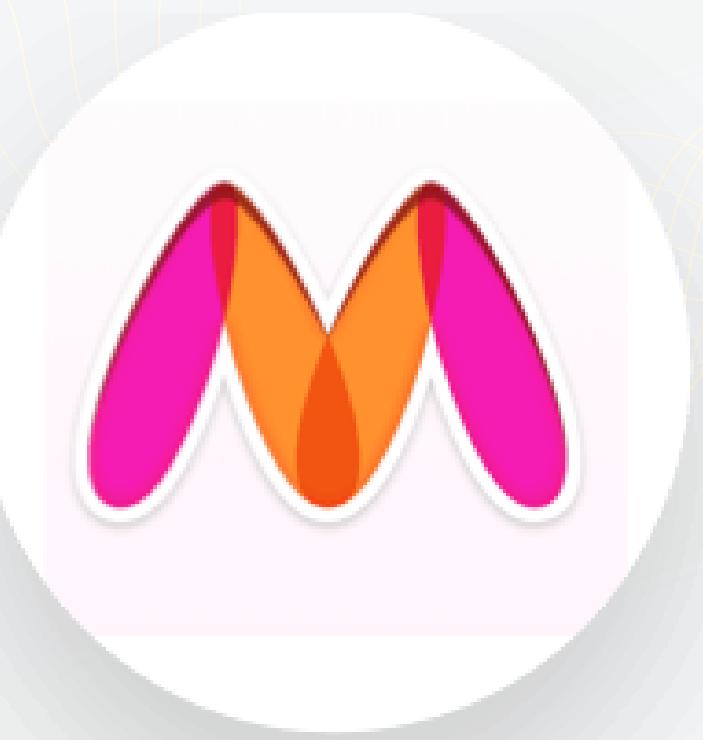
# DATA ANALYSIS



I have calculated the overall average original price for products with ratings greater than 4

V	O	P	Q	R	S
Amount	Discount Price	Column1			
74	=AVERAGEIFS(Table2[OriginalPrice (in Rs)],Table2[Ratings],>4")				
31	AVERAGEIFS(average_range, criteria_range1, criteria1, [criteria_range2, criteria2], .				
59	629	1751.028128			

# DATA ANALYSIS



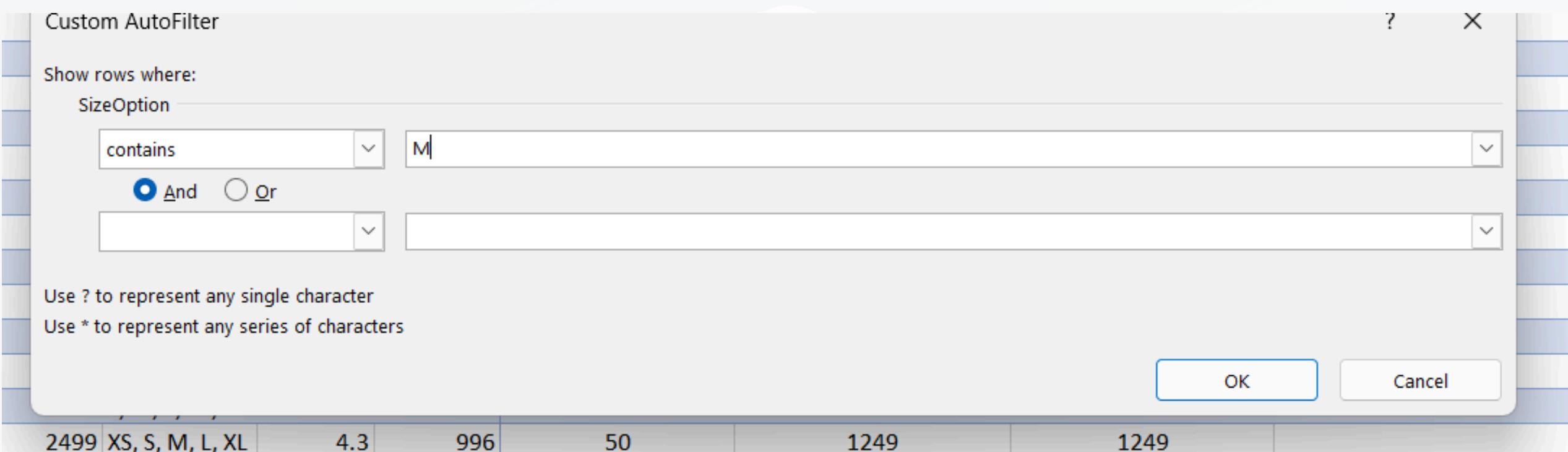
I have counted the number of products with a discount offer greater than 50% OFF.

A screenshot of a spreadsheet application showing a formula in cell Q1. The formula is =COUNTIF(M:M,>50"). A tooltip below the formula provides the definition: COUNTIF(range, criteria). The range M:M is highlighted in blue, and the criteria >50" is highlighted in orange.



# DATA ANALYSIS

I had Count the number of products available in size "M"  
By using below approach

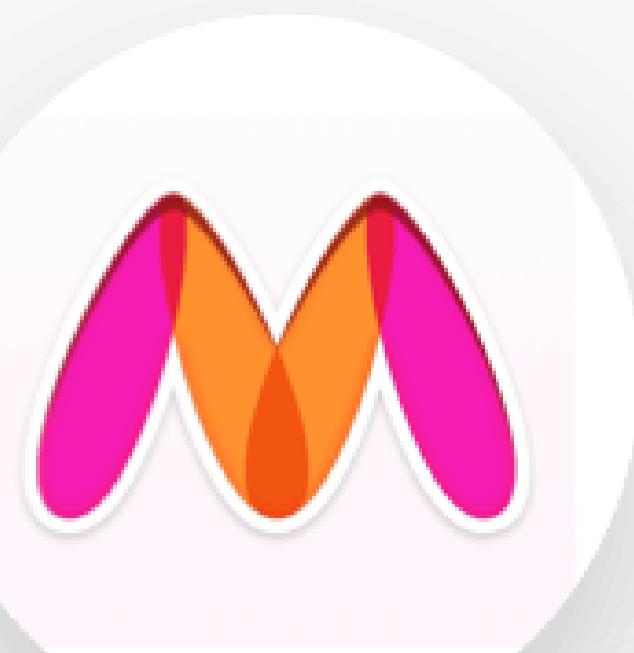


# DATA ANALYSIS



I have Created 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."

R	S	T	U
Products available in size "M"	High/Low Discount		
	=IF([@Discount Offer]>50,"High Discount","Low Discount")		
		IF(logical_test, [value_if_true], [value_if_false])	
		High Discount	
		Low Discount	



# DATA RETRIEVAL AND LOOKUP

1. I Used VLOOKUP/XLOOKUP to find the product brand, price, and rating of the product with Product\_id "11226634", separately.

S	T	U	V
High/Low Discount	Specific Product id	xlookup Product Name	
Low Discount	11226634	=XLOOKUP([@[Specific Product id]],Table2[Product_id],Table2[BrandName])	
High Discount		XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])	
High Discount			
Low Discount			
Low Discount			

U	V	W	X	Y	Z
Specific Product Name	Specific product Price				
	=XLOOKUP([@[Specific Product id]],Table2[Product_id],Table2[OriginalPrice (in Rs)])				
	XLOOKUP(lookup_value, lookup_array, return_array, [if_not_found], [match_mode], [search_mode])				

V	W	X	Y
Specific product Price	Specific Product rating		
1199	=VLOOKUP([@[Specific Product id]],Table2[[#All],[Product_id]:[Ratings]],10,0)		
	VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])		

# DATA RETRIEVAL AND LOOKUP

I had found the "DiscountPrice" for the product with the Product ID "6744434" using the INDEX and MATCH functions

X	Y	Z	AA
DiscountPrice For Specific id 2			
34 =INDEX(N:N,MATCH([@[Specific product id 2]],[Product_id],0))			
	MATCH(lookup_value, lookup_array, [match_type])		

## Final project Link:

<https://easyupload.io/om9o0p>

Thank

YOU