

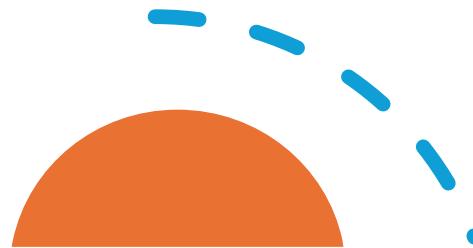
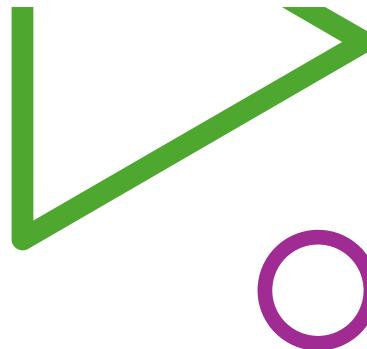


# Analysis of Myntra Apparel

- Presented By: Kunal Tripathi

## 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



### Data Cleaning and Preparation

Check for duplicate values in your dataset and remove them.

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

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

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



### Data Analysis

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.

Count the number of products available in size "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."



### Data Retrieval and Lookup

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

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

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



## Data Cleaning and Preparation

URL	Product_id	BrandName	Category	Individual category	Description	DiscountP	OriginalPr	DiscountC	SizeOptio	Ratings	Reviews
2 https://vn 2296012	Roadster	Bottom W jeans	Men	roadster n	824	1499	45% OFF	28, 30, 32,	4	999	
3 https://vn 13780156	LOCOMOT	Bottom W track-pant	Men	locomotiv	517	1149	55% OFF	S, M, L, XL	4	999	
4 https://vn 11895958	Roadster	Topwear shirts	Men	roadster n	629	1399	55% OFF	38, 40, 42,	4.3	999	
5 https://vn 4335679	Zivame	Lingerie & shapewa	Women	zivame wc	893	1295	31% OFF	S, M, L, XL	4.2	999	
6 https://vn 11690882	Roadster	Western tshirts	Women	roadster women whi	599	599	35% OFF	XS, S, M, L	4.2	999	
7 https://vn 2490950	Mast & Hi	Western tops	Women	mast harbour wome	599	599	40% OFF	XS, S, M, L	4.4	999	
8 https://vn 6744434	HIGHLANE	Bottom W trousers	Men	highlande	599	1499	60% OFF	30, 32, 34,	3.9	998	
9 https://vn 8439415	Mayra	Western tops	Women	mayra pink embroid	1395	58% OFF	S, M, L, XL	3.7	998		
10 https://vn 17381394	Roadster	Western tshirts	Women	roadster women pac	997						
11 https://vn 2359257	HERE&NO	Bottom W jeans	Men	herenow men blue s	3.5	996					
12 https://vn 7695793	HRX by Hr Sports	Weights	Men	hrx by hrit	1214	4.4	996				
13 https://vn 10307375	Roadster	Topwear tshirts	Men	roadster men navy b	4.1	996					
14 https://vn 12873874	Anubhute	Indian Wekurta-sets	Women	anubhutei 1019	4.2	996					
15 https://vn 11634538	Athena	Western jumpsuit	Women	athena women black	4.3	996					
16 https://vn 2312181	Roadster	Western tshirts	Women	roadster women ma	996						
17 https://vn 13842386	HIGHLANE	Topwear shirts	Men	highlande	516	1099	60% OFF	XS, S, M, L	4	996	
18 https://vn 10473520	Vishudhi	Indian We kurta	Women	vishudhi w	696	1699	59% OFF	39, 40, 42,	4.2	995	
19 https://vn 10561392	Sangria	Plus Size kurtas-sets	Women	sangria women gree	3999	1699	59% OFF	XS, S, M, L	4.3	995	
20 https://vn 12391750	Tokyo Tall	Western trousers	Women	tokyo talkies women	2149	2149	61% OFF	28, 30, 32,	4.1	995	
21 https://vn 2522986	DressBerr	Western tshirts	Women	dressberry women k	999	999	75% OFF	XS, S, M, L	4.3	995	
22 https://vn 17385142	Roadster	Western tshirts	Women	roadster women pac	1298	1599	55% OFF	XS, S, M, L	4.4	994	
23 https://vn 12153330	Anouk	Indian We kurta	Women	anouk wo	719	849	55% OFF	S, M, L, XL	3.8	993	
24 https://vn 1864573	Enamor	Lingerie & bra	Women	enamor black non w	849	1499	50% OFF	XS, S, M, L	4.2	993	
25 https://vn 13205276	all about y	Western shirts	Women	all about you womer	993	1205	62% OFF	XS, S, M, L	4.3	992	
26 https://vn 11535928	KASSULL'	Western tops	Women	kassull blue floral i	1299	1299	62% OFF	XS, S, M, L			

No duplicate values found.

OR

Remove Duplicates

To delete duplicate values, select one or more columns that contain duplicates.

Select All Unselect All My data has headers

Columns:

- URL
- Product\_id
- BrandName
- Category
- Individual category

OK Cancel

Ready Accessibility:Unavailable Average:4483873.161 Count:5905577 Sum:7936696+12

Average:4483873.161 Count:5905577 Sum:7936696+12

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

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

**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("%",[@DiscountedAmount])),LEFT([@DiscountedAmount],SEARCH("%",[@DiscountedAmount])-1)/100)*[@[OriginalPrice (in Rs)]],[@[DiscountedAmount]]))						
1149 55% OFF	S, M, L, XL	4	999 55% OFF	=IF(ISNUMBER(SEARCH("%",[@DiscountedAmount])),LEFT([@DiscountedAmount],SEARCH("%",[@DiscountedAmount])-1)/100)*[@[OriginalPrice (in Rs)]],[@[DiscountedAmount]]))						
1399 55% OFF	38, 40, 42, 44,	4.3	999 55% OFF	=IF(ISNUMBER(SEARCH("%",[@DiscountedAmount])),LEFT([@DiscountedAmount],SEARCH("%",[@DiscountedAmount])-1)/100)*[@[OriginalPrice (in Rs)]],[@[DiscountedAmount]]))						
1295 31% OFF	S, M, L, XL, XXL	4.2	999 31% OFF	=IF(ISNUMBER(SEARCH("%",[@DiscountedAmount])),LEFT([@DiscountedAmount],SEARCH("%",[@DiscountedAmount])-1)/100)*[@[OriginalPrice (in Rs)]],[@[DiscountedAmount]]))						
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						

**Standardize "DiscountOffer" column to uniform format.**

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

# "DiscountPrice" fill with category average.

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

The screenshot shows a table of product data. The columns include 'Original Price (in Rs)', 'DiscountOffer', 'SizeOption', 'Ratings', 'Reviews', 'Discount %', and 'Category'. A context menu is open over the 'Discount %' column, specifically over the cell containing '55% OFF'. The menu options include 'Sort Smallest to Largest', 'Sort Largest to Smallest', and 'Sort by Color'. A blue arrow from the text below points to this menu.

Original Price (in Rs)	DiscountOffer	SizeOption	Ratings	Reviews	Discount %	Category
599 60% OFF	S, M, L, XL	Sort Smallest to Largest				
599 45% OFF	S, M, L, XL	Sort Largest to Smallest				
1699 50% OFF	X5, S, M, L, XL	Sort by Color				
1198 50% OFF	X5, S, M, L, XL					
1498 50% OFF	M, S, XS, XL					
1499 60% OFF	X5, S, M, L, XL					
1799 55% OFF	X5, S, M, L, XL					
1299 60% OFF, Hurry!	X5, S, M, L, XL					
1999 55% OFF	28, 30, 32, 34					
799 50% OFF	S, M, L, XL, XXL					
899 50% OFF	S, M, L, XL, XXL					
1299 50% OFF	S, M, L, XL					
699 45% OFF	S, M, L, XL, XXL					
1295 31% OFF	S, M, L, XL, XXL					
999 51% OFF	Onesize					
1499 40% OFF	X5, S, M, L, XL					
1699 55% OFF	X5, S, M, L, XL					
2299 63% OFF	S, M, L, XL, XXL					
2699 50% OFF	28, 30, 32, 34					
399	Onesize					
2199 50% OFF	39, 40, 42, 44					
2599 66% OFF	S, M, L, XL, XXL					
1699 60% OFF	26, 28, 30, 32					
799	X5, S, M, L, XL					

Convert all numbers to percentages.

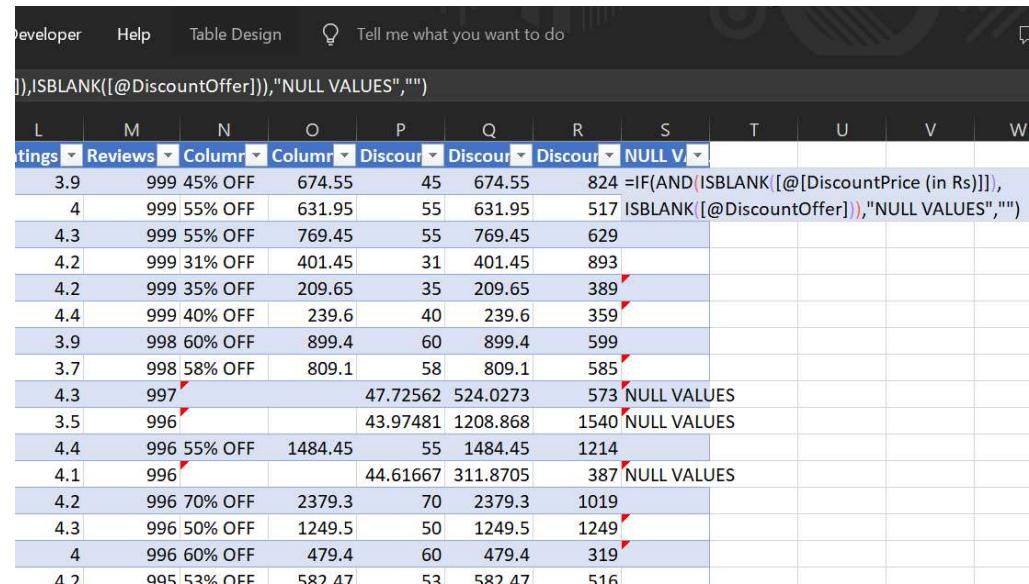
The screenshot shows an Excel spreadsheet with a table. The columns are labeled 'Discount %' and 'Category'. The formula `=IF([@Discounted %]=""",AVERAGEIFS(Table1[[#All],[Discounted %]],Table1[[#All],[Category]],[@Category]),[@Discounted %])` is entered in the first cell of the 'Discount %' column. The table contains several rows of data, mostly blank in the 'Discount %' column. A blue arrow from the text below points to the formula in the first cell.

Discount %	Category
45% OFF	674.55
55% OFF	631.95
55% OFF	769.45
31% OFF	401.45
35% OFF	209.65
40% OFF	239.6
60% OFF	899.4
58% OFF	809.1
55% OFF	1484.45

Figuring out the typical "Discount Offer" in %

## Finding rows having null values both in “Discount Price” and “Discount Offer”

- 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')`



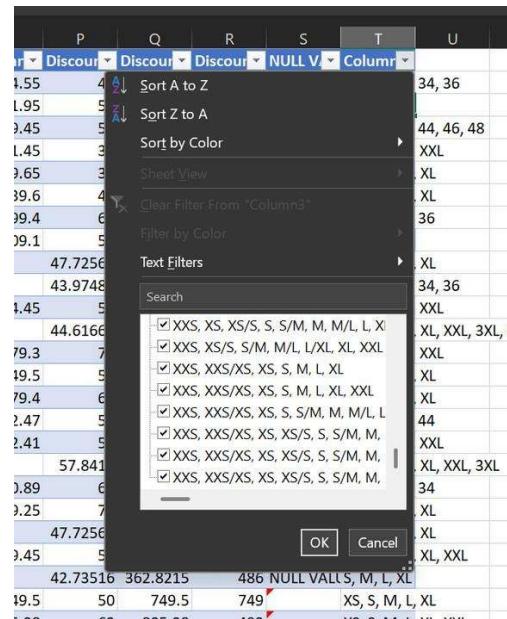
The screenshot shows a Microsoft Excel spreadsheet with a table containing product information. The table has columns L through W. The formula bar displays the formula: `=IF(AND(ISBLANK([@DiscountPrice]),ISBLANK([@DiscountOffer])), "NULL VALUES", "")`. The data in the table includes various discount percentages and prices. Several rows are highlighted in blue, indicating they contain null values in both the Discount Price and Discount Offer columns. The last row, which contains valid data, is highlighted in grey.

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]),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 53% OFF	582.47	53	582.47	516						

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	385				
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.

Tell me what you want to do								
T	U	V	W	X	Y	Z	AA	A
<b>Average of Original Price whose rating is more than 4</b> =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**

=COUNTIFS(P:P, ">50%")

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 available in size "M"**

=COUNTIF(K:K, "M")

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".

## 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."

> 50, "High Discount", "Low Discount")			
Q	R	S	T
Amount	Discounted Price	NULL V	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	Average of Original Price whose rating is more than 4
209.65	389	Low Discount	1966.667
239.6	359	Low Discount	
899.4	599	High Discount	
809.1	585	High Discount	Count the number of products with a discount offer greater than 50% OFF
524.027332	573	NULL V ALL Low Discount	526555
1208.867624	1540	NULL V ALL Low Discount	
1484.45	1214	High Discount	Count the number of products available in size "M"
311.8705433	387	NULL V ALL Low Discount	656
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

VLOOKUP		11226634
BRAND		
Maniac		
PRICE		
Sports Wear		
RATING		
tshirts		

### BRAND

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

### RATING

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

### PRICE

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

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

**Index and Match for Product ID "6744434"**

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

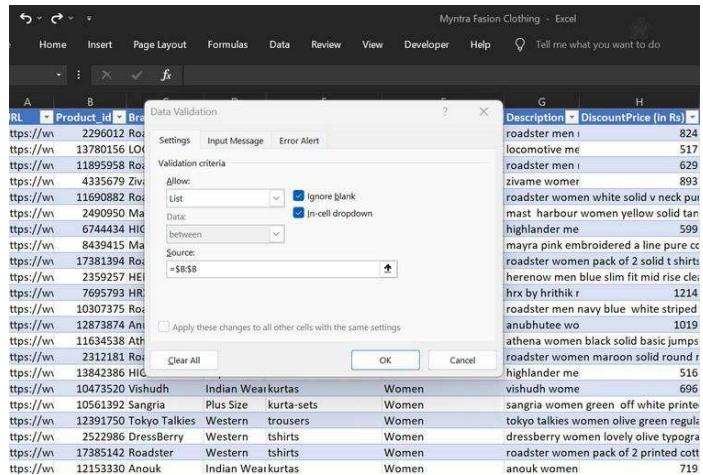
Find the "DiscountPrice" for the product with  
the Product ID "6744434"  
using the INDEX and  
MATCH functions.



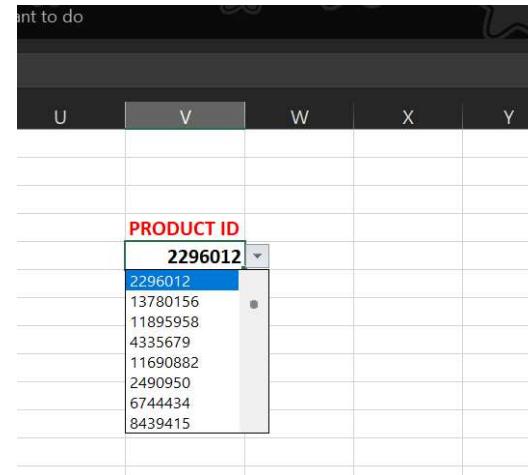
Index and Match for Product ID "6744434"	
599	

# 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.

B:B, 0), 3)				U	V	W	X	Y
Q	R	S	T					
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.

=INDEX(C:T, MATCH(V5, B:B, 0), 3)

# Thank You